



Getting Down to **FACTS**



Education Data Needs, Availability, and Access in California

Jon Fullerton
University of Southern California

May 2026



Stanford | SCALE Initiative
Accelerator for Learning

Education Data Needs, Availability, and Access in California

Jon Fullerton

University of Southern California

Over the past 25 years, schools, social service providers, and other government agencies have adopted modern information systems that capture electronically much more data about students, teachers, and schools than previously. Properly linked across levels (e.g., early childhood, K-12, postsecondary) and systems, this data can support improved educational outcomes in multiple ways. If district and school leaders can gain insights into the success of their students *after* they graduate, this can help them determine whether their educational program and career guidance is providing proper preparation. Researchers can evaluate whether programs and policies have been implemented and whether they improve the outcomes they intend to. The public can monitor system performance and press for improvements when and where necessary – identifying equity gaps in either inputs or outcomes. Finally, individual students and their families can save time and have more opportunities if properly connected data systems can guide choices and simplify processes across governmental systems – for instance, by automatically transmitting a student's transcript to colleges to which she is applying.

The Data Quality Campaign has noted the promise of linked state systems to ease students' navigation through college and career pathways, to connect students to needed supports, to help students and job seekers understand pathways into particular fields or jobs, and to allow researchers to determine what works and what does not (2023).

Relatedly, the Gates Foundation and Mathematica have created the *Education-to-Workforce Indicator Framework* that synthesizes prior research and frameworks to identify a set of critical questions and data points that should be collected to track and improve student pathways through their education and into employment. Collecting such data will allow policymakers and educational system leaders to monitor progress and improve policy with the goal of "advancing educational and economic opportunity for all." (Gonzalez et al., 2022).

Research using data linked across systems can have real impacts on major policies that might not otherwise be attainable. For instance, a research team in Michigan performed an experiment to test whether offering students from low-income families free tuition to the University of Michigan if admitted would increase application rates from these students. The researchers saw a large jump in both application and enrollment rates in the targeted student group (Dynarski et al., 2021; Burland et al., 2023). Partly as a result of the evidence provided by this and other studies, other higher ed institutions across the nation have begun guaranteeing full, blanket scholarships to students whose families fall below a set income level (see, for instance, Krupnick, 2024). But the data requirements to conduct such work are substantial. To conduct this type of research, one needs to be able to connect K-12 data administrative data to social services data to college application data to financial aid data to college enrollment data.

This report focuses on the evolving state education data landscape in California and its ability to meet the needs highlighted above. It highlights areas of substantial improvement since *Getting Down to Facts II* in 2018 as well as some ongoing gaps and challenges that remain. It will also suggest some potential changes to the state data infrastructure that could advance the ability of state data to support educational improvement and better outcomes for individual students.

A Data Use Typology

To evaluate the usefulness of California's data systems for improving educational outcomes, this report first explicates some key dimensions of data use. Specifically, a discussion of data needs must take account of different potential users, the types of uses to which they will put the data, the granularity of data required for these uses, and, finally, the required timeliness of the data.

Users

Different users naturally have substantially different things they need from education data. Below is a partial list of users and some of the uses to which they put education data.

- Students and families have a wide range of needs, from selecting schools to tracking their own academic progress to applying to college.

- Teachers, counselors, and school staff need detailed feedback on student academic performance, attendance, transcripts and courses taken, as well as insights into student progress in applying for college.
- Districts may want to look at larger trends, track and evaluate the outcomes of their improvement strategies, understand incoming students' needs and preparation, and monitor graduating students' later success.
- The state needs to understand how students are doing, ensure that students are being served appropriately, and understand the implementation and outcomes of state policies.
- Advocacy groups and the public may want to know how local institutions are performing, what groups are being underserved, and how - and with what efficacy – tax dollars are being spent.
- Researchers need detailed data to evaluate the effectiveness of programs in a rigorous manner against appropriate comparison groups.

Types of Use

The multiple needs different users have for the data suggest three broad buckets of use:

- *Incent.* Data can be used to monitor performance/progress and provide incentives for improvement, for example, data used for student report cards and data used in school and district dashboards.
- *Evaluate.* Data can be used to inform the development of policies or programs and then to evaluate their effectiveness.
- *Empower.* Data can be used as feedback to support and guide students, schools, or districts in their decision making. For instance, data monitoring progress towards completing college requirements can help guide students in their course selection.

Granularity of Data

Specific uses of data require different levels of granularity in the data.

Many uses need only *aggregate data* – that is, data reported out only for groups. Aggregate data allows for progress monitoring and can show important trends within and across groups. To protect privacy, data providers often mask reported data for "small cell sizes," i.e., when there are so few

members in a group that there is a possibility individuals (or a small number of individuals) could be identified. The vast majority of publicly available education data is aggregate data.

