CALIFORNIA Data System

Cradle-to-Career Data System First Legislative Report

December 2020

Executive Summary

Legislation passed in 2019 in California aims to link existing education, workforce, financial aid, and social service information to better equip policymakers, educators, and the public to address disparities in opportunities and improve outcomes for all students throughout the state. Over the course of 2020, in accordance with this legislation, more than 170 people from 15 state agencies and many educational institutions, research and policy organizations, and community groups have worked together to design a blueprint for the California Cradle-to-Career Data System. This report describes the planning process and outlines recommendations for the first five years of creating and implementing the data system.

Key Features

The proposed data system would be a neutral source of high-quality information, paired with professional development to help a range of stakeholders take action on this information to improve opportunities for all Californians. The public would have open access to analytical tools, including dashboards, a query builder, summaries of key student and employment outcomes, and a research library. Researchers could request access to restricted data for authorized purposes. Students and their families could use a suite of operational tools that support college and career planning, college-eligibility monitoring, electronic transcripts, and access to financial aid and other services. A support structure that includes outreach, professional development, and clear feedback loops with intended users would build the capacity of policymakers, educators, parents, and students to make better-informed decisions. The system would leverage a suite of standardized legal agreements, a modern technical architecture managed by a tech-savvy agency, a participatory governance structure,

and transparent policies to give the public confidence that individual privacy will be protected.

Types of Information

The workgroup recommends a five-year timeframe for building out the data system — referred to as ***phase one***. Initially, the system would link existing K–12, public postsecondary, employment, and financial aid data. Researchers and data providers worked together to identify more than 160 data points that relate to critical milestones in student progress, the effectiveness of financial aid, and living-wage attainment. The initial data set would also include information on participation in apprenticeships, state-subsidized early learning and care, and some social service programs. These data points would pave the way for expanded information on teacher credentialing, early learning and care, private and independent colleges, workforce training programs, and health and social services, to be included by the end of phase one.

Proposed Structure

Information from each data provider would be stored in the cloud in a secure repository. The core data set used for the dashboards and query builder would be kept in a centralized database, and other information would be linked for approved purposes. To support the ***analytical tools***, each data provider would upload a subset of their information once per year. To help safeguard individual privacy, records for individuals would be matched using a variety of data points, and unique identifiers would be assigned for each data pull. Information for the ***operational tools***, such as student planning tools and electronic transcripts, would be kept separate from information used for analytical purposes and would be refreshed on a more frequent basis than the analytical data set, to support timely student advising.

Impact

The current health and fiscal crises highlight the urgency of ensuring that the Cradle-to-Career Data System provides clear and early benefits. The proposed scope of phase one represents an important initial investment to create greater resilience and responsiveness by providing valuable, actionable information to a range of stakeholders, for example:

Students and Families

- Evaluate interests and goals for college planning, understand financial aid options
- Streamline college applications and sharing of transcripts, receive proper credit for coursework completed, alert colleges of needed food and housing supports
- Share competencies when returning to college to expand skills
- Assess the long-term benefits of early learning and care
- Understand trends in education and living-wage attainment

High School Counselors

- Identify students who are not on track for four-year institutions
- Support students who have not completed applications for college or financial aid
- Understand employment outcomes of specific education pathways
- Better support students who are first in their family to attend college

Education Leaders

- Understand trends in early learning and care, education, and living-wage attainment to support local improvement efforts and advocate for changes to local and state policy
- Accelerate access to information on prospective students
- Strengthen social service supports for students
- Identify strategies for closing equity gaps

Social Service Providers

- Understand how social supports impact education and living-wage attainment

State Agencies

- Understand the impact of interventions or develop models for new policies
- Determine student outcomes in the next phase of education or work
- Determine eligibility for social service, health, and financial aid support
- Identify strategies for closing equity gaps

Advocates and Policymakers

- Identify outcome and equity gaps to support improvement and reform efforts and secure additional resources

Researchers

- Expand access to data to track longer-term outcomes and the relationships between education, social and health services, and living-wage attainment

Governance

The Cradle-to-Career Data System would be governed by a board made up of representatives of state agencies and stakeholders who use the information. Two-thirds of the governing board's seats would be apportioned to entities providing data, and one-third to stakeholders appointed by the Governor and Legislature. The governing board would set the strategic direction and ensure that the system is supporting the state's goals. Two advisory boards would also be created, one to ensure that the system includes actionable data, and the other to ensure that the intended audiences are aware of the data system and know how to use the information it provides.

Management

The governing board would also oversee the managing entity, which would be a new program within the state's Government Operations Agency (GovOps). Consistent with its mission of supporting the work of other state agencies, GovOps would provide the services and technical expertise necessary for the data system. It would incubate the Cradle-to-Career Data System for the first five years, after which the managing entity structure would be reassessed by the governing board.

The managing entity would be responsible for implementing the data system, including assisting data providers to upload information, linking data, overseeing the development of the dashboards and query builder, producing summaries of key student and employment outcomes, helping stakeholders access and understand information, and scaling the operational tools that provide information about specific students. Although it would not produce its own research, the managing entity would play a key role in the data request process.

Accessing Data

To streamline access to information not available in the dashboards and query builder, requestors could use an expedited process to receive summary data that would not reveal individual identities. Those needing access to student-level information would follow a comprehensive review process that would include protections for human subjects, legal agreements that uphold state and federal laws, and expert review to ensure that individuals cannot be identified in resulting reports. The managing entity would support requestors in developing their proposals and would ensure that they use the most efficient means to answer their questions. The relevant data providers would

approve requests, and the managing entity would create all data sets. The managing entity would also reduce the timeframe for data access by using data-sharing agreement templates that are based on laws regarding allowable data access.

Timeline and Costs

The workgroup recommends investing in a proof-of-concept in the first half of 2021, followed by a five-year implementation process to reduce the cost of building and maintaining the system. The estimated budget for the proof-of-concept and year one (fiscal year 2021–22) deliverables is between $15 million and $20 million.

Proof-of-concept deliverables: Produce a proof-of-concept dashboard, expand access to college planning tools in low-income regions, upgrade K–12 data infrastructure for college eligibility, electronic transcripts, and application tools

Year one deliverables: Establish governance and staff, secure technology solutions, create initial analytical data set (focused primarily on K–12, public postsecondary, financial aid, and employment information), release summaries of student and employment outcomes, design dashboards and query builder interface

Year two deliverables: Launch the dashboards and query builder, train the public on using data tools, commence fulfilling data requests, expand analytical data set (teacher credentialing), expand access to college planning tools, upgrade electronic transcript infrastructure for competency-based education and social service eligibility tools

Year three deliverables: Expand analytical data set (independent and out-of-state colleges), provide electronic transcripts for all public colleges, expand access to college planning tools

Year four deliverables: Expand analytical data set (private colleges and early learning and care), provide electronic transcripts for all private and independent colleges, expand content of college planning tools

Year five deliverables: Expand analytical data set (social service, health, and workforce information), finish scaling college planning and electronic transcript tools, plan for phase two

Completing the Planning Process

With many key recommendations now in place regarding the purpose, scope, and structure of the data system, the planning process in the first half of 2021 will focus on developing policies and templates that relate to privacy and security; documenting the process for entering, correcting, and assuring the quality of data; establishing shared data definitions; refining plans for integrating early learning and care data; and crafting a structure for professional development. Information on work to date on these topics is outlined in this report and will be addressed more fully in a second legislative report that will be submitted in June 2021.

Call to Action

The intensive and thoughtful nature of the partner entities' participation in the planning process reflects their intention to leverage existing data resources, work together, and bring a large volume of actionable information into the public domain. Furthermore, the recommendations in this report reflect significant public participation that helped to forge a consensus about the scope and structure of the data system. Given the current challenges facing the state, which clarify the difficulty of using disconnected data systems to understand and respond to public health, economic, and social justice crises, the participants in this planning process urge the Legislature to endorse the creation of a stronger data ecosystem for all Californians.

CALIFORNIA Data System

Table of Contents

Cradle-to-Career Data System First Legislative Report i

Executive Summary i

Section 1: Introduction and Overview 1

Section 2: The Planning Process 2

Section 3: Purpose and Available Data 8

Section 4: Technical Structure 14

Section 5: Governance 23

Section 6: Linking and Accessing Data 30

Section 7: Privacy, Security, and Data Definitions 34

Section 8: Professional Development, Technical Assistance, and Community Engagement 39

Section 9: Implementation 41

Conclusion 51

Appendix 53

CALIFORNIA Data System

Section 1: Introduction and Overview

In 2019, the California State Legislature passed the California Cradle-to-Career Data System Act (the Act),[1] which sets out requirements for the development of a statewide data infrastructure, also known as a statewide longitudinal data system (SLDS). The Act requires that the data system ensure that existing education, workforce, financial aid, and social service information is fully leveraged to address disparities in opportunities and to improve outcomes for all Californians, from cradle to career.

Pursuant to the Act, in November 2019, the California Governor's Office launched a planning process with support from WestEd. Over the course of 2020, more than 170 people participated in a comprehensive and inclusive design process, drawn from the 16 partner entities indicated in the legislation, as well as educational institutions, research organizations, data experts, and advocacy groups. When the design process ends in June 2021, these experts will have produced detailed plans for the first five years of the state's Cradle-to-Career Data System (phase one).

The proposal that has been developed thus far through this process has drawn acclaim from national experts because it addresses challenges that have hindered other SLDSs, including:[2]

- Contextualizing education data — by integrating financial aid, social service, and employment information — to better identify structural barriers that may inhibit access and success

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

[2] See a report from Alvarez and Marsal in the November 2020 Technology & Security Subcommittee materials at https://cadatasystem.wested.org/meeting-information/technology-security-subcommittee and an assessment by the Data Quality Campaign at https://dataqualitycampaign.org/california-is-following-best-practices-to-make-its-statewide-data-system-a-reality/.

- Providing the public with access to a significant amount of information through dashboards and a query builder, paired with a streamlined data request process that aligns with state and federal privacy laws
- Addressing researchers' and policymakers' needs for access to anonymized data on groups of people over time, as well as educators' needs for highly secure information on individual students
- Including professional development to ensure that intended audiences know about and can utilize the information
- Guaranteeing a significant voice in the governance structure for members of the public to provide feedback and help design new solutions that support the state's goals for evidence-based decision-making and a more equitable future

This document, the first of two legislative reports,[3] describes the planning process and recommendations for phase one of the Cradle-to-Career Data System. After a section on the planning process are sections on the recommendations regarding the data system purpose and available data; technical structure; governance; linking and accessing data; privacy, security, and data definitions; professional development, technical assistance, and community engagement; and implementation.

Section 2: The Planning Process

Background Research

The Act requests that the planning process include reviewing and building on prior reports, existing data systems, and best practices.[4] A WestEd facilitation team compiled an evidence basis for planning in a number of ways.

First, WestEd compiled information on California-specific issues by conducting an analysis of previous planning efforts in the state and reviewing policy papers from a

[3] In March 2020, the Office of Planning and Research requested that the statutory reporting timeline be adjusted to account for an initial delay. Given that the planning process did not begin until November 2019, the legislature agreed with deferring the preliminary report of the workgroup until December 31, 2020. A progress report will be provided on April 1, 2021, and the final report on June 30, 2021.

[4] http://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml?lawCode=EDC&sectionNum=10856

number of sources. WestEd also interviewed all partner entities that are named in the legislation and surveyed stakeholders.[5]

Then, WestEd sought national perspectives by conducting interviews and reviewing documentation regarding effective practices for SLDSs. For example, the facilitation team interviewed SLDS staff in Colorado, Georgia, Hawaii, Kentucky, Maryland, Minnesota, Montana, North Carolina, Pennsylvania, Utah, and Washington, and hosted meetings of representatives from four states to describe their data systems and governance processes.[6] WestEd also worked with the following organizations, which were on hand to answer questions throughout the process: the U.S. Department of Education's Privacy Technical Assistance Center;[7] the National Center for Education Statistics State Support Team for the Statewide Longitudinal Data System Grant Program;[8] and the Data Quality Campaign.[9]

Finally, WestEd gathered information that was specific to each stage of the planning process. For example, WestEd arranged for demonstrations of existing tools, some from other states and some from California. The demonstrations were recorded and posted on a project website.[10] For many meetings during the Cradle-to-Career Data System planning process, WestEd wrote background papers that describe effective practices and provide examples specific to the topic at hand.[11]

---

[5] The partner entities include the Association of Independent California Colleges and Universities, Bureau for Private Postsecondary Education, California Community Colleges, California Department of Education, California Department of Social Services, California Department of Technology, California Labor and Workforce Development Agency, California Health and Human Services Agency, California School Information Services, California State University, California Student Aid Commission, California State Board of Education, Commission on Teacher Credentialing, Employment Development Department, and University of California.

[6] https://nces.ed.gov/programs/slds/stateinfo.asp

[7] https://studentprivacy.ed.gov/

[8] https://nces.ed.gov/programs/slds/

[9] https://dataqualitycampaign.org/

[10] https://cadatasystem.wested.org/meeting-information/potential-data-tools-webinars

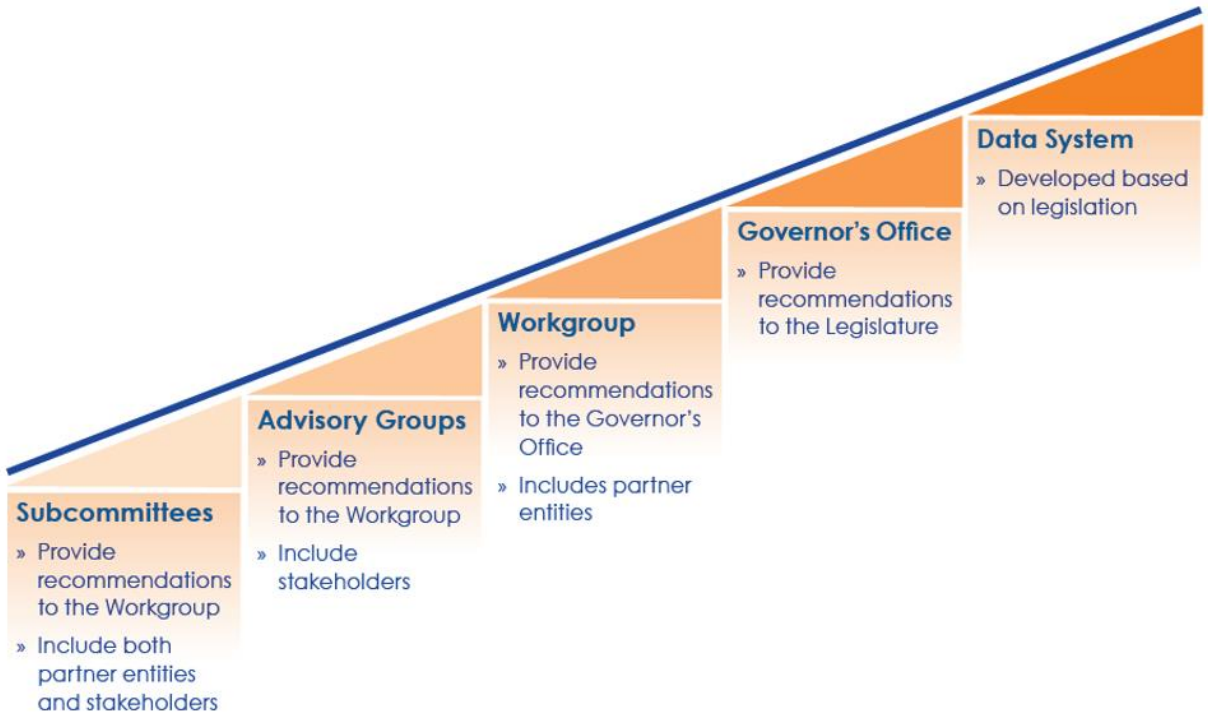
[11] For a full list of reports, see https://cadatasystem.wested.org/resources-page.

