In 2019, California enacted the Cradle-to-Career Data System Act (Act), which calls for the establishment of a state longitudinal data system to link existing education, social services, and workforce information. The Act also lays out a long-term vision for putting these data to work to improve education, social, and employment outcomes for all Californians, with a focus on identifying opportunity disparities in these areas.

The legislation articulated the scope of an 18-month planning process for a linked statewide longitudinal data system. As part of this process, the Governor’s Office appointed members to the Practice & Operations Advisory Group, whose charge is to ensure that the state data system addresses improvement efforts implemented at the level of individual institutions or regional partnerships, including tools that would be useful to students, families, and teachers. A separate Policy & Analytics Advisory Group was also appointed to ensure that the California data system is designed to support research, evaluation, public accountability, and optimization of statewide investments in education and related services. Input from both advisory groups will be used to inform final recommendations by a Workgroup consisting of the partner entities named in the California Cradle-to-Career Data System Act. These recommendations will form

1 Read the California Cradle-to-Career Data System Act at: https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=EDC&division=1.&title=1.&part=7.&chapter=8.5.&article=

2 Learn more about the planning process at: https://cadatasystem.wested.org/

3 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 Health and Human Services Agency, California School Information Services, California State
the basis of a report to the legislature and shape the state data system designs approved by the Governor’s Office.

This brief provides a discussion framework for members of the Practice & Operations Advisory Group. It includes background information on the rationale for a state data system and the authorizing legislation, in addition to five operational use cases that the state data system could support, based on priorities identified by the partner entities. Each use case describes key desired features and provides examples of how these use cases have been implemented in other states or on a limited scale in California. The brief concludes with framing questions that the advisory group will consider at their first meeting, in order to recommend which use cases should be prioritized in the first phase of state data system development.

Background

The Case for a Statewide Data System to Serve Practice and Operational Purposes

“Without access to such a system, students, families, educators, policymakers, and the public lack critical information with which to evaluate programs and interventions, illuminate roadblocks and solutions, inform decision-making, and address equity gaps.”

The Education Trust-West, 2019

Most states maintain a longitudinal data system, also known as a P20W data system because it links together information on individuals’ participation in preschool, K–12, postsecondary, and the workforce (Perez, 2016). The breadth and accessibility of these systems vary—some states also include information from health and social service agencies to provide a deeper understanding of the factors that can shape student outcomes. Other states have linked only K–12 and postsecondary data and have had
little success in leveraging this information to address state priorities (Armstrong, 2017). California is one of only eight states that have not yet created a linked data system, despite the fact that researchers, advocacy groups, policymakers, and system stakeholders have been requesting one for over a decade (Jackson & Cook, 2018) (see box on page 5). In interviews and survey responses, stakeholders throughout California have asked that a statewide data system serve not only as a mechanism for research and policy analysis, but that it include tools that can be used by educators, counselors, and other direct service providers, as well as public-facing tools that can be used by students and families.

Supporters of a state data system note that although California is rich in information, its data systems are disconnected. Because it is difficult to follow students across education and social service systems, critical questions are often left unanswered. While stakeholders differ in how they think the state data system should be designed, they generally agree that, at a minimum, it should

- produce information that contributes to a public good;
- protect student privacy;
- decrease rather than create additional burdens for system partners; and
- address equity gaps.

In the absence of a state data system, many attempts have been made to link data across sectors for practice and operational purposes. Several of these efforts involve independent partner organizations or public/private partnerships collecting student-level data and matching the information across institutions and sectors (Moore, Bracco, & Nodine, 2017). For example, the California College Guidance Initiative (CCGI) connects K-12 and postsecondary data that is used for many purposes, including support for counselors helping students to prepare for college, identification of transcript errors that can interfere with students' college and financial aid eligibility, and facilitating the electronic sharing of transcripts to inform decisions about admissions, placement, guidance, and financial aid. California Partnership for Achieving Student Success (Cal-PASS Plus) links educational records for community college students with

4 Learn more about CCGI at: https://foundationccc.org/What-We-Do/Student-Success/California-College-Guidance-Initiative
transfer, employment, and labor market information.\textsuperscript{5} A number of regions have linked education data to allow them to better understand student outcomes and work toward program improvement. For example, the Silicon Valley Regional Data Trust has linked education data in order to examine the impact of early education on primary school reading levels and absenteeism.\textsuperscript{6} In another example, a K-12-Higher Education partnership in Fresno has developed data tools that provide counselors with student-level information to inform the guidance they provide to students (Haxton & O’Day, 2015). In addition, state agencies have signed on to numerous memoranda of understanding (MOUs) that allow targeted data sharing, such as between the California Department of Education (CDE) and the California Department of Social Services to identify youth who are in the foster care system (California Department of Social Services, 2016).