However, aggregate data is not helpful for many uses. Students tracking their progress towards completing college entrance requirements (called the A-G requirements in California) need their own individual transcript, not a group summary. *Individual data* contains Personally Identifiable Information (PII) that is necessary to link the information to the right student or person. Because of the presence of PII, any sensitive data must be strictly protected, with data restricted to those who need the information to provide services (e.g., teachers, counselors) or the individuals and families themselves. This data is typically governed by privacy laws such as the federal Family Educational Rights and Privacy Act (FERPA), the Health Insurance Portability and Accountability Act (HIPAA), or California's Student Online Personal Information Protection Act.

Sometimes uses require individual-level data, but there is no need for the user to know who the specific individuals in the data set are. For instance, researchers doing impact evaluations often need to track individual students who receive or do not receive an intervention over time to estimate impact, but they do not need to know who any given student in the data is. In these cases, researchers can use *de-identified data* – that is data from which PII has been stripped and replaced with another identifier that allows the individual to be tracked across time and relevant data sets. De-identification provides some privacy protection, but, with enough data fields, a skilled analyst could potentially re-identify individuals. Such data are therefore typically restricted in access and governed by data use agreements that are in compliance with the laws mentioned above.

Timeliness of data

How "fresh" data needs to be depends upon the use.

Operational data is fresh data that gets automatically updated with the latest information. Operational data is expected to change frequently. This type of data might be found in student information systems which track absences and grades in real time or in an Operational Data Store (ODS) that brings together real-time operational data from multiple sources. Operational data can be used to track individual student progress (for instance, a student's class grades) or to identify emerging trends across a school system (for instance, absence patterns across schools for the first half of the semester).

Static data is historical data that gets refreshed periodically (e.g., once year, once a semester). Typically, such data goes through an ETL (extract, transform, load) process to validate and clean it, and then it is stored in a data warehouse. This type of data allows for evaluation of trends over time, impact evaluations, and policy analysis.

As table 1 below highlights, data system requirements in terms of granularity and timeliness will vary tremendously across specific use cases.

Table 1: Sample Use Cases

	User	Use	Granularity	'Freshness'
Impact evaluation	Researcher	Evaluate	Deidentified data	Static
Implementation study	Researcher	Evaluate	Deidentified data	Operational
Advocacy group pushing for improvement in district outcomes	Public	Incent	Aggregate data	Static
District looking to improve college and career outcomes	District leadership/ district analysts	Empower	Deidentified data	Static
Advising student on graduation pathways	Counselor	Empower	Individual data	Operational

We now turn to the data systems that California offers and evaluate how well they can meet these varied needs.

The California Landscape

Education data for school and state administrators

School districts and charter schools (collectively, Local Education Agencies or LEAs) all use student information systems (SIS) to store much of their day-to-day operational data on students (i.e., class assignments, grades, attendance, etc.). In California, providers - such as Aeries, Powerschool, and Aequitas - vary across districts, and different providers have slightly different data structures and capabilities. SISes allow teachers, principals, and district staff to pull up information about individual

students or groups of students, and most SISes also allow LEAs to create secure parent or family portals where parents can get real-time information on their child's grades and progress.

While the SIS may be the core administrative system, LEAs typically must manage numerous others as well. In California, Early Education SIS systems are generally separate from the K-12 SIS systems, capturing and reporting different data. Large districts which serve as their own Special Education Local Plan Area (SELPA) will have separate systems to manage special education plans and services such as the Special Education Information System (SEIS) or Welligent.¹ Information about teachers, employees, and finances are captured in separate enterprise resource systems such as PeopleSoft or Workday. LEAs that use interim assessments or use educational software for some instruction will often have to deal with separate systems for each. LEAs that survey students or attempt to measure socio-emotional wellness and/or school climate will have yet another data stream to consider. LEAs also leverage Learning Management Systems (such as Canvas or Google Classroom) that capture student work.

Faced with these multiple systems, larger districts may create data warehouses or operational data stores that pull together and link information from multiple data sources to make it easier for teachers and administrators to access and analyze relevant student or program data from across systems. While some districts create these warehouses themselves, third-party providers such as Panorama Education, Renaissance Learning, or Education Analytics also work with districts to link student data across systems and provide actionable insights in real time.²

County Offices of Education (COEs), another essential layer of the data system in California, can provide their member LEAs substantial information and data services – from basic SIS services (e.g., the Q SIS Consortium in Ventura COE) to data that districts might not otherwise have available (e.g., graduate college attendance from the National Student Clearinghouse in San Bernardino COE) to data warehouses, data analytics, and dashboards (San Diego COE). As noted elsewhere in this series, what county offices of education offer varies quite widely across counties (Trinidad, 2026), one result of

¹ Regional SELPAs serve as the administrative hub for special education services to multiple smaller districts.

² Artificial intelligence tools promise to greatly expand the capacity of schools to extract individualized information from LMSes and across systems over the coming years. See Gonzalez et al. (2026) in this series for the possibilities (and challenges) for AI to provide "learning experiences that matter."

which is that districts have different access to data depending on where they are geographically located in the state.