Read an analysis of prior state efforts, recommendations from policy organizations, and background papers that informed each meeting: https://cadatasystem.wested.org/resources-page.

The Planning Structure

WestEd's research into prior efforts in California and other states found that a multilayered planning process has tended to work well. In addition, the Act requires a broad range of recommendations that require different types of expertise. Therefore, the WestEd facilitation team implemented a multitiered planning structure (Figure One).

Figure One: Cradle-to-Career Data System Planning Structure



The Cradle-to-Career Workgroup named in the legislation consists of representatives from 16 partner entities. It is informed by two advisory groups that represent intended users of the data system and by six subcommittees that focus on different domains of expertise: common identifiers, community engagement, data definitions, legal, research, and technology and security (see the "Cradle-to-Career Planning Structure" box).

Most subcommittees include representatives from research, technology, and advocacy organizations, in addition to partner entities. In 2020, the subcommittees met monthly to develop recommendations, which were reviewed and voted on by the workgroup. Each quarter, the two advisory groups reviewed the recommendations of the workgroup and provided feedback. One advisory group focuses on policy and analytics — the types of data used when conducting research. The other advisory group focuses on practice and operations — the types of information used when supporting individual students. On any occasion when the advisory groups expressed concern about workgroup recommendations, a homework team made up of advisory and workgroup members worked together to craft a compromise, which was reviewed and adopted at a subsequent workgroup meeting.[12]

[12] See the appendix for a full list of all workgroup, advisory group, and subcommittee members.

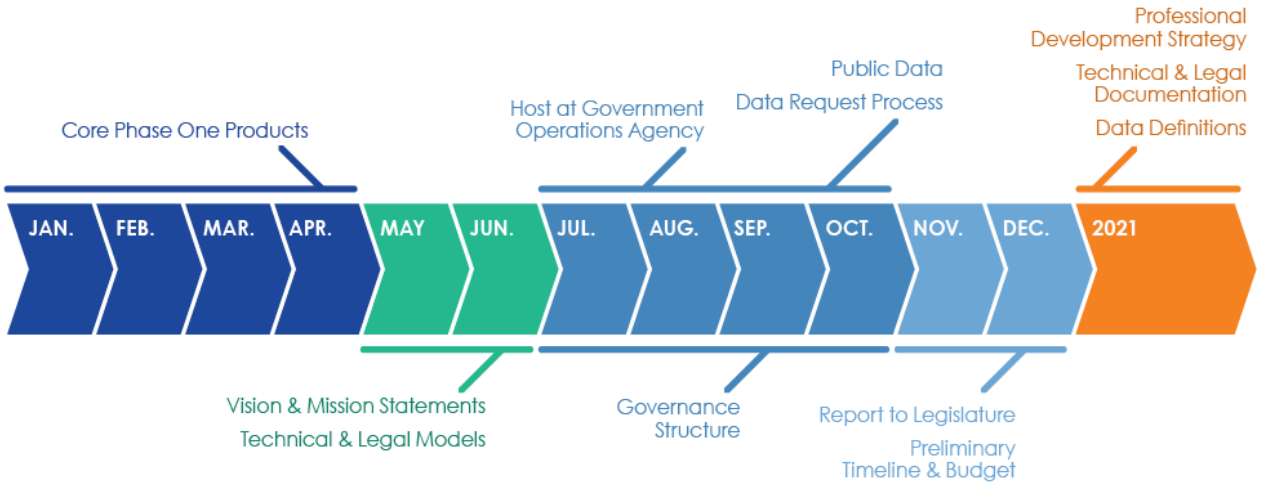
Cradle-to-Career Planning Structure

- **Workgroup:** Provides recommendations to the Governor's Office about data system development
- **Policy and Analytics Advisory Group:** Advises the workgroup on how the data system could address research, evaluation, public accountability, and optimization of publicly funded services at the system level
- **Practice and Operations Advisory Group:** Advises the workgroup on identifying how the data system could address improvement efforts at the institutional and regional levels by supporting a case management approach to service delivery and creating tools that would be useful to students, families, and teachers
- **Common Identifier Subcommittee:** Advises the workgroup about the design of the technical process that will be used to link student records across partner entities
- **Community Engagement Subcommittee** (launching in 2021): Advises the workgroup about how to create mechanisms for strong feedback loops with data users, supporting evidence-based decision-making and analytical capacity, and how to ensure equitable access to actionable information
- **Data Definitions Subcommittee:** Advises the workgroup on the documentation of technical definitions for key information that will be shared in phase one
- **Legal Subcommittee:** Advises the workgroup about the contracts, policies, and legal agreements necessary to support data sharing and privacy
- **Research Agenda Subcommittee:** Advises the workgroup on the six priority topics identified by the Cradle-to-Career Data System Act
- **Technology and Security Subcommittee:** Develops the technology specification requirements to address data structures and privacy considerations

Planning Timeline

Following work conducted in November–December 2019 to develop the facilitation process and determine workgroup, advisory group, and subcommittee members, activities have proceeded in five stages (Figure Two).

Figure Two: Cradle-to-Career Data System Planning Milestones



January–April 2020: Identify Core Phase One Products

By examining features of data systems in other states, reviewing calls to action for a longitudinal data system in California, and investigating effective projects already underway in the state, the various committees were able to identify a scope of work that addresses the state's short- and long-term needs. Although some elements of the system may take years to develop, identifying the types of functionality that will be required provides the data system with flexibility to meet future needs.

May–June 2020: Develop Data System Framework

Through crafting vision and mission statements and agreeing to foundational technical and legal models, the workgroup created a framework to evaluate subsequent governance, technical, and operational proposals, and to ensure that they align with the purposes and scope of the data system.

July–October 2020: Design Governance Process

The workgroup recommended where the data system should be housed, developed the governance structure, determined which information would be in the public domain, and established the mechanism for requesting access to restricted data. All recommendations were based on examining practices in other states and were informed by national and state experts. Several of the proposals were revised to reflect input from advisory group members.

November–December 2020: Outline Timeline and Budget

The workgroup established a preliminary timeline, budget parameters, and the content for this report to the Legislature.

January–June 2021: Create Implementation Documentation

The workgroup will develop plans for community engagement and professional development, with support from an additional subcommittee. In addition, existing subcommittees will continue to develop documentation necessary to launch the data system, such as a dictionary for all publicly available information included in the system, technical specifications, and legal templates.

View a calendar with meeting topics: https://cadatasystem.wested.org/meeting-information/calendar-for-all-meeting-dates.

Section 3: Purpose and Available Data

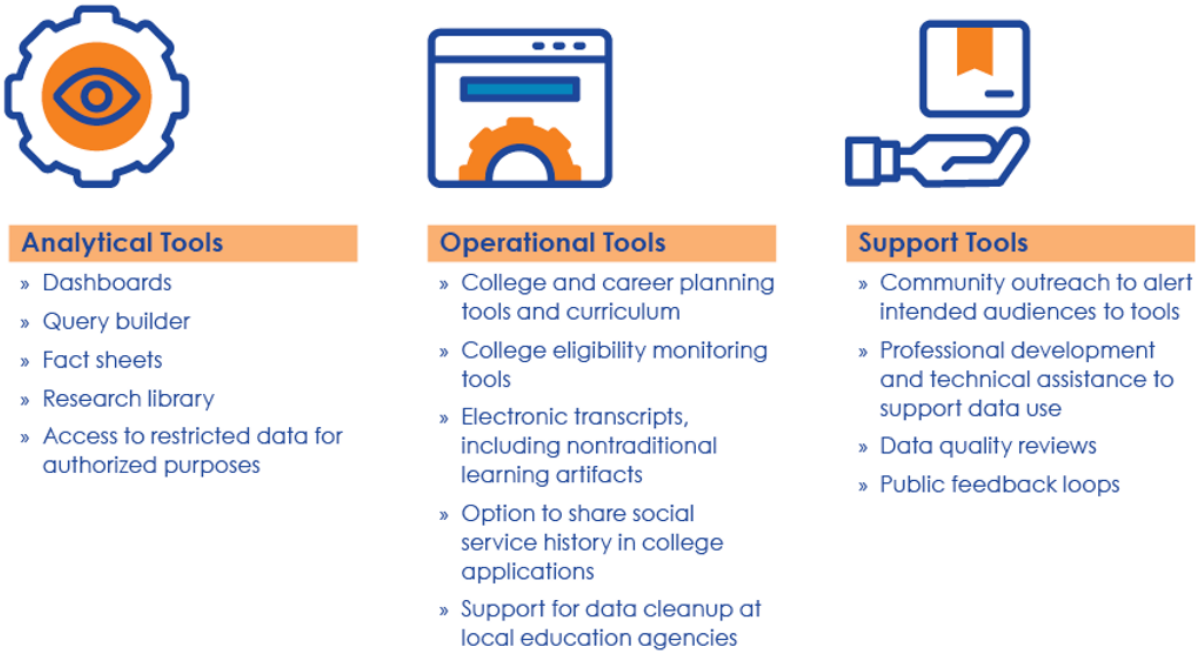
Purpose of the Data System

One of the first challenges for the planning process was to determine what purpose the data system would serve. In interviews and in the literature reviewed by WestEd at the start of the planning process, a wide range of scenarios were proposed — from the system being a research data set focused on establishing educational policies to its being a Yelp-like app for finding nearby childcare. After evaluating many potential

products for the data system, the workgroup identified a shared set of priorities for phase one of the data system, which would last roughly five years.

The workgroup recommends that the state create a neutral entity that provides high-quality information for a range of analytical and operational purposes (Figure Three). This concept is spelled out in the following vision and mission statements, and in a list of strategic priorities.

Figure Three: Cradle-to-Career Data System Core Components



Vision

The Cradle-to-Career Data System connects individuals and organizations with trusted information and resources. It provides insights into critical milestones in the pipeline from early care to K–12 to higher education, skills training, and employment. It empowers individuals to reach their full potential and fosters evidence-based decision-making to help California build a more equitable future.

Mission

To be California's source of actionable data and research on education, economic, and health outcomes for individuals, families, and communities; to expand access to tools and services to navigate the education-to-employment pipeline.

Strategic Objectives for Phase One

- Develop the architecture for linking records across agencies and creating intersegmental data sets
- Provide public-facing data visualizations, query tools, and a research library that provide actionable information on education, social services, employment patterns, and equity gaps in opportunities and outcomes
- Provide interagency data sets that enable research on factors that help Californians meet critical education milestones, evaluate the long-term impact of state-funded programs, and identify strategies for closing equity gaps
- Provide resources, training, and technical assistance that build data literacy among policymakers, practitioners, and the public
- Provide college and career planning tools, college-readiness monitoring, electronic transcripts, and confirmation of eligibility for financial aid and student supports
- Lead efforts to ensure the reliability of data contributed by the partner entities

The scope of phase one reflects a theory of action that assumes the following:

- A wide variety of Californians need access to information to support decision-making, and these stakeholders need different types of information.
- Although linking information is valuable, there will be a stronger return on investment if the data are made available to the public and if professional development is offered to help stakeholders learn about, understand, and use the information.
- By investing in a modern technology infrastructure and assuring data quality, California will be better equipped to provide useful insights for unexpected challenges while strengthening existing agency data systems.

(To see examples of how information from the data system would support this theory of action, see the "Potential Impacts of the Data System" box.)

Potential Impacts of the Data System

This list of potential impacts of the Cradle-to-Career Data System provides a sample of the many ideas identified by partner entities, advisory group members, and community members during the planning process.

- Inform state priorities by documenting the various ways that individuals transition from school or training to work and living-wage jobs over time
- Target recovery efforts from the pandemic by examining college-going, transfer, and employment patterns, and how they vary by student type, such as by race or economic status
- Support educators to develop equity strategies that contextualize education outcomes with factors such as prior course-taking patterns, school resources, and access to financial aid and social services
- Create professional development opportunities that cross-pollinate effective practices for improving student outcomes and closing equity gaps
- Develop strategies to address impending teacher shortages by identifying trends in teacher retention, preparation, diversity, and mobility
- Improve student retention and engagement by identifying those in need of food, healthcare, and housing supports
- Provide tools to support college planning, targeting interventions to ensure eligibility, and identifying progress on applications, in order to reverse the trend of fewer students applying for college and financial aid
- Reduce unnecessary hurdles for students by increasing accuracy and reducing the difficulty of sharing transcripts
- Evaluate the expansion of apprenticeships by documenting where they fit into school-to-work pathways, how they lead to living wages, and who benefits from these programs
- Improve accuracy of course-taking information in state data systems by having educators examine and correct data at the local level
- Build public trust by implementing strong and consistent privacy protections and security protocols
- Provide information to better understand and respond to unanticipated needs and crises

Proposed Information for Phase One

In naming a broad range of partner entities in the Act, California signaled its intent to build a data system that goes beyond patterns of college enrollment for K–12 students. The inclusion of early learning and care, employment, financial aid, workforce development, social service, health, and teacher credentialing data providers is in alignment with other states' efforts to link these types of information with education data. However, the Act provides flexibility regarding which data sets to link first.

The workgroup recommends focusing phase one of the data system on the six key topics that were identified in the legislation, which requires linking early care and learning, K–12, postsecondary, employment, and financial aid data (see the "Priority Research Topics" box).

For part of phase one, the workgroup also recommends including in the system a few data points from other partner entities, namely data on whether students received public assistance or subsidized food and medical care, are in the foster care system, and/or participated in pre-apprenticeship or apprenticeship programs.