While public/private and regional data-sharing efforts have produced valuable information and should not be replaced by a state data system, they entail significant costs and risks that a state data system could help address. For example, partners may spend months or years negotiating the legal framework for data sharing, which is costly and means that privacy protections may be unique to each negotiated agreement, rather than being implemented consistently statewide. Equally important, resource-strapped public entities may be unable to support the state-of-the-art security systems needed to ensure that student records are not compromised. Furthermore, it is not cost-effective to have data-sharing efforts duplicated in different regions, or to have entities in those regions searching and competing for additional funding, in the attempt to replicate successful models. Finally, because many data-sharing systems are developed to support a specific initiative or are dependent on grant funding, they may be short-lived (Moore & Bracco, 2018).

Stakeholders throughout California have asked that a statewide data system serve not only as a mechanism for research and policy analysis, but that it include tools that can be used by educators, counselors, and other direct service providers, as well as public-facing tools that can be used by students and families.

\textsuperscript{5} Learn more about CalPASS Plus at: https://www.calpassplus.org/About
\textsuperscript{6} Learn more about the Silicon Valley Regional Data Trust at: https://www.svrdt.org
Research, Policy, and Advocacy Organizations Call for the Establishment of a State Longitudinal Data System in California

The following list consists of selected publications calling for a state data system in California, with links to each report, where available.

- Association of Independent California Colleges and Universities: Report on Phase One of a Planning Grant on California Intersegmental Data and Postsecondary Educational Metrics
- California Competes: Out of the Dark: Bringing California’s Education Data Into the 21st Century
- Campaign for College Opportunity: Building a Student-Centered Data System in California
- Education Insights Center: California Education Policy, Student Data, and the Quest to Improve Student Progress
- The Education Trust-West: Data for the People campaign
- Policy Analysis for California Education: Making California Data More Useful for Educational Improvement
- Public Policy Institute of California: Increasing the Usefulness of California’s Education Data and Modernizing the State’s Education Data System
- Senate Select Committee: Longitudinal Data Systems

The California Cradle-to-Career Data System Act

In 2019, California enacted the Cradle-to-Career Data System Act which outlines the scope of an 18-month planning process for a P20W system, allocates $2 million to

7 https://californiacompetes.org/publications/out-of-the-dark
8 https://collegecampaign.org/wp-content/uploads/2019/05/Longitudinal-Data-FINAL.pdf
9 http://edinsightscenter.org/Publications/Research-Reports-and-Briefs//ctl/ArticleView/mid/421/articleId/2198/California-Education-Policy-Student-Data-and-the-Quest-to-Improve-Student-Progress
10 https://west.edtrust.org/dataforthepeople/
12 https://www.ppic.org/content/pubs/report/R_813PWR.pdf
support that process, and earmarks an initial $10 million toward the development of a state data system.

The Act also lays out a long-term vision for putting data to work to improve outcomes for all Californians, with a focus on identifying disparities in opportunities. By securely linking data that schools, colleges, social service agencies, financial aid providers, and employers already collect, the data system will

- enable users to identify the types of supports that help more students learn, stay in school, prepare for college, graduate, and secure a job;
- provide information that teachers, parents, advisors, and students can use to identify opportunities and make decisions;
- help agencies plan for and improve education, workforce, and health and human services programs; and
- support research to ensure policy effectively supports individuals from birth through career.

Recognizing that the data system will need to be built in phases, the Act lays out several priorities:

- **Linking existing information in the system.** The first data sets to be linked should be existing K–12 and college data sets, followed by employment and earnings data, early childhood education information, and social services information, although this order can be amended if the advisory groups and partner entities identify different priorities. Included in this priority is the need for the new data system to disaggregate information by several student characteristics in order to identify equity gaps.