With multiple systems across many providers, data is not standardized at the operational level. However, the California Department of Education's (CDE) *California Longitudinal Pupil Achievement Data System* (CALPADS) requires districts to upload individual student, teacher, course, and program data in a standardized format. Operational since 2009, CALPADS is a data warehouse and the primary statewide data source on K-12 students. The state uses CALPADS for many things – from determining program compliance to reporting data to the federal government to calculating accountability measures to reporting publicly available statistics about districts and schools. As part of the move to CALPADS, the state adopted a Statewide Student Identifier (SSID) in 2005-06. This allows for students to be tracked across LEAs in the state and is essential for calculating measures such as the 4-year cohort graduation rate even if students move or transfer over the relevant time period.

The data set is rich enough that CALPADS also serves as a historical data warehouse for districts (they can only see their own data, of course). However, much of the data that districts collect does not get captured in CALPADS. For instance, student marks in elementary school are not part of the data collection. Nor is student performance on non-state assessments. Nor are measures of social-emotional well-being. Districts are on their own to integrate and analyze such data.

While CALPADS is the central data repository for the CDE, the state collects other data as well. The *California Basic Education Data System* (CBEDS) requires LEAs and Counties to provide information about their schools and districts as well as aggregate data on students and staff. The CDE also collects adopted budgets and detailed financial information from LEAs, IEP implementation data, and preschool and childcare data for programs it administers (the CAPSDAC, CDMIS, and PLIS systems).

California's three public higher education systems (the University of California System, the California State System, and California Community Colleges) all have their own, separate data systems. These have not been connected in the past, though the Cradle-to-Career data system (discussed below) promises to smooth some data sharing between these systems for students who transfer across them. The California Student Aid Commission maintains data on financial aid eligibility and use for students in California public higher education.

Public Facing Data

The state provides several publicly accessible data sources that allow users to examine aggregate historical data from the above systems.

- *DataQuest*, hosted by the CDE, provides aggregate historical data at the school, district, county, and state levels, leveraging CALPADS as the primary data source. The database provides a wide range of data, from enrollment and assessment results to graduation and college attendance rates to staffing patterns.
- *Ed-Data*, a partnership between the CDE, EdSource, and the Fiscal Crisis and Management Assistance Team/California School Information Services (FCMAT/CSIS), offers much of the same data found in DataQuest, but with the addition of financial data and some additional data points. Ed-Data also uses graphs for a more intuitive presentation of the data than the tables found in Dataquest.
- *The California School Dashboard* is intended to provide information on school and district performance and progress to parents and educators. The metrics used are sometimes different from those in DataQuest. For instance, the Dashboard reports on student performance in terms of "distance (in scaled score points) from standard" while DataQuest reports percent meeting standard. (See Haderlein and Polikoff, 2026).
- *The School Accountability Report Card* provides additional information on school performance and resources, again using somewhat different measures than those in the School Dashboard.
- *The California Cradle-to-Career Data System (C2C)* is a new data system that will provide aggregated longitudinal data that cross system boundaries to the public. Currently, only dashboards on pathways to and through college are available to the public; however, many more dashboards are planned. C2C is an extremely important development in California's data ecosystem and will be discussed in detail below.

Somewhat confusingly, the first four sources contain much of the same underlying data but the data are reported in slightly different ways across the sources. Consolidating this data and providing more explanations on how to do a crosswalk between two measures of the same academic outcome might allow for better access and understanding by the public.

Student- and Family-Facing Data

Students and parents also have access to some operational data from education agencies to help them navigate the school system and pathways to higher education. As mentioned above, most school districts have a parent portal that is powered by their SIS, allowing parents to track their children's progress.

The California College Guidance Initiative (CCGI) is an operational data system that connects district data systems to state institutions of higher education (IHEs) and the California Student Aid Commission. The tools CCGI provides allow students and their parents to track their completion of California's A-G 4-year college requirements, apply to public universities (automatically sending transcript data), and apply for financial aid. Schools and counselors can use the system to understand their students' progress against the above and track student trends in college readiness more broadly. CCGI was initially voluntary for districts to join and link to. However, as of June 30, 2026, all districts serving students in grades 9-12 are statutorily required to link their data to CCGI (CDE, 2025 October 3).

Data for research and improvement

As noted above, LEAs in California face multiple internal operational data systems that need to be standardized and linked to develop educational strategies, evaluate programs, and monitor progress. Ideally, LEAs would also be able to pull in data from other sources so that they could understand other systems that may be impacting student experiences and needs (e.g., social services data) and/or monitor the success of their students after they leave the district (e.g., college and workforce data). In addition, it is useful for LEAs to compare themselves to others, identifying common challenges and learning from those who have solved them.