These data points are of high value for understanding education outcomes, particularly given the disruptions of the pandemic and given an increased awareness that structural factors shape education and economic outcomes. By contrast, Figure Four is a visualization of Oregon's longitudinal data system which is limited in that it includes *only* education outcomes. To better clarify why some Californians are more likely to attain a postsecondary credential than others, the workgroup recommends that the data system also indicate whether students are struggling with their basic needs, such as inadequate housing and food or are pursuing alternative routes to living-wage jobs.

Furthermore, by including a few targeted non-education data points, potential technical and legal challenges to incorporating multiple kinds of data can be identified early on. Social service and health data are governed by different legal frameworks than education data, which require careful review to determine allowable use. Furthermore, workforce data are stored in multiple databases, utilizing different underlying definitions for participation and demographics, which can complicate efforts to combine data sets.

Priority Research Topics

The Act identifies six topics on which the data system should inform policy:

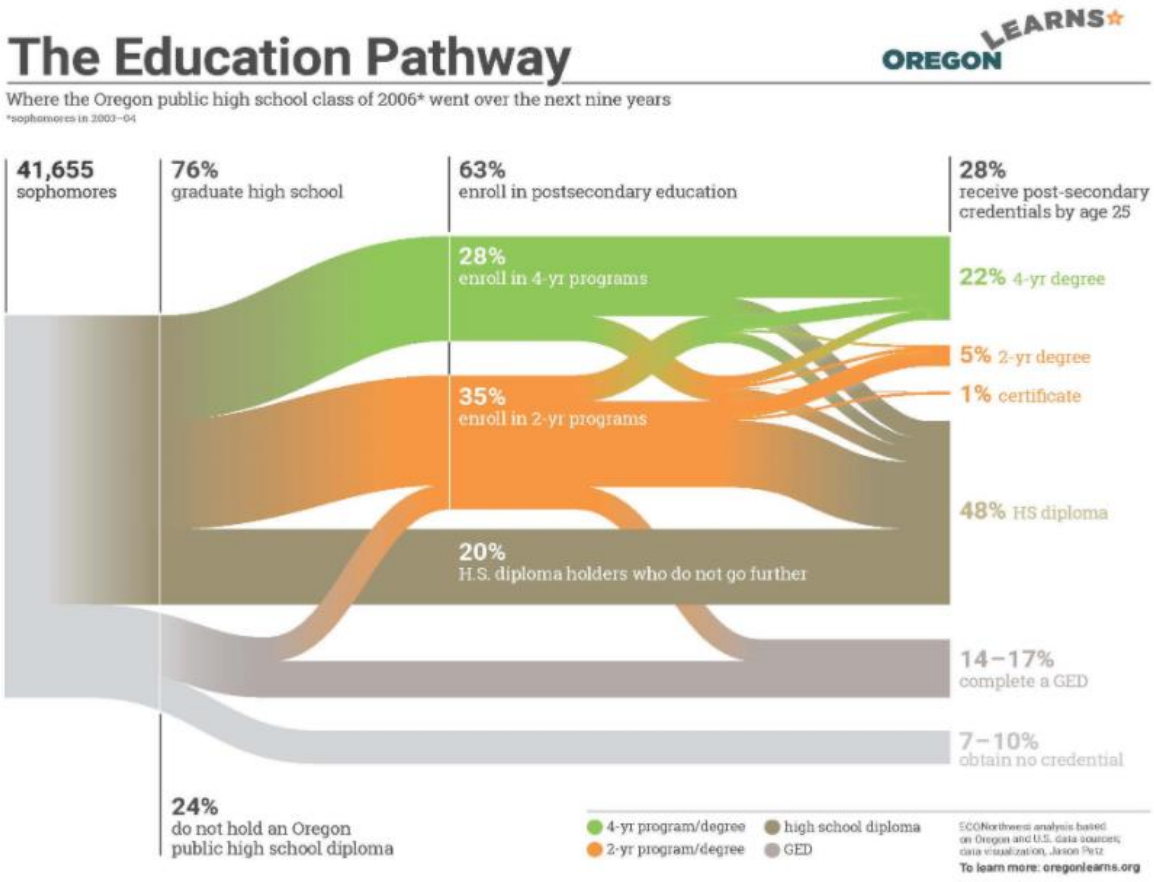
- The impact of early education on student success and achievement as a student progresses through education segments and the workforce
- The long-term effect of state intervention programs and targeted resource allocations in primary education
- How prepared high school pupils are to succeed in college
- How long it takes students who transfer from community college to the University of California, the California State University, or another four-year postsecondary educational institution to graduate with a baccalaureate degree
- College access, completion, and long-term effects of access to state financial aid
- The workforce effect of graduation from high school, community college, and four-year postsecondary educational institutions

For phase one, the workgroup recommends that the focus on time-to-baccalaureate degree be expanded to examine a broader range of issues related to transfer support and post-transfer milestones.

During spring 2021, the workgroup will aim to align data points related to early learning and care in accordance with the recently released Master Plan for Early Learning and Care.

Because there is insufficient information to link California Department of Education records with data from the Employment Development Department, the workgroup recommends that a future phase of the data system examine career outcomes for students who do not enroll in college.

Figure Four: Lack of Contextual Data Can Limit Understanding of Long-Term Outcomes



Section 4: Technical Structure

Analytical Tools

For the analytical tools that would allow users to access the data, the workgroup recommends using a hybrid data architecture. Hybrid data architecture refers to the method whereby some information would be linked and placed into an online repository, which the workgroup defines as the P20W data set.[13] Other information would be linked only when authorized by legal agreements that ensure allowable use. Once per year, each data provider would upload a subset of their information into

[13] P20W refers to linking records from Pre-K through graduate school (20th grade) to workforce information.

individualized cloud-based repositories associated with the Cradle-to-Career Data System, using a secure transfer mechanism. Records for individuals would be matched within the cloud environment,[14] and the managing entity would be responsible for creating all data sets (Figure Five).

The system would be designed to optimize security and privacy. Each partner entity would designate who within their organization has access to the information in the cloud and would tag the uploaded information by using a data classification scheme to indicate how sensitive each data point is. Information that is released through public dashboards and query tools would be subject to a de-identification protocol developed by the partner entities in compliance with the federal statutes for the Family Educational Rights and Privacy Act (FERPA),[15] the Health Insurance Portability and Accountability Act (HIPAA),[16] and California privacy laws, such as the Information Practices Act.[17] When third parties access restricted data sets, such as for evaluating the outcomes of a program, which requires matching specific student records, they would conduct their analysis in a secure data enclave, similar to the system used by the Texas SLDS. Researchers would not be able to remove from the system any information that could reveal an individual's identity and would sign data-sharing agreements that spell out data security provisions.

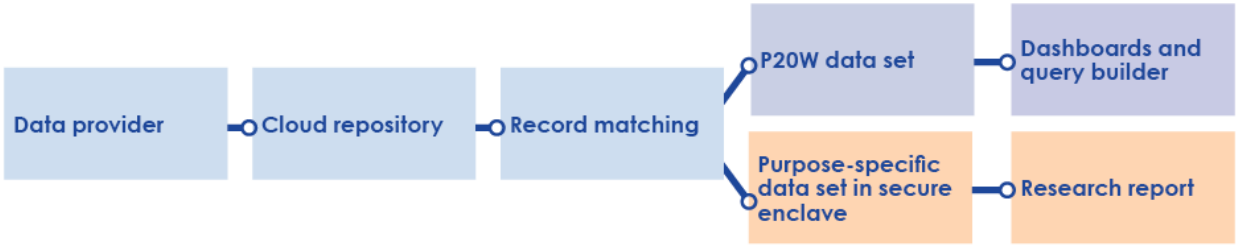
[14] The State of California requires that state departments shift toward cloud computing solutions for all information technology projects, in accordance with the State Administrative Manual, Sections 4983 and 4983.1. For more information, see the California Department of Technology's Technology Letter: https://cdt.ca.gov/wp-content/uploads/2017/03/TL-14-04-Cloud-Computing-Policy.pdf.

[15] https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html#:~:text=The%20Family%20Educational%20Rights%20and,privacy%20of%20student%20education%20records.&text=Parents%20or%20eligible%20students%20have%20the%20right%20to%20request%20that,to%20be%20inaccura

[16] https://www.cdc.gov/phlp/publications/topic/hipaa.html#:~:text=The%20Health%20Insurance%20Portability%20and,the%20patient's%20consent%20or%20knowledge

[17] https://leginfo.legislature.ca.gov/faces/codes_displayexpandedbranch.xhtml?tocCode=CIV&division=3.&title=1.8.&part=4.&chapter=1.&article=

Figure Five: Cradle-to-Career Data Flow for Analytical Tools



Data for Analytical Purposes

The Research Agenda Subcommittee, consisting of representatives from partner entities and research organizations that have experience working with intersegmental data, used the six priority research topics to develop a list of more than 160 data points that would be included in phase one of the data system. For each of the topics, WestEd provided a background paper that summarized the information that is available in other state data systems or has been utilized in prominent research studies. Subcommittee members reviewed the background papers, identified relevant questions for California, and specified likely actions that various stakeholders could take with that information. The stakeholders they considered included students and families, teachers and counselors, administrators and policymakers, and advocates and researchers. Finally, subcommittee members listed the more than 160 data points that would be necessary to provide the desired information, and they recommended that California purchase data from the National Student Clearinghouse[18] to provide information on students who enroll in colleges in other states.

This list of data points was shared with each of the data providers, which reviewed the suggested data points to clarify whether the providers collect this information and if the data are of sufficient quality to place into the public domain. In October 2020, the workgroup approved the list of data points to include in phase one of the California Cradle-to-Career Data System. Additional work during spring 2021 will help to align the early learning and care data points with the Master Plan for Early Learning and Care.

Members of the Research Agenda Subcommittee and the two advisory groups identified other data points, in addition to the list of approved P20W data points, that they would like to see included in future phases of the data system (see the "Examples

[18] https://www.studentclearinghouse.org/

of Desired Data Points for Future Phases of the Data System" box). Most of these other data points are not currently collected at the state level and thus would require an investment of time and resources to gather. However, because these data points frequently address the context for learning, or new issues that have emerged in the wake of the pandemic, the state might benefit from a systemic review of data points that would be of high value for policy and implementation purposes.

Examples of Desired Data Points for Future Phases of the Data System

- Adverse Childhood Event (ACES) scores
- Familial relationships
- Children in private early care facilities
- Participation in industry-based high school academies
- Work-based learning, such as internships
- State licensure information
- More nuanced information on online and hybrid courses
- Occupation
- Number of hours worked per fiscal quarter

Accessing Analytical Data

The workgroup recommends that information pertaining to the six priority research topics be accessed through four mechanisms or tools (Figure Six). Most of the data provided through these means would be at the summary level — that is, only information on groups of students would be presented, to protect individual privacy.

- **Dashboards:** infographics that provide users with the ability to drill down to more complex charts, including the ability to see comparisons and disaggregated results by student characteristics, institutions, geographies, and timeframes, plus the ability to export files with summary data
- **Query builder:** a tool that allows users to drag and drop data points to create simple queries, with the ability to export spreadsheets and charts with summary data

- **Reports:** a searchable research library that allows users to find summaries of key student and employment outcomes as well as research reports developed using Cradle-to-Career data
- **Data requests:** processes for requesting information not available through the dashboards or query builder, including an expedited review of summary data or, if information is requested at the level of individual students, a comprehensive review and approval process to ensure individuals' privacy will be protected in compliance with all applicable laws

Figure Six: Accessing Cradle-to-Career Data

The workgroup recommends providing information in a variety of formats, to best meet the needs of various constituencies. For students and families, an infographic format could be useful for clarifying existing patterns of education attainment and employment outcomes. For example, they could use the dashboards or receive a fact sheet that shows the proportion of students at their high school who enroll in postsecondary education, where they enroll, attainment of first-year milestones and awards, and time to award (see the "Proposed Dashboard Visualizations" box).

Educators, advocates, and policymakers could use the query builder to find answers to specific questions that emerge in planning and policy-development processes. For example, information could be pulled by Assembly or Senate district to document factors influencing the number of students earning bachelor's degrees in Science,

Technology, Engineering, and Math (STEM) subjects, such as the proportion of students enrolled in high school career and technical education courses such as engineering, information technology, or health, the proportion of students taking transfer-level math courses at community colleges, and the characteristics of individuals who major in STEM disciplines. Or, users could access the research library to learn about research that has already been conducted about STEM pathways.

Proposed Dashboard Visualizations

- Pathway diagram showing the full trajectory of individuals, from early childhood education through postsecondary education and beyond, including when they entered the workforce and attained a living wage
- Primary school outcomes based on participation in early learning and care, and quality of that care
- Relationship between chronic absenteeism and high school and college outcomes, in the context of counseling ratios
- Proportion of secondary students who enroll in postsecondary education, where they enroll, attainment of first-year milestones and awards, and time to award
- Proportion of community college students who become eligible for transfer, if and where they transfer, time to transfer, and baccalaureate degree attainment
- Proportion of students who apply to and enroll in college, based on whether they receive financial aid, plus the relationship of financial aid to common majors, time to award, employment industry, living-wage attainment, and debt
- Employment match rates, earnings, earnings gains, and living-wage attainment

State agencies, researchers, and evaluators could request data that would help them to better understand the impact of interventions or to develop models for new policies. For example, information from the Cradle-to-Career Data System could be leveraged to document the long-term outcomes of teacher preparation programs, such as where graduates are employed, whether they represent the demographics of the students they are teaching, and which programs train teachers who are more likely to stay in the

profession. Researchers could additionally identify characteristics of schools that are successful at retaining their teachers, particularly in high-poverty districts.

Because the federal laws that protect the privacy of an individual's data impose limitations on the use of real data, the workgroup also recommends that the state evaluate the feasibility and value of developing synthetic data sets to support future research. Rather than providing access to data on individual students, synthetic data sets substitute values found in original data sets with similar values to create files that mimic student-level data but do not contain information that could be tracked back to specific individuals.

Read a summary of the goals and content for the dashboards and query builder, user stories that describe who would use each tool, the list of data points, and additional data points desired for future phases of the data system: https://cadatasystem.wested.org/recommendations.

Operational Data

In addition to noting the statutory requirement that the Cradle-to-Career Data System transfer high school pupils' educational records to postsecondary educational institutions, the workgroup recommends the statewide adoption of tools that support more equitable access to college and that respond to the needs of adult learners.