- **Guaranteeing privacy and security.** The system cannot be built until clear guidelines and legal agreements have been established to ensure that information will be securely gathered and stored in compliance with federal and state laws and in accordance with privacy best practices, that the identity of sensitive populations such as undocumented Californians will be protected, and

14 https://www.documentcloud.org/documents/4638909-Staff-Report.html
that an appropriate managing entity has been identified to control access to data.

- **Providing information for students, families, and educators.** The system will include an interface for sharing information with teachers, parents, advisors, and students to support decisions. An advisory group that includes educators, community organizations, and advocates will help to identify which types of information would be most useful for these groups.

- **Facilitating analyses for researchers and policymakers.** The legislation identifies six priority areas that will be examined using data from the new system.

- **Assuring quality.** The legislation addresses the need to improve the quality and reliability of education information, both within and between agencies and other entities providing data. This improvement effort will include creating a single identification number for each student to be used across all public entities; establishing consistent data definitions; and developing processes for correcting data. The Act also notes that new types of information may be needed to answer pressing questions.

Once the data system has been built, and data security and privacy safeguards are in place, the state data system can be leveraged to streamline service delivery. For example, in other states, data systems automatically identify students who are eligible for such benefits as free or reduced-price lunch. When information can be shared between agencies, families do not need to fill out as much paperwork to access social and financial supports, and they can be informed about the full range of options available to them.

In addition, data could be shared with entities and partnerships that are using information to support local and regional efforts. For example, if a collaborative of high schools and colleges in a large urban area is working together to improve college advising for low-income high school students, it could obtain information about whether students from the region enrolled in college in other parts of the state. Or, educators and employers in a rural region who are designing stronger pathways toward emerging careers could determine whether students in those programs are getting jobs and making living wages.
Practice and Operations Tools

Based on feedback from the partner entities, several potential use cases emerged for a state data system:

- Automate eligibility for coursework, social services, and financial aid to remove barriers and ensure that more students receive support;
- Provide support to individual students by providing student-level data that can be used to better coordinate services and education planning as students transition between institutions;
- Create electronic transcripts that help students transfer their education records as they move from one institution to another and plan for a career;
- Combine academic and social service information to better understand education and employment outcomes; and
- Improve program design through planning tools and technical assistance that integrate intersegmental information.

The sections below describe each use case, including summarizing what stakeholders are requesting, with some illustrative examples of the kinds of functionality desired. The section also highlights examples of how entities in California, as well as across the country, are currently utilizing linked data for these purposes.

Automate Eligibility

One potential purpose of a state data system could be to provide colleges and universities with quicker and more reliable access to information on a student’s high school grades, course-taking patterns, and test scores to determine the appropriate placement of students for first-year mathematics and writing courses. New policies require both the California community colleges and the California State University (CSU) system to use multiple measures rather than a single high-stakes test to determine whether students are likely to succeed in transfer-level math and English courses. Connected systems that automate the upload and analysis of multiple measures data could ease the burden on institutions to make these determinations.

Another type of automation could be to remove barriers to support services by eliminating the need to navigate eligibility requirements and to submit duplicative
paperwork. The burden on students and families—particularly those facing a higher degree of socioeconomic challenges—would be reduced and access to services that foster stronger education and employment outcomes would be enhanced if application tools notified students about their eligibility for other programs. For example, students receiving Cal Grants could be alerted if they also qualify for Cal Fresh and request that their eligibility information be entered into a Cal Fresh application form.

In California, the Multiple Measures Assessment Project (MMAP)—a collaborative effort led by the Research and Planning Group for California Community Colleges (RP Group) and Cal-PASS Plus—developed a placement tool that uses multiple measures. Researchers matched student-level success rates in transfer-level math and English courses to high school data including GPA, grades in math and English courses, highest level of math and English courses taken, and various test scores. California community colleges that used multiple measures to increase the number of students eligible for non-remedial coursework saw significant gains in the number of Latinos and African Americans who completed transfer-level math and English courses in the first year of college (Mejia, Rodriguez, & Johnson, 2019). A state data system could help to streamline the process of updating placement recommendations as more colleges adopt multiple measures.