In response to these needs, California has developed numerous research practice partnerships (RPPs) and district improvement networks that integrate district data, pull in data from across systems, and/or allow districts to look at data across multiple districts. These include:

- RPPs with individual districts. The Los Angeles Educational Research Initiative (LAERI), the Oakland Unified School District – UC Berkeley Research-Practice Partnership, the Stanford-San Francisco Unified District Partnership, the San Francisco Unified School District-UC Berkeley Research-Practice Partnership, and the San Diego Education Research Alliance are all examples

of long-term partnerships that connect university researchers with districts to work on mutually determined topics (National Network of Research-Practice Partnerships, 2025). Leveraging a carefully constructed research data warehouse of administrative data, LAERI, for instance, has worked with LAUSD on rigorous studies of college access and improving middle-school math policy (Miller, Phillips, and Ahearn, 2024; Miller and Phillips, 2025). Dee and Huffaker at Stanford have likewise been able to provide San Francisco with a rigorous evaluation of the effects of detracking math in the 9th grade.

- District improvement networks. The CORE districts are a collaborative of eight of the largest school districts in the state that have developed student surveys, dashboards, reports, and analytics that allow deeper insights than the state provided information discussed above. Districts beyond the eight largest may join the data collaborative and receive benefits as well. The Stanford-Sequoia K-12 Research Collaborative connects 1 high school district and its 8 feeder elementary school districts to improve outcomes across district boundaries. The Orange County Educational Advancement Network (OCEAN) partners UCI researchers with local school leaders to identify critical questions they have and to execute research to answer them. USC's Education-to-Workforce District Network works with districts in Southern California to improve academic and workforce outcomes for students by leveraging data and analytics (including data on post-secondary outcomes) and giving participating districts the opportunity to work and learn from one another. Likewise, UC Davis supports three regional K-16 collaboratives in Northern California working to improve educational attainment and align programs with workforce needs in their area (Hurt et al., 2026).
- Cross-sector partnerships. In prior years Cal-PASS Plus allowed districts and community colleges to sign an MOU to connect their student data to community college data allowing them to track students' success and progress in community colleges. While districts did not look at individual student data, they could track students through their community college experience. With the advent of the California Cradle-to-Career Data System, however, some of the linked data collection for Cal-PASS Plus ended in 2023 with much of the community college specific data migrated to the community college system's Data Vista platform.

Finally, there are also several research centers focused on state policy that integrate data across systems and segments. The *California Education Lab* is a research center out of UC Davis that focuses on college and career readiness and college access, connecting data from the CDE, California's three higher ed systems, and the California Student Aid Commission. The *California Policy Lab* is a research institute based out of UC Berkeley and UCLA that draws from faculty across the UC system to explore a range of issues using state (and other) administrative data. CPL has successfully connected anonymized data from the California Department of Social Services to college data from the three state systems.

As can be seen above, there is a great deal of work that is ongoing in the state to improve the insights LEAs can gain from data and to connect LEA data to other systems' data. However, what is available varies by region, the choices of local governmental agencies, and the ability to develop (and fund) research partnerships. In addition, LEAs have widely differing internal capacities to link and study data. While mega-districts such as LAUSD or Long Beach USD often have their own internal research and analysis capacity, smaller districts may have only one person (if that) to handle the bulk of reporting to the state, leaving no time for analyzing or using data. The end result of all these factors can be highly uneven access to cross-system data and the insights they can provide (Phillips et al., 2018). California's newly launching Cradle-to-Career Data System is intended to fundamentally change the availability of data for improvement and to provide those with less data capacity or expertise with insightful data stories that are easily understood.

The Cradle-to-Career Data System

The Promise and Progress of the California Cradle-to-Career Data System

For the previous iteration of *Getting Down to Facts* in 2018, researchers noted the complex data ecosystem in California discussed above and that California lagged other states in its multitude of unlinked data systems and in not having a comprehensive data system that could connect individual data across sectors (e.g., K-12, higher ed, workforce, social services). This lack of infrastructure has hampered California's ability to coordinate services across sectors/levels, to learn from and evaluate the impact of policy and practices, and to generate new knowledge about how to improve education in

the state (Phillips et al. 2018). The authors urged California to develop such an infrastructure, leveraging other states as examples.

In 2019 California passed the Cradle-to-Career Data Systems Act establishing a planning process for a longitudinal data system to solve this issue, linking education, social service, and workforce data. This Cradle-to-Career (C2C) data system would "be California's source of actionable data and research on education, economic, and health outcomes for individuals, families, and communities"(N.A. 2020). The system was to be used to:

- (A) Address disparities in opportunities and outcomes,
- (B) Support student guidance,
- (C) Foster continuous improvement, and
- (D) Address the needs of researchers.

In 2021 a Trailer Bill established an Office of Cradle-to-Career Data (California Cradle-to-Career Data System Act).

With this organization in place, work began on establishing a Participation Agreement with different governmental agencies providing data and setting up a new data system to link across departments, despite some skepticism (Walters, 2022). In 2022-23 the Cradle-to-Career Data System began ingesting historical data from departmental providers; in 2023-24 links across K-12, college, and career data points were instantiated; and C2C released the first public dashboard on student pathways through college in the Spring of 2025 (California Cradle-to-Career Data System, 2025).