- **College and career planning:** provide K–12 districts with tools and curricula needed to systematize guidance practices
- **College-eligibility monitoring:** allow students, parents, educators, and counselors to monitor factors that influence college-going and retention rates, such as completing eligibility requirements for four-year colleges and submitting financial aid applications
- **Electronic transcripts:** provide a consistent platform that streamlines and automates college, financial aid, and transfer-application processes for students and institutions, and expand the types of records that can be shared for adults pursuing employment-related training, such as competency-based credentials, experiential transcripts, e-portfolios, certifications, and badges

- **Eligibility:** enable students to authorize information-sharing that would qualify them for support services, such as information on whether they are in the foster care system or have received public services, including CalFresh, CalWORKS, and Medi-Cal
- **Data clean up:** identify instances of inaccurate information in local student information systems, and support educational institutions to correct this information

The workgroup determined which operational tools to include through a multistep process, including reviewing recommendations from the Practice and Operations Advisory Group and public surveys, examining research on tools implemented in other states and within California, and participating in demonstrations of possible tools. Given the directive in the Act to build upon existing data systems where appropriate, the workgroup gave a preliminary endorsement to scale the California College Guidance Initiative (CCGI)[19] and eTranscript California,[20] conditional upon further review.

To further inform this preliminary endorsement, WestEd then developed a report that includes the following:

- a review of the literature and examples from states that are using similar tools on a regional or statewide basis
- lists of practices and features considered to be most effective, as well as key implementation concerns for scaling
- a description of how CCGI and eTranscript California seek to address key goals that were established by the workgroup and are found in effective practices
- a chart summarizing the current and proposed scale for CCGI and eTranscript California
- technical charts showing how data would flow from local or state agencies through the tools

The workgroup reviewed the report and voted to recommend the scaling of CCGI and eTranscript California in phase one of the data system.[21]

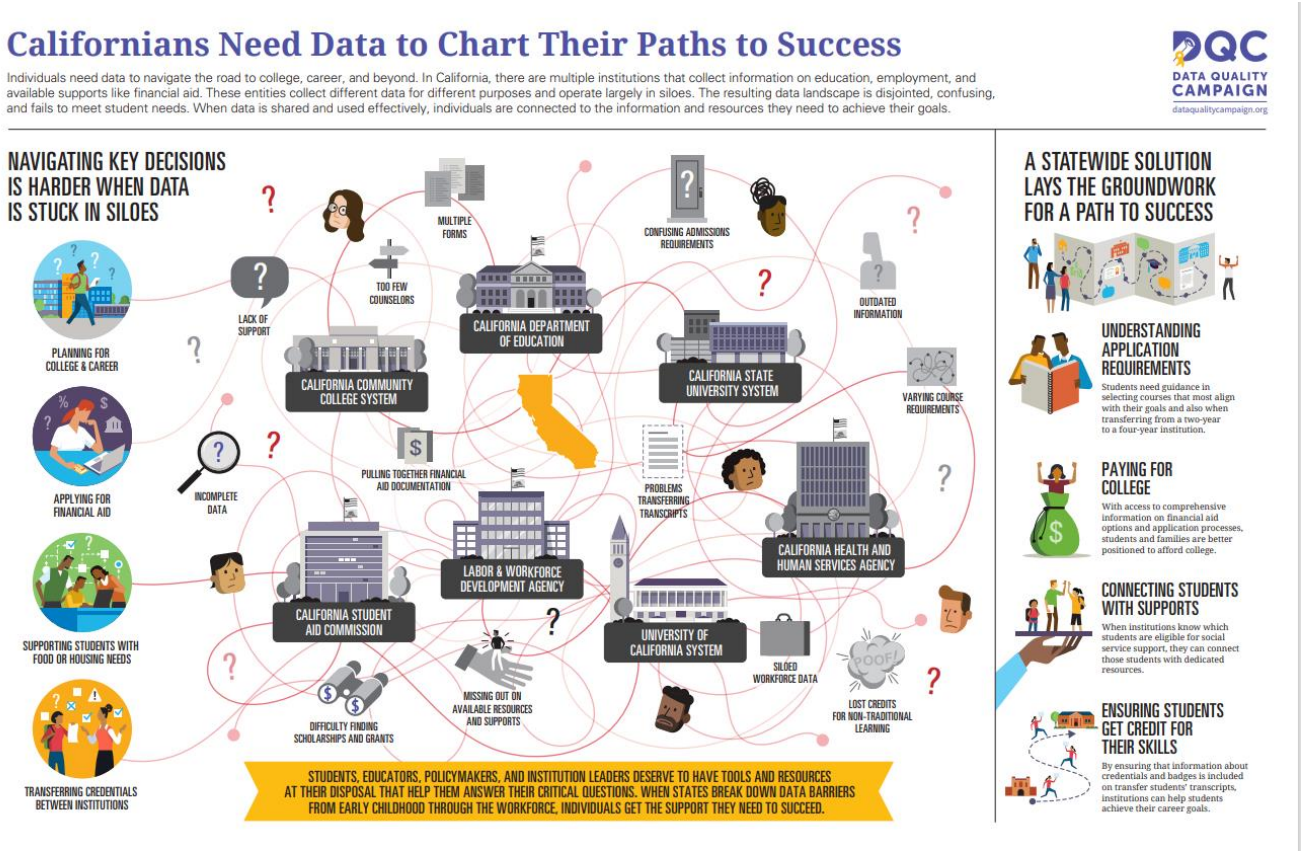
[19] Also known as CaliforniaColleges.edu: https://www.californiacolleges.edu/#/

[20] https://etranscriptca.org/

[21] The University of California plans to participate in phase two of the project, pending outcomes of phase one.

Figure Seven provides an illustration of key challenges facing Californians and how operational tools would support addressing these challenges. As this figure demonstrates, navigating key decisions, such as planning for college or a career, can be difficult when data are stuck in siloes. The workgroup recommendations for the Cradle-to-Career Data System would provide the operational tools needed by students and families to make better-informed decisions.

Figure Seven: Key Challenges and How Operational Tools Would Support Californians



Accessing Operational Data

Although making broad comparisons or understanding trends (such as knowing that 30 percent of students experience food insecurity) can be helpful, having more specific information (such as knowing which students are hungry) is more meaningful and useful for crafting on-the-ground interventions. However, examining data at this level introduces security and privacy considerations, because information can be linked to individuals. Accordingly, the workgroup recommends that the data for the operational tools be kept separate from the data for the analytical tools.

The workgroup also recommends leveraging the operational tools to improve the quality of data in state systems. For example, as CCGI works with high schools to improve the accuracy of the schools' local data regarding which courses are properly registered with the University of California (UC) and California State University (CSU) as CSU-eligible and UC-eligible "a–g" coursework, these improvements would yield higher-quality data at the state level as well. Furthermore, the proposal to scale CCGI includes a shift from collecting data from each individual local educational agency to collecting data from the California Department of Education's (CDE's) California Longitudinal Pupil Achievement Data System (CALPADS).[22] This change would streamline the process for local K–12 schools by reducing the need for additional legal agreements or numerous file uploads to multiple entities, while fostering the ability of the CDE to collect information more frequently from districts.

Similarly, the workgroup's recommendation for scaling eTranscript California includes moving the existing platform to an open-source format that would allow for the inclusion of nontraditional learning elements such as badges. Expanding the types of information that could be included on transcripts would remove a major impediment to offering credit for prior learning, which many community colleges are moving to adopt in the wake of the pandemic, particularly given the urgent need for retraining.

Read documentation on the operational tools, including a chart summarizing the scaling proposal, the proposed technical data flows for CCGI and eTranscript California, and the WestEd report: https://cadatasystem.wested.org/recommendations.

Section 5: Governance

The workgroup recommends a governance structure that emphasizes reviewing the efficacy of the data system related to a set of strategic goals. It also recommends clear structures for integrating feedback from the public (Figure Eight). This focus stands out from the governance approach of most other states' data systems, which emphasize

[22] CALPADS is a longitudinal data system used to maintain individual-level data, including student demographics, course data, discipline, assessments, staff assignments, and other data for state and federal reporting.

procedures such as establishing a shared research agenda, approving data requests, and troubleshooting technical implementation.

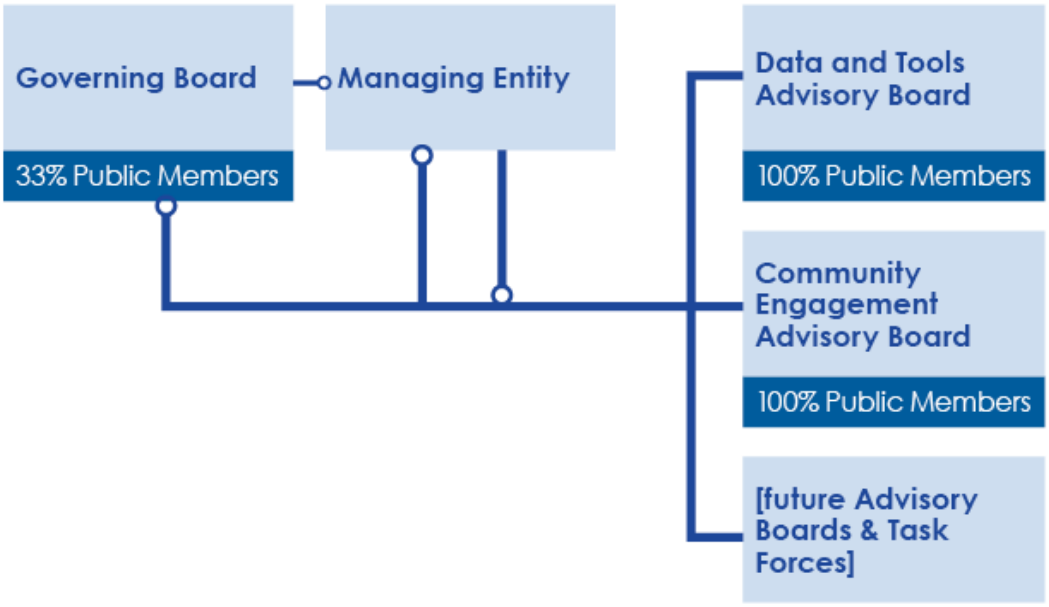
The workgroup developed the proposed governing board structure in partnership with advisory group members, with a focus on addressing two priorities: securing a sufficient vote for the public and ensuring that appropriate data will be shared.

The proposed approach, which reserves one-third of governing board seats for the public and provides them with the same voting rights as data providers, would be the most inclusive state data system governing board in the country. The proposal is also unique because it creates a structured flow of information on access to actionable information through the two advisory boards and it tasks one advisory board with examining trends in denied data requests.

However, when the number of seats on the governing board was adjusted later in the planning process to provide expertise on workforce training programs, members of the Policy and Analytics Advisory Group and Practice and Operations Advisory Group requested further conversation about the balance of the votes.

The workgroup believes that the continuous improvement mechanisms that are built into the governance proposal will provide an opportunity for data providers and public representatives to work together to ensure that the data system addresses the public's need for data while ensuring that information is of high quality and safeguards individual privacy. Additional conversations in early 2021 will seek to resolve remaining concerns.

Figure Eight: Relationship Between the Governing Board, Managing Entity, and Advisory Boards



Governing Board

The proposed 18-member governing board for overseeing the Cradle-to-Career Data System represents the entities that would provide the information necessary to compile a nuanced understanding of progress from cradle to career, with seats allocated based on the entities that provide information (Figure Nine). For example, information on children in early learning and care programs would be shepherded by representatives from both the CDE and the California Department of Social Services. The Association of Independent California Colleges and Universities would hold a seat to facilitate the inclusion of data from independent colleges, separate from a seat for the Bureau for Private Postsecondary Education, which will be developing its own data system in tandem with the Cradle-to-Career Data System.[23] Also, the California Labor and Workforce Development Agency and the California Health and Human Services Agency would hold seats to support planning for expanding workforce, health, and social service data.

[23] https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201920200AB1340

In addition to the seats for data providers, a third of the seats on the governing board would be reserved for members of the public, with the goal of ensuring that the data system aligns with the vision and mission.

Because K–12 education is represented by only one agency (the CDE), whereas higher education, health and social services, and workforce segments have more than one representative on the governing board, the Governor would appoint two K–12 seats for members of the public — one for a teacher and one for an education leader — in addition to the CDE representative, providing a total of three K–12 seats.

Four other public seats would round out the one-third of the board's seats to be held by community members. Specifically, the Assembly would appoint one of these seats, the Senate would appoint one seat, and the Governor would appoint two more of the public seats (in addition to the two K–12 public seats also appointed by the Governor). These public seats would be allocated to constituencies that are intended to benefit from the system and/or are affected by the data, such as practitioners, families, students, adult learners and workers, community organization staff, research organization staff, and advocacy organization staff.

The workgroup recommends that the governing board make decisions using a two-thirds majority vote and be responsible for the following:

- Ensuring that the Cradle-to-Career Data System is serving its intended purpose
- Providing operational oversight of the managing entity
- Overseeing participation in the Cradle-to-Career Data System and governance structure
- Determining the types of information available through the Cradle-to-Career Data System
- Recommending improvements to the mechanisms for accessing information in the Cradle-to-Career Data System
- Monitoring technical, legal, and data implementation of the Cradle-to-Career Data System
- Hiring, evaluating, and firing the executive director

Figure Nine: Governing Board Seats

Advisory Boards

The workgroup recommends that the governing board be supported by two advisory boards that would be named in the implementing legislation for the Cradle-to-Career Data System.

- **Data and Tools Advisory Board:** examines whether the data system is providing actionable information and identifies ways to improve access to that information, including through visualizations on the dashboards, data points in the query builder, the data request process, and the content of the operational tools
- **Community Engagement Advisory Board:** examines whether the managing entity is creating strong feedback loops with data users, supports evidence-based decision-making and analytical capacity, and ensures equitable access to actionable information

The intent is for the advisory boards to serve a similar purpose as the public consultation requirements in the state's Local Control and Actionability Plan. The advisory boards are meant to ensure that the governing board consults with data experts and intended data users. Advisory board members would be appointed by the governing board and would be made up of practitioners, families, students, adult learners and workers, community organization staff, research organization staff, and advocacy organization staff. All meetings would be subject to the provisions of the Bagley-Keene Act.

Finally, the workgroup recommends that the governing board be empowered to create and sunset additional advisory boards and task forces, taking into consideration cost, size, and purpose, to address other issues that may arise during the implementation of the Cradle-to-Career Data System.

Managing Entity

Consistent with the vision and mission statements, the workgroup recommends that the managing entity for the Cradle-to-Career Data System be a neutral body that would house the data system and provide reliable information to support decision-making and advance equitable outcomes. For determining where to house the data system, the workgroup established a set of criteria that include the principles of adaptability, administrative flexibility, cost, durability, independence, neutrality, oversight and accountability, transparency and credibility, and trust and security. The workgroup also

determined that the managing entity should be a governmental body, rather than a nonprofit or an academic institution, but should not be one of the data providers.