Other states such as Arkansas have leveraged their data systems to increase the number of students receiving financial aid. The Academic Challenge Scholarship provides funding for higher education, however, many eligible students were unable to secure aid because they had to provide a high school transcript as part of the application process. Now, when students submit the online application form, transcript information is automatically transferred from the Arkansas state data system and students are immediately informed whether they meet the requirements necessary to receive the grant. Not only did this innovation increase the number of students who received aid, postsecondary enrollment also increased.16

State data systems that link education information to multiple social service data sets have been able to expand the number of students receiving benefits. For example,

15 Learn more about MMAP here: https://rpgroup.org/All-Projects/ctl/ArticleView/mid/1686/articleId/118/Multiple-Measures-Assessment-Project-MMAP
16 Information provided in personal communication with Paige Kowalski, January 6, 2020.
since 2008–09, the federal government has required states to remove barriers to participation in school-based free breakfast and lunch programs by using a direct certification process with the Supplemental Nutrition Assistance Program (SNAP). Many states use centralized state education data as part of the match process but have not been able to identify all eligible students using just education and SNAP data. In 2016–17, New York City expanded its match process to include Medicaid data, which resulted in more than a quarter of a million additional children participating in federal nutrition programs. As a result, Medicaid data will be included in direct certification across the state (Food Research and Action Center, 2018).

**Provide Support to Individual Students**

A second category of tools that stakeholders are requesting of a state data system would enable service providers to deliver more timely supports and better coordinate services for individual students. For example, the tool could provide information to counselors that would guide recommendations about appropriate services, clarify college costs based on financial aid eligibility, and identify majors at a nearby college that align with a desired job. Providing a more comprehensive picture of the supports that students are accessing, in the context of their education pathways, can also help educators and providers understand factors that influence student success.

In California, the Silicon Valley Regional Data Trust (SVRDT) is a collaborative research organization. A secure system enables data-sharing across three counties (San Mateo, Santa Clara, and Santa Cruz) and multiple public agencies, including K–12, behavioral health, child welfare services, and juvenile justice. SVRDT uses cross-disciplinary data to inform school and agency staff who directly serve children or manage youth programs, to educate policymakers who govern schools and agencies, and to support researchers who partner with practitioners and policymakers. By applying advanced data analytics, SVRDT aims to create a personal blueprint for students that clarifies the myriad factors influencing their lives and improves the effectiveness of services and academic outcomes for all children, but especially for children of poverty. A state data system could allow similar analyses in additional counties.

17 Learn more about the Silicon Valley Regional Data Trust at: https://www.svrdt.org
Georgia has embedded information from the state data system into local K-12 student information systems, which allows teachers to access student-specific information on grades, support services, and test scores, starting with early childhood interventions (Center for Educational Leadership and Technology, n.d.). They recently added postsecondary enrollment data that enables teachers to see outcomes for their students for two years after high school graduation. In addition, aggregated information from the state data system has been compiled on a website that helps employers learn about the number of students who have been trained in various disciplines in specific locations and for students to better understand the relationship of K-12 career education programs to regional labor markets.18

The Michigan Data Hub (MiDataHub) allows information to flow between local providers and the state in real time so that data can be integrated across systems, including assessment, special education, food services, and transportation information. The state aims to present the information back to educators in the form of dashboards and early alert systems (Project Unicorn, 2019). For example, MiDataHub was integrated with the KReady System, which provides access to children’s early learning assessments and preparation for kindergarten, so that teachers have easy access to information that could inform personalized learning plans.19

While several states and regions have developed systems that allow users to share data in real time, it is important to note that these are predicated on either: 1) requiring all education institutions to deploy the same student information system software and each social service agency use the same client services software, 2) the state negotiating with all software vendors and custom-build local applications used in the state to create plug-ins that allow for data to be exchanged, or 3) having local education institutions tag data in their information systems to align with a common data standard and frequently exporting this information to a state data system.

18 Learn more about Georgia’s system at: https://www.gadoe.org/Technology-Services/SLDS/Pages/SLDS.aspx and https://gacareerpipeline.gadoe.org/
19 Learn more about Michigan’s system at: https://www.midatahub.org/
Create Electronic Transcripts

California’s state data system has the potential to facilitate the electronic transmission of student transcripts, easing a burdensome process for both students and institutions. Electronic transcripts can help to ensure that student identification numbers are carried from one institution to another, speed up the receipt of transcripts at the end of an academic year, and reduce the need for verification of grades, courses, or graduation status on the part of staff at a receiving institution.