The C2C Data System is governed by a 21-member Governing Board composed of representatives from the data providers and selected members of the public. The Governing Board has final say on what data points are included in the C2C database, data standards used, expansions of data to include to new providers, privacy and security policy, and oversight of the managing entity. Importantly, data providers (each of whom is represented on the governing board) "retain sole control over their source data and may reject, add, or remove data elements contributed to the P20W data set, as reflected in its participation agreement with the managing entity" (California Cradle-to-Career Data System Act,

10861). In other words, data providers retain ownership of the data, and what data is shared is ultimately up to the providers. The Governing Board thus plays a critical role in advancing the work publicly and collaboratively across agencies.

In addition to the Governing Board, C2C leverages an extensive series of advisory boards, committees, and community engagement meetings – almost all of which are public meetings. The goal is to develop a data system "by and for Californians," robust to changing needs in the state (C2C, 2026d).

As of March 2026, C2C has received more than 10 years of data from 18 data providers covering more than 250 data elements across 12 domains (California Community Colleges, 2026). (See Figure 1 listing data domains and providers.) In theory, researchers will one day be able to track students from preschool to K-12 to college and workforce training programs to employment. Districts will be able to compare the college outcomes and earnings of their graduates to others to better understand whether they are preparing students for postsecondary success and whether they are equally successful with all groups of students. The public will be able to understand the impact of early education on later success in school.

C2C envisions this database to power multiple products aimed at the different users discussed earlier in this paper.

Dashboards will be aimed at public users and provide specific "data stories" with a focus on clear, easy to understand data presentation. The first dashboard released is on student college pathways to college, allowing users to know what pathways to and through California public colleges students typically take, whether there are differences in pathways between different demographic groups, and labor market outcomes for college completers.³

Figures 2 and 3 show two example analyses from this set of dashboards. Figure 2 tracks whether students attend state higher ed institutions and their paths through the institutions – i.e., how many students who enroll in a two-year college graduate, transfer, receive a certificate, etc.. Figure 3 shows how long it typically takes students to graduate. Both figures are designed to be easy for users to understand. While Figures 2 and 3 show outcomes for California as a whole, the dashboard allows

³ C2C is working to expand coverage to include out-of-state college enrollment and private college enrollment in later versions.

users to examine the data for individual school or legislative districts and/or different student populations. To protect privacy, C2C aggregates the data used in the dashboards, with no data point with less than 11 members reported.

K-12 leaders and advocacy groups could leverage this data to understand which students are succeeding after graduation, compare their district results to others, and (as more years of data become available) monitor changes over time. Counselors could use this data to help students understand the labor market outcomes of an associate degree vs. a bachelor's degree in their geographic region.

C2C is also developing **Query Builder** functionality that will allow users to re-cut the data to perform analyses that are not in the pre-built dashboard displays. Such analyses might include looking at trends across multiple districts or connecting district level dashboard outcomes in a new way. For instance, the Public Policy Institute of California used query builder data to compare college completion across high school districts (Johnson and Mejia, 2025).

While Query Builder data are public, users of the Query Builder are likely to be somewhat more sophisticated in analytics than those at whom the dashboards are aimed. For instance, the initial "Query Builder" data released to support the student pathways dashboard consists of spreadsheet files containing separate rows of data for every district and population used in the dashboard. To cut the data in new ways, users need to know how to use either spreadsheet or statistical software at an intermediate level.

The data in the Query Builder will be the same as what underlies the dashboards, including the same rules of data suppression (no data point with less than 11 members reported).

Finally, a **Data Enclave/Research Database** is being created to which only approved researchers on approved projects will have access. Researchers using this database will have access to de-identified, individual student-level data, in an "enclave" environment (i.e., researchers will log into a C2C provided environment and perform any analysis there, with no de-identified, individual-level data leaving the C2C environment). C2C has contracted with NORC at the University of Chicago to provide this infrastructure and leverage NORC's existing expertise and processes (Bates, 2026).

While the processes for accessing this database are still being finalized, the broad approach will be as follows: Researchers who only need *aggregate data* will follow an expedited process. While procedures for releasing aggregate data are still under development, an early proposal was that the C2C office will review the request and send the request to the relevant data providers for approval. Upon approval, C2C will provide the relevant aggregate data to the researchers. Researchers who need *individual data* will fill out a request form that will be evaluated by each data provider and an Institutional Review Board. Data providers may go back and forth with the requester to ensure that the data requested is necessary. Upon approvals, researchers must provide proof of appropriate research training and sign legal documents ensuring confidentiality. The data approval process was initially expected to take around three months, though a pilot is underway to determine whether this is a realistic timeline. Finally, before publication, researchers must submit their reports for review to the data providers to ensure data privacy has been maintained in the published findings (C2C, 2026c).