Some members of the workgroup recommended housing the data system at a newly established independent commission, similar to the California Student Aid Commission or the California Commission on Teacher Credentialing. However, the other workgroup members were concerned that this approach might be more costly and could leave the managing entity without sufficient policy and political standing.

Ultimately, the workgroup agreed that the managing entity should be incubated from a new program or office within the Government Operations Agency (GovOps). With the managing entity placed within an agency that focuses on state operations and coordinates among myriad government programs, the data system is less likely to become politicized. Furthermore, given that GovOps is leading several efforts to improve the technical infrastructure of state programs, the creation and operation of the data system would benefit from that relevant expertise. Finally, by tapping into the extensive operational services within GovOps, early work on the data system could focus on the mechanics of linking and validating data and engaging in a user-centered design process for public tools, rather than having to address the operational start-up issues that would be encountered in other options. The workgroup recommends that after five years, at the end of phase one, the arrangement should be reviewed to determine whether having GovOps be the managing entity remains preferable or whether the managing entity should become an independent commission.

The managing entity would be tasked with the following roles and responsibilities:

- Support the governing board, advisory boards, and data providers
- Manage administrative functions for the Cradle-to-Career Data System
- Manage the technical infrastructure
- Implement public tools and support their use
- Support data requests
- Coordinate efforts to improve data quality
- Ensure legal compliance

The managing entity would focus on implementing the recommendations of the governing board and should have expertise in data and analytics to support the collection, provision, analysis, visualization, and use of data. Although the entity should provide written summaries of information available in the data system that relate to the

public good and equitable opportunities and outcomes for students, it should remain neutral and not create policy recommendations or establish its own policy agenda.

The managing entity would also play a critical role in staffing the data system. California's data contributors are disparate and largely independent agencies. By contrast, in other states, a single entity may represent early learning and care, K–12, and teacher credentialing. In addition, in most states, postsecondary partners are overseen by a coordinating board or a single state entity. However, California has nine separate entities and data systems for these same segments. By having the Cradle-to-Career Data System managing entity be responsible for coordinating directly with relevant data providers, or convening small teams to troubleshoot shared issues, the managing entity could help to streamline operations and reduce the burden on data providers, thereby ultimately better serving the public. Similarly, the managing entity could oversee additional advisory groups that would support implementation, such as convening counselors and students to provide input on college and career planning tools.

Read about the structure, membership, and decision-making rules for the governing board and the responsibilities of the managing entity: https://cadatasystem.wested.org/recommendations.

Section 6: Linking and Accessing Data

Record Linking

One of the foundational elements of longitudinal data systems is the ability to link records for individuals across various data providers. Historically, entities within California have relied on a shared identifier, such as using a social security number to link education and employment records, or using the CDE's statewide student identifier (SSID)[24] to link K–12 and postsecondary records. However, other states have moved to a more secure and accurate approach known as "master data management." The

[24] https://www.cde.ca.gov/ds/sp/cl/ssid.asp

workgroup recommends that the Cradle-to-Career Data System use master data management for linking individual records.

Master data management leverages all available shared elements to form a match, so that social security numbers and education identifiers can be paired with other variables such as name, address, date of birth, and demographic characteristics. Master data management processes include algorithms that address common variations in the data, such as typos or derivations of names (for example, William, Will, Bill, and Guillermo). The resulting linkages can be evaluated based on the reliability of the match, to determine whether matches are sufficiently strong for the intended use. For example, a researcher may find a 90 percent reliability rating sufficient for a trend analysis involving large numbers of students, whereas an evaluator seeking to understand a discrete intervention may seek a higher reliability rating.

Once records are matched, the system assigns each individual a new unique identifier, with different identifiers created for each data set. As a result, a researcher could request a data set and receive a list of records with no names but with a unique identifier for each student, created just for that study. The key for the linked records could be stored so that the data set could be reconstructed at a future date, but the data do not need to remain linked in the Cradle-to-Career Data System, thus reducing the risk that identities could be disclosed.

To obtain a better understanding of available master data management tools, the workgroup partnered with the California Department of Technology to release a Request for Information (RFI). This market scan found a large number of companies qualified to provide master data management, including off-the-shelf solutions, custom-development options, and providers that offer both master data management and other solutions that would be needed for the Cradle-to-Career Data System, such as data warehousing and dashboards. Most respondents had experience implementing SLDSs, were already supporting one of California's state agencies, or had experience with education and health and human services data.

Data Request Process

The workgroup recommends that the Cradle-to-Career Data System — in addition to making a significant amount of information freely available through its query builder — should have a data request process that streamlines access to data by distinguishing the level of information requested (Figure Ten).

- **Expedited review process:** Requests for de-identified aggregate data that are not available through the query builder would require a data request form to be filled out, reviewed by the managing entity, and approved by the relevant data providers. Data provided through the expedited process would not require a legal agreement.
- **Comprehensive review process:** Requests for anonymized individual-level data would require a comprehensive review. The requestor would fill out a data request form; the request would be reviewed by the managing entity; the request would be evaluated by the data providers and an institutional review board (IRB); the requestor would sign a legal agreement and provide proof of research training; analytical outputs would be reviewed to ensure data privacy; and data providers would review any reports submitted for inclusion in the Cradle-to-Career research library.

For both expedited and comprehensive reviews, the managing entity would work with requestors, as needed, to clarify their research question(s), explain which data are available through the Cradle-to-Career Data System, and ensure that the appropriate level of information is being requested. Once a request is fully developed, the managing entity would forward it to the appropriate data providers for review, along with a neutral summary that clarifies the nature of the request and its relevance to the Cradle-to-Career Data System vision and mission.

Documentation would be posted on the Cradle-to-Career Data System website to provide a list of all requests for data, who made each request, the project abstract, and the status of each request. Data providers would act in good faith to approve requests that are allowable by law. If data providers have questions or concerns about a data request, they would work with the requestor to modify the request. If a data provider denies a request, a rationale would need to be provided, which would be posted on the website.

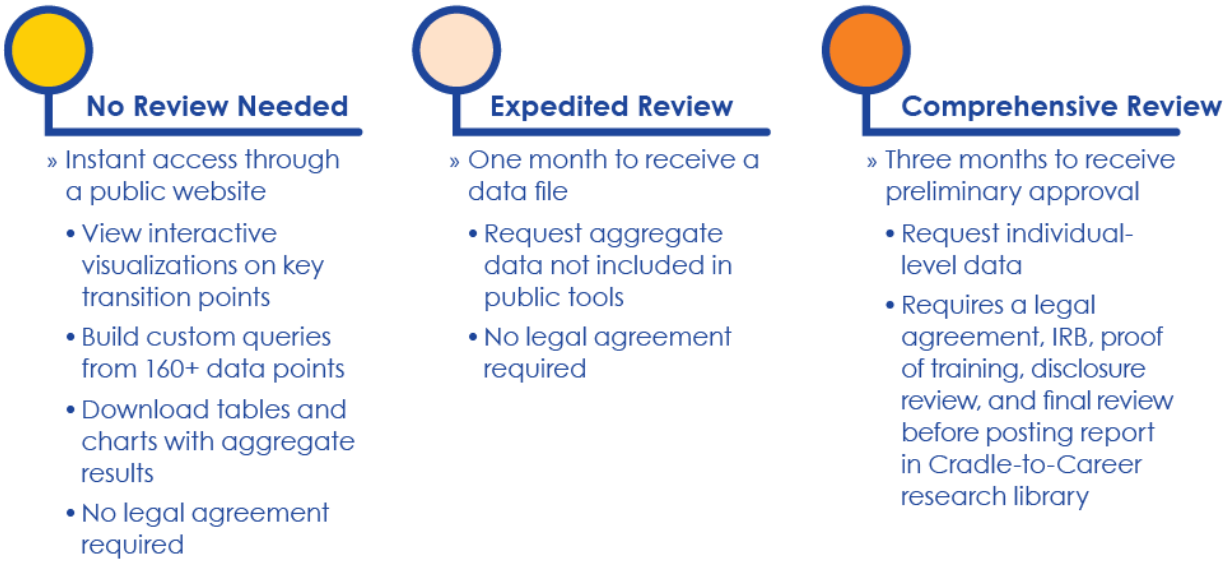
The proposed process for accessing information on individual students includes several checks and balances to ensure that individual identities would be protected, and that information would only be released in accordance with state and federal laws. For example, all requests for individual-level data must receive IRB approval from the Commission for the Protection of Human Subjects,[25] which is authorized to review education and health data for federal projects and which operates in compliance with

[25] https://oshpd.ca.gov/data-and-reports/data-resources/cphs/

the Information Practices Act. Furthermore, all outputs from the data system must be reviewed by a data disclosure specialist to ensure that identities cannot be discerned.

After year one of implementation, the data request process would be reviewed by the Data and Tools Advisory Board, with input from the managing entity and the data providers, to determine if the process could be improved. Each subsequent year, the Data and Tools Advisory Board would review the requests that were approved and denied and would then identify mechanisms that might better foster the fulfillment of appropriate data requests.

Figure Ten: Multiple Ways to Access Data



Read a full description of the data request process and view a visual summary: https://cadatasystem.wested.org/recommendations.

Section 7: Privacy, Security, and Data Definitions

Legal and Privacy Framework

The workgroup recommends leveraging the Interagency Data Exchange Agreement (IDEA) structure for legal agreements between the managing entity and all data providers that are state agencies. IDEA, which was originally developed to foster data sharing among the programs and departments in the California Health and Human Services Agency, is now being adopted across all state agencies. It pairs an overall agreement that establishes foundational legal assurances with a Business Use Case Proposal (BUCP) that identifies specific allowable uses of shared data.

To ensure that IDEA would meet the needs of the Cradle-to-Career Data System, the Legal Subcommittee submitted revised text regarding FERPA compliance and will develop a Cradle-to-Career Data System version of the BUCP. The Legal Subcommittee is also developing templates that would govern access to sensitive data for authorized users, such as to fulfill data requests. These templates will address each of the allowable use cases under FERPA and HIPAA, so that whenever a data request is approved, the relevant information can be integrated directly into the appropriate form. By removing the need to develop each data-sharing agreement from scratch, the timeline for fulfilling requests can be greatly reduced while also ensuring that uniform terms are applied. Finally, the Legal Subcommittee recommended that the data system not be subject to Public Record Act requests.

In spring 2021, the Legal Subcommittee will develop legal frameworks that undergird privacy laws, such as the way that individual records would be de-identified and the classification scheme that would determine the level of sensitivity for each data point (see the "Legal and Privacy Documentation" box).

Legal and Privacy Documentation

The Legal Subcommittee has developed several documents that are required for the data system, including:

- De-identification definitions
- FERPA addendum to the IDEA
- Opt-out policy
- Payment policy
- Personally Identifiable Information definition
- System disclaimer

Other policies that will be developed by June 2021 include:

- Cradle-to-Career Data System version of IDEA BUCP
- Managing entity data-sharing agreement
- Third-party data-sharing agreements
- Privacy policy

Read the legal documentation: https://cadatasystem.wested.org/meeting-information/legal-subcommittee.

Technical and Security Framework

Due to the sequencing of the workgroup's decisions, the Technology and Security Subcommittee is just starting to develop policies that address security and privacy considerations, including how information would be loaded into the data system, how it would be kept secure once in the data system, and how information would be made available to partner entities or third parties (see the "Technical and Security Documentation" box). Much of this work will be completed in spring 2021, after the legal framework has been established.

Technical and Security Documentation

The Technology and Security Subcommittee has developed several documents that are required for the data system, including the following:

- Data classification scheme
- Data security framework

Other topics that will be addressed by June 2021 include the following:

- Data quality management
- Extract, transform, and load requirements
- Security authentication methodology
- Security requirements for the cloud repository and data warehouse

Read the technical documentation: https://cadatasystem.wested.org/meeting-information/technology-security-subcommittee.

Data Definitions and Data Quality

The Act charges the workgroup with describing a process to develop common data definitions. As part of assessing data quality, the partner entities have already begun developing definitions for the P20W data points that would be used in the dashboards and query builder.

For each of the P20W data points, each partner entity will record whether and how the information has been captured, the dates for which the information is available, the values associated with the data point, and the quality of the information. After examining the similarities and differences between agency definitions, the Data Definitions Subcommittee will determine how the information should be displayed in the dashboards and query builder.

For example, the partner entities have found that they identify "economically disadvantaged" status in a variety of ways, including by whether students are eligible for free-or-reduced-price meals, receive financial aid, or have a parent with a college degree. Therefore, the partners determined that these more precise variables should be used, rather than creating a single data point for "economically disadvantaged" that would be defined differently by each data provider.

As part of this process, members of the Data Definitions Subcommittee also determine whether the definitions are ready for implementation, require further review once data are compiled, or would not be appropriate for the public tools and should be accessed by request only. The subcommittee also will determine whether particular data should come from a single partner entity or consist of data combined from multiple providers, based on quality and the sources of the information.

For example, because information on expected family contribution to college costs is derived from the Free Application for Federal Student Aid (FAFSA), which many community college students do not complete, the Data Definitions Subcommittee recommended a review of this data point after data are compiled, to see if the data provide a distorted view of how postsecondary students cover the cost of college. Similarly, the subcommittee recommended that information on the duration of financial aid be accessed by request only, because that information needs to be contextualized by factors that a casual user of the query builder might not be aware of, such as information on whether a student was continuously enrolled or was in college for so long that they stopped being eligible for financial aid.

The process of establishing data definitions has also helped identify ways to improve the information available to the public. For example, the definition for "foster youth" was crafted with partner entities, academic researchers, and advocacy organizations. Additional data-quality evaluation processes, such as how to correct data, will be addressed in spring 2021.

Data Definitions

Definitions for the public display of information have already been developed for the following kinds of data points, including identifying which entities collect each item, determining data quality, and reconciling differences in definitions:

- Variables needed for record matching
- Region
- Individual characteristics (age, sex/gender, race/ethnicity, parental education, foster youth, military, homeless, disabled, migrant, language)
- Financial aid participation
- Social service participation and free-and-reduced-price meals
- K–12 assessment scores, grade point average, a–g eligibility
- Postsecondary enrollment, grade point average, satisfactory progress
- Course characteristics (math pathways, career and technical education, development education)
- College goal and major

Review the public display options for data points: https://cadatasystem.wested.org/meeting-information/definitions-subcommittee.

Section 8: Professional Development, Technical Assistance, and Community Engagement

Early in the planning process, workgroup and advisory group members noted the unconscious tendency to list actions that they expected the data system would take. However, information cannot act for itself. Therefore, a critical component of the Cradle-to-Career Data System would be an outreach mechanism to ensure that a broad range of stakeholders are aware of the information and know how to use it to inform their decision-making.