Transcript data can also be combined with other tools that enable students to explore education pathways, identify possible careers, and create financial plans. For example, students could determine whether they are candidates for selective universities based on their GPA, test scores, and course-taking patterns. They could also explore careers that align with the skills they gained in high school and learn more about the amount of education required to qualify for those jobs.

The California College Guidance Initiative (CCGI) streamlines the transmission of transcript information. Students can launch admissions applications for either the CSU system or community colleges from within the tool, which then automatically populates information on the student’s courses and grades. Because the information comes directly from the transcript, admissions officers do not have to separately verify the accuracy of the information with the high school or district. In addition, because the transcript is electronically transferred, both the CSU and community college systems have access to an individual applicant’s K-12 student identifier (SSID), which improves accuracy of data matches for research or service coordination purposes. Finally, CCGI integrates education and career planning tools.20 Currently, districts that are not part of subsidized pilots in the Central Valley and Inland Empire must pay to participate in CCGI. A state data system could support the scaling of this service to all schools.

North Dakota provides parents and students with electronic secondary school transcripts that can be shared with colleges and universities. In addition to providing families with access to real-time transcript information through a web portal, the North

20 Learn more about CCGI at: https://foundationccc.org/What-We-Do/Student-Success/California-College-Guidance-Initiative
Dakota system allows for secure, immediate, and cost-effective transfer of student records.21

**South Carolina** has used information from its state data system and a partnership with Parchment, a third-party digital credentialing service, to develop an e-transcript system that is free to South Carolina students and families. In addition to allowing South Carolina high schools to seamlessly transfer student transcript records to colleges and universities, South Carolina colleges and universities are able to transfer student transcript data to other postsecondary institutions.22

**Combine Academic and Social Service Information**

Many partner entities and advisory group members wanted to combine education, employment, social service, and health information. Linked data could be used to inform schools and colleges about the proportion of homeless or incarcerated students they serve, so that educators better understand the challenges faced by the people they serve and focus on providing students and their families with tools that address the areas of greatest need. Education and employment information would also be of value for health and social services—for example, medical providers, criminal justice authorities, and welfare and nutrition programs—to measure outcomes for service recipients.

The **California** CalWORKs Outcomes and Accountability Review (Cal-OAR) Data Dashboard contains real-time county and state-level information needed to analyze and understand the outcomes of the state’s welfare-to-work program. For example, users can see education attainment levels, earnings by education level, welfare-to-work activities, and wage progressions for CalWORKs participants. To make the data more actionable, the dashboard provides trends over time and displays results by race, ethnicity, gender, and assistance unit type.23 Linking Cal-OAR with a state data system

21 Learn more about the North Dakota system at: https://www.nd.gov/dpi/parentscommunity/etranscripts

22 Learn more about the South Carolina system at: https://ed.sc.gov/data/information-systems/power-school/sc-etranscripts/

23 Learn more about the Cal-OAR dashboard at: https://www.cdss.ca.gov/inforesources/calworks/cal-oar/cal-oar-data-dashboard
would help to provide additional information about education milestones for students who participate in CalWORKs.

Researchers in Washington state linked information from the state data system to juvenile justice information to create a number of research products, including the Juvenile Justice Dashboard. A public website allows users to compare high school and GED completion, postsecondary enrollment, community college and baccalaureate completion, employment hours, and earnings for students who had been involved in the justice system. Results are displayed for all students, as well as showing outcomes by gender and race.24

The District of Columbia manages a system that shares real-time information between education and homeless services agencies. Schools receive current information on which of their students are homeless. Service providers can base education placements on children’s grade levels and special education needs, without putting the burden on parents to collect this information. The data system has also been used by researchers to identify the number of homeless children who are not enrolled in school and by administrators to secure additional funding for homeless services (Institute of Education Sciences, 2016).