The research database is intended for research professionals and academic researchers who need access to individual data and possess considerable methodological and programming sophistication.

The three products discussed thus far are all designed on a new, static data warehouse. Data providers will transfer data to the warehouse at regular intervals (generally once per year) and the data, once integrated into the warehouse, should not change save for any corrections or updates from data providers themselves.

However, such data is not useful for supporting all individual student guidance (a goal in the guiding legislation) around topics such as course decisions as it is updated at a far slower pace than students, counselors, and administrators need. Therefore, C2C is also partnering on the scaling of operational data systems such as the California College Guidance Initiative (CCGI) and the eTranscript California.

As noted above, CCGI is a set of college planning tools that allows students to identify career interests, apply to California public colleges (both community and 4-year) by transferring transcript and other data from districts to the colleges, and monitor their progress towards completing California college prerequisites (the A-G requirements). CCGI also allows counselors and teachers to monitor their students' progress against the A-G requirements and state college applications. CCGI thus connects *live*

transcript data from K-12 Student Information Systems with data from the application systems of California colleges. Initially, districts volunteered to take part in this system; however, as of 2026, all districts are statutorily required to establish a data sharing partnership (CDE, 2026). While limited only to California colleges, this system provides data in a timeframe that is useful for individual decision making.

eTranscript California is a system that allows community colleges and participating 4-year colleges to transfer transcript information in an automated manner. If a California high school student takes one or more courses at a community college, they are required to obtain transcripts from those community colleges as part of their postsecondary application process. As more California K-12 students take advantage of growing dual enrollment opportunities in their districts and at their local community colleges, the burden on students is growing. In addition, students must report what classes they have taken at community colleges to counselors to include that information in their A-G planning.

In 2024, C2C ran the eTranscript California & Career Passport task force that recommended a) integrating eTranscript California with district data systems to support better student guidance and reduce administrative burden and b) the creation of “Career Passport” that would capture both academic and workforce training information for students (C2C, 2024). In 2025 funding was appropriated to revamp eTranscript California and begin the implementation of these recommendations.

Potential Challenges

The emerging C2C Data System represents a tremendous opportunity to monitor and improve educational outcomes in the state, and the final system promises to be one of the most in-depth integrations of data across the 50 states (Data Quality Campaign, June 2022). Moreover, it is evolving rapidly, with new functionalities and procedures scheduled to come online over the course of the year (C2C, 2026a).

That said, C2C faces several potential challenges that could hold it back from having the maximum potential impact on education in the state. I divide these challenges into two major types – those of access and those of coverage.

Potential Challenges: Access

As noted above, C2C has created a data warehouse that links individuals across systems while masking their identities. To further protect individual privacy, in its public facing tools (dashboards and query builders), C2C imposes a minimum cell size of 11. In other words, if a particular reported outcome (for instance, "students who first enroll in a 4-year college and complete an associate degree" – presumably due to transfer) has less than 11 individuals in it, it will not be reported.

However, merely suppressing that one cell is insufficient. If a particular outcome combines two (or more) categories, an enterprising member of the public could recover the suppressed numbers by taking the total in a given outcome (for instance, "completed associates degree") and subtracting off the non-suppressed cells (for instance, "students who first enroll in a 2-year college and complete an associate degree"). Therefore, C2C deploys "complementary suppression" in which both the suppressed cell and the next larger cell are not reported publicly (C2C, 2026b).

While this works well for the state and the very largest districts, in practice most districts end up with a substantial amount of their outcomes data suppressed due to complementary suppression. As Figure 3 shows, for Glendale Unified, a relatively large district serving over 25,000 students, the C2C dashboard represses pathways data for three out of four outcomes. Even for San Francisco Unified, with almost 56,000 students, pathways data is suppressed for 2 out of four outcomes. These results get worse if one wants to examine results for a specific population. For San Francisco, three out of four college completion outcomes on the dashboard are suppressed when looking only at the subset of males, and *all* of the outcomes are suppressed when looking at outcomes for students who experienced homelessness during their K-12 education.

As a result, most districts (and local advocacy groups) are limited in their ability to use this data to identify challenges, guide strategy development, and monitor outcomes. Simply too many of the outcomes are suppressed in the dashboards. In some ways, the current data could be a step back from data districts and community colleges formerly had access to through CalPass Plus – which allowed for more detailed reporting on the progress of students, at least into community colleges, although C2C's scope is broader than that of CalPass Plus (all 2- and 4- year colleges – not just those that signed up to share data).

Happily, there are a few potential solutions for this challenge. First, C2C could design dashboards to allow users to pool data across years. So, instead of looking at pathways and outcomes for one year's class of students, users could combine two, three, or four years of students, creating cells of sufficient size to prevent suppression.