Advisory group members have emphasized the importance of building the capacity of many different constituencies to use data. Building capacity includes ensuring that constituencies are aware of the ways that data — particularly information accessed through dashboards or a query builder — can reinforce bias. When charts can be produced that show different outcomes for one racial group compared to another, blaming students for those outcomes can be too easy in the absence of understanding the different educational opportunities and learning context available to each group.

As a first step in community engagement, the planning process has included several opportunities to solicit community input to support the data system design. In March 2020, WestEd hosted public webinars about possible data tools, and followed up with a survey. In October 2020, more than 400 people attended four webinars targeted for data experts, educators and practitioners, and students and families. WestEd reached dozens more practitioners, such as teachers and researchers, through eight presentations at conferences. After learning about the proposed tools, participants weighed in through Zoom polls, surveys, and a blog. The vast majority indicated that they would find the resources useful and indicated an intention to use them. Furthermore, most felt that the tools would be helpful for reducing equity gaps in access and opportunity (see the "Community Input on the Proposed Cradle-to-Career Data System Tools" box).

For example, one participant, reflecting on the college eligibility monitoring tools, noted they "would help parents understand what the student needs and as a parent it empowers us to ask the counselors for the necessary classes and support the student." Another noted that the tools "can help simplify the process which can help reduce fear of the unknown that first-generation college students and their families generally face." However, participants also noted that the usefulness of the tools will depend on the accuracy and timeliness of the data, the ability of guidance counselors to interpret the

information, the ability to make parents aware of the tools and how to use them, and the accessibility and user-friendliness of the tools and the technology.

When reflecting on the analytical tools, participants were encouraged that the data system allows for easier access to data that is all in one place, supports identification of patterns of inequity, and can potentially make demonstrating progress easier. However, some cautioned that the tools provide just a first step. As one participant, a data expert noted, "I think dashboards can be useful in identifying where disparate outcomes are occurring, but they're not nuanced enough to get to why. It will take follow-up research, discussions, and interventions to drive change."

To address these concerns, a new subcommittee, called the Community Engagement Subcommittee, will develop a set of recommendations for how the governing board and managing entity could implement a user-centered design process for the dashboards and query builder, build awareness about the tools, and offer training on appropriate data use across a range of communities.

Community Input on the Proposed Cradle-to-Career Data System Tools

- 96% of data expert and 81% of educator/practitioner respondents are likely to use the dashboards.
- 96% of data expert and 71% of educator/practitioner respondents are likely to use the query tools.
- 87% of educator/practitioner respondents are likely to use the operational tools.
- 96% of student/family respondents are likely to use the college and career planning tools.
- 100% of student/family respondents are likely to use the course monitoring/"a–g" tracking tool.
- Over 94% of survey respondents indicate that information could generate insights to help advocate for more equitable opportunities and outcomes.

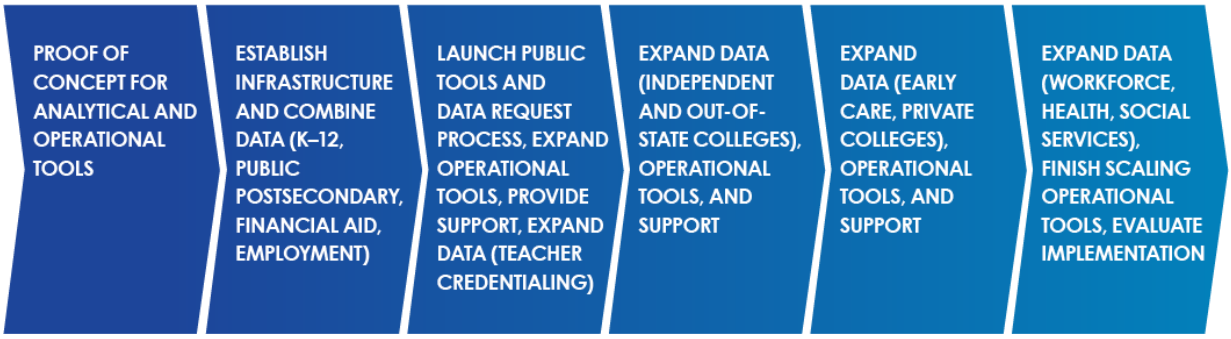
Read a feedback summary from the October 2020 community outreach campaign: https://cadatasystem.wested.org/feedback.

Section 9: Implementation

Sequencing and Timeline

The workgroup recommends that phase one of implementing the Cradle-to-Career Data System be sequenced as outlined in Figure Eleven and in the detailed timeline that follows. This timeline presumes that the system would be fully funded and that implementation would begin in July 2021, based on proof-of-concept investments in the first half of 2021. If work begins at a later date, or if the tasks of establishing the governing board, hiring staff, and procuring technical solutions takes an extended time, deliverables would be delayed. Although the proposed timeline is broken into fiscal years, some deliverables could be rolled out on a quarterly basis.

Figure Eleven: Sequence of Data System Development in Phase One



Timeline

Timeframe	Component	Item	Responsible Parties
Current Year: Conduct Proof-of-Concept Projects			
February 2021–June 2021	Analytical Tools	Test analytical tool process and create dashboard on teacher retention	California Department of Education (CDE), Commission on Teacher Credentialing (CTC), California State University (CSU), Government Operations Agency (GovOps)
	Operational Tools	Support college planning in low-income regions, begin upgrades to technical infrastructure	California College Guidance Initiative (CCGI)
Year One: Establish Infrastructure			
July 2021–June 2022	Governance	Establish membership and governance structure for the governing board and advisory boards, create the managing entity	Governor, Legislature, data providers
	Operations	Hire staff	Managing entity
	Operations	Create website and communication materials	Managing entity
	Operations	Develop an annual evaluation plan	Managing entity
	Analytical Tools	Procure and/or develop master data management, cloud storage, data warehousing, secure data enclave, and research request process tools	Managing entity

Timeframe	Component	Item	Responsible Parties
	Analytical Tools	Partners tag and load P20W data into the cloud, match records, build data warehouse	California Community Colleges Chancellor's Office (CCCCO), University of California (UC), Employment Development Department (EDD), California Health and Human Services Agency (CHHS), California Labor and Workforce Development Agency (CLWDA), California Student Aid Commission (CSAC), CDE, CSU, managing entity
	Analytical Tools	Launch website and provide summary of key student and employment outcomes	Managing entity
	Operational Tools	Integrate CCGI with CALPADS and upgrade data sharing with CSAC and CSU	CCGI, CDE, CSAC, CSU, managing entity
	Support Tools	Create user-centered design for dashboards, query builder, and research library	Community Engagement Advisory Board, managing entity
Year Two: Launch Data System			
July 2022–June 2023	Governance	Review and approve operational plans	Governing board
	Governance	Conduct evaluation and prepare annual report for the legislature	Governing board, managing entity
	Analytical Tools	Examine and improve data quality within the P20W data set	Managing entity
	Analytical Tools	Integrate data with public-facing tools	Managing entity

Timeframe	Component	Item	Responsible Parties
	Analytical Tools	Provide summary of key student and employment outcomes	Managing entity
	Analytical Tools	Release dashboards on a rolling basis throughout the year	Managing entity
	Analytical Tools	Release query builder	Managing entity
	Analytical Tools	Approve first set of data requests	Managing entity, relevant data partners
	Analytical Tools	Integrate data from teacher credentialing	Managing entity, CTC
	Analytical Tools	Secure legal agreements with independent colleges and the National Student Clearinghouse	Managing entity
	Operational Tools	Expand the number of districts implementing CCGI	CCGI, managing entity
	Operational Tools	Upgrade infrastructure for eTranscript California, develop permissions tracking for social service data	eTranscript California, California Department of Social Services (CDSS), managing entity
	Support Tools	Implement awareness campaign and begin professional development campaign	Community Engagement Advisory Board, managing entity
Year Three: Evaluate and Improve Data System			
July 2023–June 2024	Governance	Review and approve operational plans	Governing board
	Governance	Conduct evaluation and prepare annual report for the legislature	Governing board, managing entity

Timeframe	Component	Item	Responsible Parties
	Analytical Tools	Produce summary of key student and employment outcomes	Managing entity
	Analytical Tools	Integrate data from the independent colleges and the National Student Clearinghouse	Independent colleges, managing entity
	Analytical Tools	Evaluate efficacy of data request process	Governing board, Data and Tools Advisory Board, managing entity
	Analytical Tools	Update and improve tools, based on community input	Data and Tools Advisory Board, managing entity
	Analytical Tools	Evaluate feasibility and value of a synthetic data set	Governing board, Data and Tools Advisory Board, managing entity
	Operational Tools	Continue expanding the number of districts implementing CCGI, begin development of CTE and apprenticeship pathways tools, integrate labor market information	CCGI, managing entity
	Operational Tools	Launch eTranscript California for public postsecondary institutions	CCCCO, CSU, UC, eTranscript California, managing entity
	Support Tools	Continue professional development campaign, begin technical assistance campaign	Community Engagement Advisory Board, managing entity

Timeframe	Component	Item	Responsible Parties
Year Four: Expand Features			
July 2024–June 2025	Governance	Review and approve operational plans	Governing board
	Governance	Conduct evaluation and prepare annual report for the legislature	Governing board, managing entity
	Analytical Tools	Produce summary of key student and employment outcomes	Managing entity
	Analytical Tools	Integrate private college and early learning and care data sets	CDSS, CDE, CTC, Bureau for Private Postsecondary Education (BPPE), managing entity
	Analytical Tools	Plan for expansion of social service and health, workforce data sets	Governing board, Data and Tools Advisory Board, managing entity
	Analytical Tools	Potentially build a synthetic data set	Governing board, Data and Tools Advisory Board, managing entity
	Analytical Tools	Update and improve tools, based on community input	Data and Tools Advisory Board, managing entity
	Operational Tools	Continue expanding the number of districts implementing CCGI, launch CTE tools	CCGI, managing entity
	Operational Tools	Launch eTranscript California services for private and independent postsecondary schools	Association of Independent California Colleges and Universities (AICCU), BPPE, eTranscript California, managing entity

Timeframe	Component	Item	Responsible Parties
	Support Tools	Continue professional development campaign, begin technical assistance campaign	Community Engagement Advisory Board, managing entity
Year Five: Complete Implementation and Plan for Phase Two			
June 2025–July 2026	Governance	Evaluate whether the managing entity should continue as a program within GovOps	Governing board, managing entity
	Governance	Assess whether the data system is yielding expected outcomes	Independent evaluator
	Governance	Review and approve operational plans	Governing board
	Governance	Establish strategic objectives, evaluate whether to recruit additional data providers	Governing board, Data and Tools Advisory Board, Community Engagement Advisory Board, managing entity
	Analytical Tools	Produce summary of key student and employment outcomes	Managing entity
	Analytical Tools	Integrate and test social service, health, workforce, and teacher credentialing data sets	CHHS, CLWDA
	Analytical Tools	Assess analytical tools and develop an improvement strategy	Governing board, Data and Tools Advisory Board, managing entity
	Operational Tools	Complete scaling of CCGI to all districts	CCGI, managing entity
	Operational Tools	Assess operational tools, develop a strategy for improving them	Governing board, Data and Tools Advisory Board, managing entity

Timeframe	Component	Item	Responsible Parties
	Support Tools	Assess support tools, develop a strategy for improving them	Governing board, Community Engagement Advisory Board, managing entity

Proof-of-Concept Project

In addition to allocating funds to support the planning process, the Act included funds that can be spent "with the approval of an expenditure plan by the Department of Finance and notification to the Joint Legislative Budget Committee based upon reporting from the planning facilitators." The Governor's Office is proposing to use these funds to launch a proof-of-concept project that leverages existing resources to demonstrate how the Cradle-to-Career Data System would work.

Analytical Data

The Commission on Teacher Credentialing (CTC), CDE, and CSU have volunteered to provide the data necessary to analyze teacher workforce retention trends and document the impact of teacher preparation programs across the state. Connecting this information will realize an effort begun 15 years ago to develop a longitudinal teacher information data system. It will also provide a proof-of-concept for the proposed data system, including testing the legal templates, data coding and uploading process, use of a master data management solution to match records, review and certification of data tables, and implementation of visualization tools that make the information available to a broad range of potential users. This practical experience will help to refine requirements for technology procurements that will be implemented in year one of the data system.

Given the impending workforce shortage, which has been accelerated by the strain of teaching during the pandemic, it is critical for the state to understand where educators have historically been retained in their jobs, the subject areas that they teach, their movement between regions, and which students are being served by multiple educators due to turnover. Understanding where turnover is higher and examining educator profiles will allow teacher preparation programs to determine if changes can be made to mitigate this churn. Having information about turnover is especially critical regarding teachers of color, to foster greater understanding of how they enter and progress through the educator development system. The project lays the groundwork

for additional data matches that could produce information on the relative diversity of educators compared to the students they serve, common educational pathways that lead to teaching, and the occupations that teachers enter when they leave the teaching profession.

Operational Data

In the Central Valley and Inland Empire, CCGI has played a key role in helping districts support students in preparing and submitting financial aid and college applications. For example, high schools in these regions used CCGI tools in fall 2020 to pinpoint outreach to students who had begun but not completed their applications, which helped with attaining goals that had been set before the pandemic.[26] As part of the proof-of-concept undertaking, funds would be provided to CCGI to support its efforts in regions of the state that have the lowest college-going rates.

In addition, proof-of-concept funds would be allocated to begin the overhaul of CALPADS that is necessary to provide up-to-date information that will be used in college eligibility and application tools, as well as improve the technical infrastructure used to exchange information between CCGI, CSAC, and CSU.

Estimated Costs for Year One (2021–22)

Developing the Cradle-to-Career Data System will entail two types of investments: one-time expenses to set up the system and annual costs to maintain and improve it. At this point in the planning process, developing a definitive budget is challenging because the exact costs of the technical infrastructure are unclear. However, given expenses that have already been documented, the workgroup estimates the fiscal impact would be $15–$20 million for 2021–22. Additional estimates regarding ongoing costs will be provided in the next legislative report.

The proposed costs are based on identifying budget categories, examining budgets for data systems in other states, including the costs for scaling CCGI, and the proof-of-concept. While some of these costs are one-time, many would require a commitment of ongoing funding for the data system to function. This commitment is particularly

[26] https://calmatters.org/commentary/my-turn/2020/11/navigating-the-transition-between-k-12-and-higher-education-virtually/

important for funding to support staff positions, to ensure that the appropriate talent could be recruited and retained.

To develop the 2021–22 fiscal impact estimate, the workgroup used the following categories and considerations:

Staffing and Operations: Other state data systems that include dashboards, allow for research requests, and use vendors to implement their technical infrastructure (as opposed to creating a home-grown system) employ between 11 and 13 staff members. The managing entity responsibilities outlined by the workgroup indicate California would likely need a similar number of positions. Other costs include the administrative and managerial costs for GovOps to administer the system, such as office space, equipment, external consultants, and materials associated with community outreach.