**Improve Program Design**

Another potential benefit of a state data system is the ability to provide information and support that can be used to strengthen instruction and service delivery. For example, high schools could receive feedback reports that share data on students who transition to postsecondary institutions, including where they enrolled, which major they selected, what types of academic supports they accessed, and whether they graduated. Information on students’ performance in the next phase of education or in a career can help educators better understand whether their programs and services are resulting in the desired outcomes, as well as whether there are equity gaps in educational attainment and employment.

24 Learn more about the Washington dashboard at: https://erdc.wa.gov/data-dashboards/juvenile-justice-dashboard
Many states have found that providing data tools is not sufficient to generate change at the institution level. Pairing information with support services such as technical assistance and planning tools can support stronger data use among educators, while simultaneously providing a feedback loop that helps ensure data tools are responsive to the needs of their users (Institute of Education Sciences, 2019).

In California, eight school districts have formed a collaborative called CORE, including Los Angeles, Long Beach, Fresno, Santa Ana, San Francisco, Garden Grove, Sacramento City, and Oakland. The CORE districts built and maintain a comprehensive school improvement system that provides access to information that is not collected by or available through the state. School and district profiles include data on student-level academic growth, middle-school indicators of high school readiness, students’ social-emotional skills, and school climate measures, along with traditional measures of test scores, graduation rates, and absenteeism. Districts can compare their schools’ performance to that of similar schools across the state to clarify strengths and challenges in the different areas. CORE also convenes member districts regularly to share findings and improve data use.²⁵ A state data system could store valuable information to supplement information captured by the California Longitudinal Pupil Achievement Data System (CALPADS), the data system maintained by CDE.

The Minnesota Department of Education uses data from its Early Childhood Longitudinal Data System to provide technical assistance for early care programs, such as helping providers reach new audiences and evaluate long-term outcomes for children. In addition, information from Minnesota’s Statewide Longitudinal Education Data System is used to inform school district planning and curriculum decisions. For example, the Saint Paul public schools incorporated performance measures from the state data system into its strategic plan to gauge how well its diverse graduates are prepared for college, career, and life (National Center for Education Statistics, n.d.).²⁶

Tennessee pulls data from the state data system into a centralized location for use by academic institutions, with a focus on strengthening local planning processes. InformTN includes information on academic achievement and growth, college and career

25 Learn more about the CORE districts at: https://coredistricts.org/our-work/improvement-communities/data-collaborative-community/
26 Learn more about the Minnesota data system at: http://sleds.mn.gov/
readiness, as well as non-academic information related to school climate and access. School and district leaders can build plans within the tool and see data visualizations that provide feedback on priority measures, including outcomes for at-risk students (Data Quality Campaign, 2019).

**Action Steps for the Practice & Operations Advisory Group**

A state data system cannot solve every problem, but it does have the potential to include tools that can provide valuable information to students, families, educators, counselors, and other service providers.

Given the wide range of uses for the California data system, a key challenge in designing it is to keep the vision expansive enough to meet critical needs, while also understanding how best to stage system development—to identify what can be done in the short term and what steps must be taken to reach longer-term goals. Interviewees and experts who provide support for state data systems note the importance of finding “low hanging fruit” that can be harvested in the first phase of system development to generate momentum and demonstrate the usefulness of the work (Clark et al., 2016).

In considering how to best phase in the priorities, advisory group members will be asked to consider the guidance provided in the legislation, the short- and long-term utility of the priorities identified by the partner entities, and the complexity required to implement each priority. The following questions can be helpful in identifying what the first phase of development should focus on:

- What are the key user needs that should inform the first phase of system development?
- Which use cases would you prioritize to address the most critical needs?
- Are there existing tools that could/should be expanded for operational use statewide?
- Are there tools that could/should be piloted on a smaller scale initially to test efficacy?

---

27 Learn more about the Tennessee planning system at: https://edplan.tn.gov/public/plans/
• What additional use cases need to be anticipated as part of the first phase of design?
• What would change if the first phase of the system were in place?
• Would all students benefit from this change?

The rationale for building a state data system that links information from various education, employment, and service entities is clear, particularly when considering the types of data tools that have already been built in California as well as in other states. The priorities recommended by the Practice & Operations Advisory Group will help ensure that the new data system will provide valuable tools that will be of use to a broad set of stakeholders, and that the state data system builds upon successful work already underway to foster more equitable outcomes for all Californians.
References