A second possibility would be to drop the cells that cause data suppression earlier in the pathway. In Figure 3, the group that is causing the data suppression on the various leaves are those who start out in a 4-year college and then transfer to a 2-year college. This is a small group overall and, when split across three different outcomes (did not complete, completed a certificate, and completed an associate degree), creates cells below the cutoff. If instead a single cell of "transferred to 2-year college" was created without attempting to distribute this group to individual outcomes, almost all the information in the leaves of the graph would be preserved. (In fact, the data query tool dataset does this, but not the dashboard).

Alternatively, the dashboard could introduce a small amount of "noise" into the various nodes of the graph – randomly adding or subtracting a small number of students at each node. For even relatively small groupings, this would make very little difference in their overall outcomes but would prevent analytic recovery of exact numbers in cells with less than 11 members. Given the use cases for this data, small changes in outcomes should not overwhelm the larger learnings that districts and schools can glean.

Finally, C2C governing board could consider creating a restricted access version of the dashboard – giving approved district administrators who have signed an appropriate data-confidentiality agreement access to the data with lower cell-size repression limits.

A second potential access challenge relates to researchers. As noted above, the research application process is not yet finalized, but one notable part of the draft process is that all research data access requests will flow up to the providing agencies for approval – which the agencies might deny. This could lead to data access being denied (or a report being prevented from publication) if an agency perceives that the research might uncover something unflattering.

Indeed, in 2023, the California Department of Education threatened to cut off access to data from researchers testifying against it in an unrelated lawsuit (Fensterwald, 2023). While the state ultimately backed down, there could be worries about the agencies pre-emptively vetoing research they do not want to see for ideological or political reasons.

To prevent this, the draft C2C guidance states that “Data providers will act in good faith to approve requests that are allowable by law. If data providers have questions or concerns about a data request, they will work with the requestor to modify the request. If a data provider denies a request, a rationale will need to be provided, which will be posted on the website.” (C2C, 2026c). By requiring agencies to state their rationale publicly and keeping "allowable" rationales to deny within a specific set of reasons, the hope is that agencies will choose not to deny research, even if they may not want it done.

How well this will work remains to be seen, and it is possible that agencies will be adept at coming up with legally plausible, but ultimately specious, reasons for denying projects.

A related access issue that is not currently clear in the guidelines is around capacity. We do not yet know how many researchers will want access to the data. Given the size and import of California, the richness of the data set, and the number of education researchers, there could easily be hundreds of requests to access this data each year. This could overwhelm both the C2C office and the data providers’ ability to approve data access requests.

A possible solution here might be to establish a research approval committee that would set priorities for research and evaluate proposals to select the strongest. Obviously, a poorly constituted committee could replicate the issue of vetoing research for ideological or political reasons. Such a committee would need to be composed of a carefully chosen mix of researchers (evaluating proposed study strength) and policy leaders (evaluating which are the most critical questions to answer) to ensure that important, but uncomfortable, questions are explored.

Potential Challenges: Data Coverage

A second set of challenges that could limit impact concerns data coverage – that is some critical data is either not available or the data that is available is in a form that limits its utility.

In the next section, we will highlight critical gaps in the data the state collects. Obviously, if no state agency is collecting certain categories of data, that data cannot be part of the C2C data system.

There are, however, some data that state agencies collect that cannot be consistently linked to other student data. Information on wages is probably the most important member of this category. In the student pathways dashboards, agencies can look at the average wages of students who graduated or get a certificate from either a four-year or two-year public university; however, C2C cannot currently look at wage outcomes for those who did not attend either of these. This group includes both those who go to private or out-of-state universities and those who do not pursue higher education at all. Understanding long-term outcomes for both these groups is important, but it is particularly important for those who do not attend college. California has made substantial investments in improving pathways to career readiness in high school (California Governor's Council for Career Education, 2025) and will need to monitor whether students who go directly into the workforce after completing a career pathway in high school are economically successful.

The challenge arises from the fact that, unlike colleges, school districts in California are prohibited from collecting data such as the social security numbers of their students. As a result, the Employment Development Department (the provider of wage data) is unable to match students who did not go to college to its records, and this critical data is not available.

C2C is aware of this challenge and is actively looking for solutions to link employment data to those who did not attend college. One promising potential solution explored by other states is to use driver's license/state id data to allow for more complete matching (Data Quality Campaign, 2025).

Even in cases where data exists and can be linked, it is sometimes not being imported in its most useful form for research or the goals of C2C. For instance, state assessment scores will be reported by level (i.e., did not meet standard, met standard, exceeded standard) and not by scaled scores or "distance from standard" as is used in the California school dashboards. If the CDE provided scaled scores instead, this would give substantially more information that is easier for analysts to work with and more likely to allow researchers to capture program impacts. Capturing only categorical levels will constrain the usefulness of these data points.

Finally, for many projects researchers will need to bring in outside data with information on things such as program participation, dosage, or demographic characteristics not captured in the standard C2C data fields. The rules for what data can be brought in for specific projects are not yet clear and, of course, C2C will need to transform identifiers in the data to those used in the C2C data enclave.