Governing Board: A line item should be included for governance-related costs, such as travel and per diem for quarterly meetings of the governing board and the two advisory boards.

Support for Partner Entities: During the first year of implementation, the partner entities providing the bulk of the information (in California's case that includes CDE, CSU, UC, CSAC, California Community Colleges Chancellor's Office, and Employment Development Department) would need additional support and resources in order to tag, load, and test the data to ensure the record matching algorithm and the information produced are reliable. These tasks cannot be easily outsourced to or by GovOps because the work requires a deep knowledge of segment-specific data to ensure that the information produced is of high quality. Furthermore, because the CDE is in the pivotal position of supporting both the implementation of the college planning and eligibility tools and of providing the core data for the P20W data set, the CDE may need additional resources during both the startup and implementation phases of this work. Finally, because it is unclear how much time would be required of data providers once the data request system becomes available, or if some segments will have a higher level of demand than others, data provider staffing will need to be assessed in future years.

Analytical Tools: In the first year, the managing entity would need to procure or build the technical tools necessary to compile and provide access to information. These tools include mechanisms for data loading, data hosting, data warehousing, data documentation, data matching and integration, security, data request tracking, public dashboards, the query builder, and a secure data enclave.

Operational Tools: To scale CCGI, upgrades are needed to the underlying data infrastructure managed by CDE. CALPADS was designed for the purpose of periodic reporting, primarily for accountability purposes, and needs to be modernized before it can accommodate more frequent exchanges of information. CCGI would then build an integration with CDE and upgrade existing integrations with postsecondary and financial aid partners.

Conclusion

When the planning process started for a California Cradle-to-Career Data System, the state faced a very different landscape than it does today. Californians have had to stay resilient through a global pandemic, an economic downturn, a state budget shortfall, the transformation of education to an online format, reckoning with racial injustice, devastating wildfires, and significant political upheaval. Even as the members of the planning process were called to address these unprecedented challenges, they did not shy away from their work to design the Cradle-to-Career Data System. If anything, the crises of 2020 made the need for a longitudinal data system — one that looks beyond college-going rates — even more clear and compelling.

The proposed model would provide a critical information infrastructure to help a wide range of stakeholders determine the best path forward and support California's recovery:

- Students and families could plan for their futures and seek the support services they need to be successful.
- K–12 educators could provide stronger guidance to students and increase the number who apply to college and for financial aid.
- Researchers and policymakers could understand how the crises of 2020 impacted education and employment outcomes.
- Advocates could raise awareness about which Californians are attaining their goals and which are left behind.
- Education leaders could evaluate interventions, identify bright spots, and clarify how public services work in concert to advance equity.

Although start-up and implementation will require an investment of one-time and ongoing funds, it would strengthen agency data systems. The Cradle-to-Career Data System is a worthwhile investment that is crucial for giving students and families the

information they need to pursue their dreams, despite the challenges they face, and for helping the state develop public policy based on reliable and actionable information.

Appendix

Location of Legislative Requirements in the Report

Requirement	Code Section	Legislative Report Location
First Legislative Report Requirements[27] A planning facilitator that contracts with the Office of Planning and Research pursuant to subdivision (b) of Section 10855 shall report to the Department of Finance and the Legislature by July 1, 2020, on the proposed structure of the data system, including, but not limited to, the entity charged with managing the data system, the architecture of the data system, and the information that will be available on the data system		
Entity: Entity charged with managing the data system	Education Code Section 10856 (a)	Section 5: Governance | Managing Entity, pages 28–30 and https://cadatasystem.wested.org/recommendations
Architecture: Architecture of the data system	Education Code Section 10856 (a)	Section 4: Technical Structure | Analytical Tools pages 14–15 and https://cadatasystem.wested.org/recommendations
Available Data: The information that will be available on the data system	Education Code Section 10856 (a)	Section 3: Purpose and Available Data | Proposed Information for Phase One, pages 12–13 and https://cadatasystem.wested.org/recommendations
Disaggregation: How the system could have the capacity to disaggregate: race, ethnicity, region, gender, military status, parents' education, and age	Education Code Section 10856 (b)(1)	Section 7: Privacy, Security, and Data Definitions | Data Definitions and Data Quality, pages 36–38 and https://cadatasystem.wested.org/meeting-information/definitions-subcommittee

27 http://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=EDC&division=1.&title=1.&part=7.&chapter=8.5.&article=

Requirement	Code Section	Legislative Report Location
Transferring Educational Records: How to transfer high school pupil educational records to postsecondary educational institutions	Education Code Section 10856 (b)(2)	Section 4: Technical Structure | Operational Data, pages 20–23 and https://cadatasystem.wested.org/recommendations
Early Education: The impact of early education on student success and achievement as the student progresses through education segments and the workforce	Education Code Section 10856 (c)(1)	Section 4: Technical Structure | Accessing Analytical Data, pages 17–20 and https://cadatasystem.wested.org/recommendations
Primary Education: The long-term effect of state intervention programs and targeted resource allocations in primary education	Education Code Section 10856 (c)(2)	Section 4: Technical Structure | Accessing Analytical Data, pages 17–20 and https://cadatasystem.wested.org/recommendations
College and Career Readiness: How prepared high school pupils are to succeed in college	Education Code Section 10856 (c)(3)	Section 4: Technical Structure | Accessing Analytical Data, pages 17–20 and https://cadatasystem.wested.org/recommendations
Transfer Outcomes: How long it takes students who transfer from community college to the University of California, the California State University, or another four-year postsecondary educational institution to graduate with a baccalaureate degree	Education Code Section 10856 (c)(4)	Section 4: Technical Structure | Accessing Analytical Data, pages 17–20 and https://cadatasystem.wested.org/recommendations
Financial Aid: College access, completion, and long-term effects of access to state financial aid	Education Code Section 10856 (c)(5)	Section 4: Technical Structure | Accessing Analytical Data, pages 17–20 and https://cadatasystem.wested.org/recommendations
Workforce: The workforce effect of graduation from high school, community college, and four-year postsecondary educational institutions	Education Code Section 10856 (c)(6)	Section 4: Technical Structure | Accessing Analytical Data, pages 17–20 and https://cadatasystem.wested.org/recommendations

Requirement	Code Section	Legislative Report Location
Prior Efforts: It is the intent of the Legislature that the workgroup review and build upon prior reports, including those produced pursuant to Senate Bill 1298 of the 2007–08 Regular Session (Chapter 561 of the Statutes of 2008), and review and build upon existing data systems and best practices, as appropriate and relevant	Education Code Section 10856 (d)	Section 2: The Planning Process | Background Research, page 2 and https://cadatasystem.wested.org/resources-page
Second Legislative Report Requirements A planning facilitator that contracts with the Office of Planning and Research pursuant to subdivision (b) of Section 10855 shall report to the Department of Finance and the Legislature by January 1, 2021, its recommendations on all of the following:		
Data Access: How access to data will be controlled and authorized	Education Code Section 10857 (a)(1)	Section 6: Linking and Accessing Data | Data Request Process, pages 31–33 and https://cadatasystem.wested.org/meeting-information/technology-security-subcommittee
Common Data Definitions: A means of developing common data definitions	Education Code Section 10857 (a)(2)	Section 7: Privacy, Security, and Data Definitions | Data Definitions and Data Quality, pages 36–38 and https://cadatasystem.wested.org/meeting-information/definitions-subcommittee
Additional Data Elements: Additional data elements necessary for partner entities to collect for future linkage to the data system	Education Code Section 10857 (a)(3)	Section 4: Technical Structure | Data for Analytical Purposes, page 16 and https://cadatasystem.wested.org/recommendations

Requirement	Code Section	Legislative Report Location
Statutory Changes: Any specific changes to state law that may be necessary for implementation	Education Code Section 10857 (a)(4)	For scaling of the California College Guidance Initiative, https://cadatasystem.wested.org/recommendations Other items will be addressed in second legislative report
Training and Support: Plans for training and support of users	Education Code Section 10857 (a)(5)	Will be addressed in second legislative report
System Access: How users will be able to access the data system	Education Code Section 10857 (a)(6)	Section 4: Technical Structure | Accessing Analytical Data, pages 17–20 and https://cadatasystem.wested.org/recommendations
Data Entry and Correction: How authorized users will enter and correct data	Education Code Section 10857 (a)(7)	Will be addressed in second legislative report
Data Quality: How to ensure data quality from each component of the education system and participating entities	Education Code Section 10857 (a)(8)	Will be addressed in second legislative report
Security and Data Privacy: How to address security and data privacy considerations, including compliance with existing state and federal data privacy and use laws, and standard elements that will be included in agreements necessary to share data among partner entities	Education Code Section 10857 (a)(9)	Section 7: Privacy, Security, and Data Definitions | Security Framework, pages 35–36, https://cadatasystem.wested.org/meeting-information/legal-subcommittee , and https://cadatasystem.wested.org/meeting-information/technology-security-subcommittee
Memoranda of Understanding: How to connect data in a manner that limits the number of memoranda of understanding necessary and maximizes efficiencies	Education Code Section 10857 (a)(10)	Section 7: Privacy, Security, and Data Definitions | Legal Framework, pages 34–35, https://cadatasystem.wested.org/meeting-

Requirement	Code Section	Legislative Report Location
		information/legal-subcommittee
Data Request Process: How to respond to requests from researchers to access data	Education Code Section 10857 (a)(11)	Section 6: Linking and Accessing Data | Data Request Process, pages 31–33 and https://cadatasystem.wested.org/recommendations
Data Expansion: How the data system should be expanded to incorporate childcare, early education, workforce, financial aid, and health and human services data	Education Code Section 10857 (a)(12)	Section 9: Implementation | Sequencing and Timeline, pages 41–48
Implementation Timeline: An implementation timeline, including key project milestones and sequencing for functionality expansions and enhancements	Education Code Section 10857 (a)(13)	Section 9: Implementation | Sequencing and Timeline, pages 41–48
Fiscal Estimate: The estimated fiscal impact of developing the data system and the cost of recommended expansions and enhancements, including the ongoing management costs	Education Code Section 10857 (a)(14)	Section 9: Implementation | Estimated Costs, pages 49–51 Ongoing costs will be addressed in second legislative report
Prioritization of Key Components: A prioritization of key components needed to best enhance and expand the data system if available funding is insufficient to address all desirable elements	Education Code Section 10857 (a)(15)	Executive Summary, pages i–vi
Public Interface: How to create a public-facing interface to share information with the public that can help inform decisions	Education Code Section 10857 (a)(16)	Section 4: Technical Structure | Accessing Analytical Data, pages 17–20 and https://cadatasystem.wested.org/recommendations
Statewide Student Identifier (SSID): How to identify and track students who do not have a preexisting	Education Code Section 10857 (a)(17)	Section 6: Linking and Accessing Data | Record Linking, pages 30–31

Requirement	Code Section	Legislative Report Location
statewide student identifier from the State Department of Education		
Prioritization for Implementation: To the extent practicable, the workgroup shall prioritize the implementation of the data system in the following order: (A) Phase 1: K–12 and higher education (B) Phase 2: Workforce (C) Phase 3: Early care and education (D) Phase 4: Health and human services and other data connections	Education Code Section 10857 (b)(1)	Section 3: Purpose and Available Data | Proposed Information for Phase One, pages 12–13 and https://cadatasystem.wested.org/recommendations

Documentation Referenced in the Report

Background Research (Page 2)

https://cadatasystem.wested.org/resources-page

- Analysis of prior state efforts
- Recommendations from policy organizations
- Background papers that informed each meeting
- Acronym glossary

Planning Timeline (Page 7)

https://cadatasystem.wested.org/meeting-information/calendar-for-all-meeting-dates

- Calendar with meeting topics

Analytical Data (Page 16)

https://cadatasystem.wested.org/recommendations

- Summary of the goals and content for the dashboards and query builder
- User stories that describe who would use each tool
- Publicly available data, including content of dashboards and data points in the P20W data set
- Additional data points desired for future phases of the data system

Operational Tools (Page 20)

https://cadatasystem.wested.org/recommendations

- User stories that describe who would use each tool
- Chart summarizing the scaling proposal for the operational tools
- Proposed technical data flows for CCGI
- Proposed technical data flows for eTranscript California
- WestEd report on operational tools

Governance (Page 27)

https://cadatasystem.wested.org/recommendations

- Structure, membership, and decision-making rules for the governing board
- Responsibilities of the managing entity

Data Request Process (Page 31)

https://cadatasystem.wested.org/recommendations

- Description of the data request process
- Flow chart for the data request process

Legal Documentation (Page 35)

https://cadatasystem.wested.org/meeting-information/legal-subcommittee

- De-identification definitions
- FERPA addendum to the IDEA
- Opt-out policy
- Payment policy
- Personally identifiable information definition
- System disclaimer

Technical Documentation (Page 36)

https://cadatasystem.wested.org/meeting-information/technology-security-subcommittee

- Data security framework
- Data classification scheme
- Technical visualization of data flow

Data Point Documentation (Page 38)

https://cadatasystem.wested.org/meeting-information/definitions-subcommittee

- Variables needed for record matching
- Region

- Individual characteristics (age, sex/gender, race/ethnicity, parental education, foster youth, military, homeless, disabled, migrant, language)
- Financial aid participation
- Social service participation and free-and-reduced-price meals
- K–12 assessment scores, grade point average, a–g eligibility
- Postsecondary enrollment, grade point average, satisfactory progress
- Course characteristics (math pathways, career and technical education, development education)
- College goal and major

Community Engagement (Page 39)

https://cadatasystem.wested.org/feedback

- Feedback summary from the October 2020 community outreach campaign

Planning Process Participants

Cradle-to-Career Workgroup

Association of Independent California Colleges and Universities

- Thomas Vu, Vice President for Policy

Bureau for Private Postsecondary Education

- Michael Marion, Chief Executive Officer
- Leeza Rifredi, Deputy Bureau Chief

California Commission on Teacher Credentialing

- Michele Perrault, Director of the Administrative Services Division

California Community Colleges Chancellor's Office

- Barney Gomez, Vice Chancellor for Digital Innovation and Infrastructure Division

California Department of Education

- Cindy Kazanis, Director of the Analysis of the Measurement, and Accountability Reporting Division
- Sarah Neville-Morgan, Deputy Superintendent of the Opportunities for All Branch
- Mary Nicely, Senior Policy Advisor to the State Superintendent of Public Instruction

California Department of Social Services

- Akhtar Khan, Branch Chief for the Research Services Branch, Research, Automation, and Data Division
- Natasha Nicolai, Chief Data Strategist

California Department of Technology

- Brenda Bridges Cruz, Deputy Director of the Office of Professional Development

California Health and Human Services

- Elaine Scordakis, Assistant Director of the California Office of Health Information Integrity

California Labor and Workforce Development Agency

- Jeanne Wolfe, Assistant General Counsel

California School Information Services

- Amy Fong, Chief Operations Officer

California State Board of Education

- Sara Pietrowski, Policy Consultant

California State University, Office of the Chancellor

- Ed Sullivan, Assistant Vice Chancellor for Academic Research and Resources

California Student Aid Commission

- Patrick Perry, Division Chief of Policy, Research, and Data

Employment Development Department

- Amy Faulkner, Chief of the Labor Market Information Division

Government Operations Agency

- Joy Bonaguro, Statewide Chief Data Officer

University of California, Office of the President

- Chris Furgiuele, Director for Institutional Research and Planning

Policy and Analytics Advisory Group

Asian Americans Advancing Justice

- Liza Chu, California Policy Manager

California Budget and Policy Center

- Kristin Schumacher, Senior Policy Analyst

California Competes

- Su Jin Gatlin Jez, Executive Director

California Policy Lab

- Evan White, Executive Director

California Teachers Association

- Brian Guerrero, Former President and 7th grade Language Arts teacher at Lennox Middle School

Child Trends

- Carlise King, Executive Director of the Early Childhood Data Collaborative

Children's Data Network

- Emily Putnam-Hornstein, Director

Children Now

- Samantha Tran, Senior Managing Director of Education Policy

Education Insights Center

- Andrea Venezia, Executive Director

Education Trust-West

- Christopher Nellum, Deputy Director of Research and Policy

Educational Results Partnership

- James Lanich, President and Chief Executive Officer

GreatSchools.org

- Orville Jackson, Senior Director of Data Strategy

Policy Analysis for California Education

- Heather Hough, Executive Director

Public Advocates

- Liz Guillen, Director of Legislative and Community Affairs
- Rigel Spencer Massaro, Senior Legislative Counsel

Public Policy Institute of California

- Jacob Jackson, Research Fellow

The Institute for College Access and Success

- Angela Perry, Policy Analyst

Practice and Operations Advisory Group

Bakersfield College

- Craig Hayward, Dean of the Office of Institutional Effectiveness

California Federation of Teachers

- Reynaldo Dulaney, Jr., High School Teacher, Thurgood Marshall High School

Campaign for College Opportunity

- Sara Arce, Vice President of Policy and Advocacy

Child Care Resource Center

- Susan Savage, Director of Research

CORE Districts

- Rick Miller, Executive Director

Foothill College

- Laurie Scolari, Associate Vice President of Student Services

Futuro Health

- Anthony Dalton, Chief Data and Technology Officer

Lancaster Unified School District

- Michele Bowers, Superintendent

Linked Learning Alliance

- Roneeta Guha, Vice President for Strategy and Impact

Poway Unified School District

- Kathleen Porter, Executive Director of Career Technical, Adult, and Alternative Education

Reinvent Stockton Foundation

- Lange Luntao, Executive Director

Riverside County Office of Education

- Catalina Cifuentes, Executive Director of College and Career Readiness

Santa Clara County Office of Education

- Joell Hanson, Career Technical Education Coordinator
- Marcy Lauck, Retired Senior Advisor for Data Initiatives
- Nabil Shahin, Director of K–12 Data Governance

uAspire

- Jaclyn Pinero, Chief Regional Officer & Co-Chief Executive Officer
- Tyler Wu, California Policy Director

UNITE-LA

- David Rattray, President

West Hills College Lemoore

- Kristin Clark, President

Common Identifier Subcommittee

Association of Independent California Colleges and Universities

- Jonathon Chillas, Chief Data Officer and Vice President of National University

Bureau for Private Postsecondary Education

- Scott Valverde, Chief of Student Assistance and Relief

California College Guidance Initiative

- Ben Baird, Director of CaliforniaColleges.edu Data Infrastructure

California Commission on Teacher Credentialing

- Michele Perrault, Director of the Administrative Services Division

California Community Colleges Chancellor's Office

- Todd Hoig, Director of the Management Information Systems

California Department of Education

- Channa Hewawickrama, Education Research and Evaluation Consultant for the Early Learning and Care Division
- Jerry Winkler, Division Director of the Educational Data Management Division

California Department of Social Services

- Jayson Hunt, Research Data Specialist
- Akhtar Khan, Branch Chief, Research Services Branch, Research, Automation, and Data Division

California Department of Technology

- Janet Buehler, Enterprise Architect

California Health and Human Services Agency

- Chris Krawczyk, Chief Analytics Officer
- Jennifer Schwartz, Chief Counsel

California Labor and Workforce Development Agency and Employment Development Department

- Amy Faulkner, Division Chief of the Labor Market Information Division

California School Information Services

- Greg Scull, Information Systems Officer

California State Board of Education

- Sara Pietrowski, Policy Consultant

California State University Office of the Chancellor

- Jeff Whitney, Director of Implementation Services

Cal Poly Pomona

- Joseph Hackbarth, Director of Enrollment Management Data and Operations Analysis

California Student Aid Commission

- Patrick Perry, Division Chief for Policy, Research, and Data

Educational Results Partnership

- Dan Lamoree, Principal Data Architect

University of California, Davis

- Paco Martorell, Associate Professor for the School of Education

University of California Office of the President

- Eric Goodman, Data Architect and Identity Access Management Lead

University of Southern California

- John Prindle, Research Assistant Professor

Definitions Subcommittee

Association of Independent California Colleges and Universities

- Todd Britton, Acting Vice President and Chief Information Officer for University of LaVerne
- Randy Tarnowski, Director of Research

Bureau for Private Postsecondary Education

- Joanna Murray, Senior Education Specialist

California Commission on Teacher Credentialing

- Phi Phi Lau, Consultant
- Erin Skubal, Director, Certification
- Marjorie Suckow, Consultant

California Community Colleges Chancellor's Office

- Todd Hoig, Director of Management Information Systems

California Department of Education

- Channa Hewawickrama, Education Research and Evaluation Consultant for the Early Learning and Care Division
- Jerry Winkler, Division Director of the Educational Data Management Division

California Department of Health Care Services

- Muree Larson-Bright, Staffing Audits and Research Branch Chief
- Phil Heinrich, Senior Information Systems Analyst

California Department of Social Services

- Patrick Delaney, Research Data Specialist
- Akhtar Khan, Branch Chief for the Research Services Branch, Research, Automation, and Data Division

California Department of Technology

- Janet Buehler, Enterprise Architect

California Health and Human Services Agency

- David Sanabria, Data Architect
- Jennifer Schwartz, Chief Counsel

California Labor and Workforce Development Agency

- Patrick Getz, Information Technology Project Manager for CAAL-Skills

California Office of Systems Integration

- David Sanabria, Data Architect

California School Information Services

- Rima Mendez, Requirements Officer

California State University Office of the Chancellor

- Monica Malhotra, Director of Strategic Analytics

California State Board of Education

- Sara Pietrowski, Policy Consultant

California Student Aid Commission

- Adrian Felix, Research Data Specialist

Employment Development Department

- Margo Gonzales, Deputy Division Chief of the Labor Market Information Division

University of California Office of the President

- Chris Furgiuele, Institutional Research Manager

West Coast University

- Valerie Mendelsohn, Vice President of Compliance and Risk Management

Legal Subcommittee

Association of Independent California Colleges and Universities

- Veronica Villalobos Cruz, Managing Director of Cruz Strategies

Bureau for Private Postsecondary Education

- Douglas Smith, Counsel

California Commission on Teacher Credentialing

- Katie Elliott, Attorney
- Linh Nguyen, Chief Deputy Director

California Community Colleges Chancellor's Office

- Kathy Lynch, Deputy Counsel

California Department of Education

- Bruce Yonehiro, Deputy General Counsel

California Department of Health Care Services

- Cynthia Bosco, Attorney III

California Department of Social Services

- Akhtar Khan, Branch Chief for the Research Services Branch, Research, Automation, and Data Division
- Carolyn Kubish, Senior Counsel

California Department of Technology

- Jennifer Marquez, Attorney
- Kary Marshall, Attorney

California Government Operations Agency

- Gabriel Ravel, Deputy Secretary and General Counsel

California Health and Human Services Agency

- Jennifer Schwartz, Chief Counsel

California Labor and Workforce Development Agency

- Jeanne Wolfe, Assistant General Counsel

California School Information Services

- Rima Mendez, Requirements Officer

California State Board of Education

- Sara Pietrowski, Policy Consultant

California State University Office of the Chancellor

- Arun Casuba, Executive Director and Strategic Sourcing and Chief Procurement Officer
- Ed Hudson, Chief Information Security Officer

California Student Aid Commission

- Julia Blair, Chief Legal Counsel

Employment Development Department

- Brian Davis, Senior Attorney

Government Operations Agency

- Gabriel Ravel, Deputy Secretary and General Counsel

University of California Office of the President

- Stella Ngai, Senior Counsel, Public Accountability and Governance

Research Agenda Subcommittee

Association of Independent California Colleges and Universities

- Thomas Vu, Vice President for Policy

Bureau for Private Postsecondary Education

- Alma Mededovic, Compliance Analyst

California Commission on Teacher Credentialing

- Tine Sloan, Commission Chair

California Community Colleges Chancellor's Office

- John Hetts, Visiting Executive for Research and Data for the Educational Services and Support Division

- Valerie Lundy-Wagner, Assistant Vice Chancellor of Research and Data for the Educational Services and Support Division

California Department of Education

- Ben Allen, Education Programs Consultant for the Early Learning and Care Division
- Jonathan Isler, Administrator for the Data Visualization and Reporting Office, Analysis, Measurement, and Accountability Division

California Department of Social Services

- Natasha Nicolai, Chief Data Strategist

California Department of Technology

- Janet Buehler, Enterprise Architect

California Health and Human Services Agency

- Jennifer Schwartz, Chief Counsel

California Labor and Workforce Development Agency

- Dan Rounds, Deputy Director for Policy, Research, and Legislation

California Office of Statewide Health Planning and Development

- Chris Krawczyk, Chief Analytics Officer

California School Information Services

- Martha Friedrich, Client Services Officer

California State University Office of the Chancellor

- Nathan Evans, Chief of Staff and Senior Advisor for Academic and Student Affairs

California State Board of Education

- Sara Pietrowski, Policy Consultant

California Student Aid Commission

- Jessica Moldoff, Research Data Specialist

Centers of Excellence for Labor Market Research

- Laura Coleman, Statewide Director

Employment Development Department

- Muhammad Akhtar, Deputy Division Chief of the Labor Market Information Division

Gurnick Academy of Medical Arts

- Abraham Cicchetti, Campus Director

Hoss Lee Academy, Inc.

- Lisa Lee, Owner/Director

Research and Planning Group for California Community Colleges

- Alyssa Nguyen, Director of Research and Evaluation

UnitekLearning.com

- Steve Watkins, Vice President of Program Development and Regulatory Affairs

University of California, Berkeley

- Tameka McGlawn, Executive Director of the College and Career Academy Support Network
- Jesse Rothstein, Professor of Public Policy and Economics

University of California, Davis

- Michal Kurlaender, Professor and Department Chair for the School of Education
- Sherrie Reed, Executive Director of the California Education Lab

University of California Office of the President

- Tongshan Chang, Director for Institutional Research and Planning

University of California, Santa Barbara

- Russell Rumberger, Professor Emeritus of the Department of Education

Technology and Security Subcommittee

Antelope Valley Union High School District

- Formeka Dent, District Data Technician

Association of Independent California Colleges and Universities

- Helen Norris, Vice President and Chief Information Officer for Chapman University

Bureau for Private Postsecondary Education

- Jason Piccione, Deputy Director and Chief Information Officer

California College Guidance Initiative

- Ben Baird, Director of CaliforniaColleges.edu Data Infrastructure

California Commission on Teacher Credentialing

- Andy Manguia, Information Technology Supervisor
- Don McGillivray, Information Technology Specialist

California Community Colleges Chancellor's Office

- Barney Gomez, Vice Chancellor, Digital Innovation and Infrastructure Division
- Daryl Lal, Information Security Specialist

California Department of Education

- Alan Nakahara, Manager for the Application Development and Maintenance Office of the Technology Services Division
- Rodney Okamoto, Director for the Technology Services Division

California Department of Social Services

- Karissa Vidamo, Acting Information Security and Privacy Officer

California Department of Technology

- Janet Buehler, Enterprise Architect
- Vitaliy Panych, Acting State Chief Information Security Officer
- Michele Robinson, Manager for Security Risk and Governance

California Health and Human Services Agency

- Adam Dondro, Agency Chief Information Officer
- Lloyd Indig, Agency Information Security Officer

California Labor and Workforce Development Agency

- Douglas Leone, Agency Information Security Officer

California School Information Services

- Greg Scull, Information Systems Officer

California State University Office of the Chancellor

- Subash D'Souza, Director of Cloud Data Engineering
- Ed Hudson, Chief Information Security Officer

California State Board of Education

- Sara Pietrowski, Policy Consultant

California Student Aid Commission

- Gurinder Bains, Chief Information Officer

CORE Districts

- Noah Bookman, Executive Director

Educational Results Partnership

- Dan Lamoree, Principal Data Architect

Employment Development Department

- Todd Ibbotson, Information Security Officer

Modesto Junior College

- Jenni Abbott, Dean of Institutional Effectiveness

University of California Office of the President

- Matthew Linzer, Information Security Manager
- Hooman Pejman, Data Architect

WestEd Facilitation Team

- Lupita Alcala, Director of Education Policy and Outcomes
- Rachel Antrobus, Senior Program Associate
- Rhegille Baltazar, Program Coordinator
- Allie Bollella, Program Coordinator
- Kathy Booth, Project Director for Educational Data and Policy
- Kathy Reeves Bracco, Consultant
- Erin Carter, Web Producer/Administrator
- Jessica Chittaphong, Program Coordinator
- Rosemary De La Torre-Trigueros, Program Assistant
- Neal Finkelstein, Senior Managing Director
- LeAnn Fong-Batkin, Senior Project Manager
- James Gilroy, Program Coordinator
- Melissa Josue, Administrative Assistant
- Mansi Master, Operations Analyst
- Marion McWilliams, Legal Counsel
- Michael Medina, Graphic Designer
- Sara Miller, Research Associate
- Baron Rodriguez, Director of Privacy and Data Security
- Marisela Sifuentes, Research Assistant
- Ravinder Singh, Senior Program Associate
- Cameron Sublett, Senior Research Associate

- Amy Supinger, Consultant
- Noel White, Senior Editor