Gaps in the data

While C2C will link data across different systems and sectors in California, it cannot link data that does not exist. There are several critical areas where the state either does not collect data or collects insufficient data to evaluate outcomes, to determine whether different groups are getting appropriate levels of services, or to guide managerial improvement.

Early Childhood Education (ECE) data

ECE data is well known for being complex and hard to capture (Fullerton, 2021). Multiple departments typically administer ECE programs, and, even within departments, there can be multiple data systems. California is no exception. The California Department of Education administers the California State Preschool Program (CSPP) and Transitional Kindergarten (TK) while the California Department of Social Services administers a multitude of other childcare programs. Even within the programs CDE administers, data collection is fragmented. Districts gather TK data through their K-12 (or TK-12) SIS system, which is ultimately reported into CALPADS. CSPP data is reported into separate systems, and districts typically have completely different SIS systems for their PK programs than what is used for TK-12.

Moreover, the data that is collected (particularly outside of districts) are often snapshots of data that do not have sufficient information to connect families and students across time or to K-12 systems and later educational outcomes. Individual children in multiple programs are counted multiple times, distorting insights into supply and demand of programs. Finally, outside of the Desired Results Developmental Profile (DRDP) which is collected only for a subset of children receiving services, California does not collect any standardized academic outcomes data for children before grade 3 (Stipek and Meloy, 2026).

Early childhood workforce data is similarly fragmented across state departments and, outside of district administered programs, thin, and often voluntary or incomplete (Stipek and Meloy, 2026).

As the technical report in this series by Stipek and Meloy on early education data highlights, this fragmented system undermines the state's ability to learn which ECE programs are most effective at preparing children for school, tracking the flow of ECE workers across systems (TK generally offers substantially higher pay than other PK programs), or evaluating whether teachers that have different credentials have differential impacts of children. While the state has begun work on an Early Childhood Integrated Data System (ECIDS), it is still in early stages and collecting only basic information.

Expanded Learning data

In response to declining student engagement and test scores due to COVID, in 2021 the California Legislature passed Assembly Bill 130, which established the Expanded Learning Opportunities Program (ELO-P) to help school districts provide comprehensive support to "unduplicated" students (i.e., students who fall into one or more of the following groups: English learners, students from low-income families, unhoused, and foster youth) in Grades K–6 beyond the standard school day (A.B. 130, 2021). A.B. 130 allocated \$1.75 billion to ELO-P in 2021–22, and funding was increased to \$4 billion a year in 2022–23 (A.B. 181, 2022), 2023–24 (S.B. 114, 2023), and 2024–25 (S.B. 153). \$4.6 billion was allocated in 2025–26. This new funding is to be used to support "before school, after school, summer, or intersession learning programs that focus on developing the academic, social, emotional, and physical needs and interests of pupils through hands-on, engaging learning experiences" (California Department of Education, 2025). These programs were to "complement, but . . . not replicate, learning activities in the regular school day and school year" (California Department of Education, 2025).

This is a major investment, with districts with large numbers of unduplicated students receiving over \$2,700 per eligible student (Legislative Analyst's Office, 2026). Given that many eligible students do not take advantage of the ELO-P program, funding is likely substantially higher per participating student.

Amazingly, prior to the 2025–26 school year, districts were not required to report which students received ELO-P services. Though one hopes that districts were keeping track of this, the state has no

way of knowing whether those who are receiving services had the most need or whether the services received are related to improvements in any outcomes.

Beginning this year, the state is requiring districts to report student days in expanded learning attendance for individual students as part of CALPADS data collection (CDE, 2025, November). However, the reporting requirement is quite light – with 5 minutes attendance counting the same as a full day. No information is being captured about the specific expanded learning opportunities of which students take advantage. In other words, the state will not be able to know whether a student came to a before-school program for 10 minutes or spent six hours in a summer program. Given the size of the investment and the varying types of programs students can participate in (even within one district), a great deal of potential learning about what is working across the state is being lost.

Even if this is solved, the state would only be collecting data for expanded learning programs offered by public local education agencies. A significant number of expanded learning opportunities are offered by towns or community-based organizations that are not on school grounds nor funded by school systems. These programs can play a critical role in child development and safety. However, even identifying what programs are offered – much less tracking individual student participants – is a challenging task (National Academies of Sciences, Engineering, and Medicine, 2025, Fullerton et al., 2025).

Student Intervention Data⁴

This points to a closely related gap in California's K–12 data. Not only do we not know about expanded learning programs, but we also do not know much about other major interventions schools are trying. In the aftermath of COVID, California has spent over \$7.2 billion on a Learning Recovery Block Grant. While the point of this grant was to allow districts to invest the funds as they saw fit, little was done to systematically allow for learnings to be captured.

For instance, if no one beyond the school knows which students received a particular intervention (e.g., high-dosage tutoring, extended math classes, etc.), it can be guaranteed that systems will not be able

⁴ This section is largely taken from Fullerton, 2021.

to determine which interventions worked, their cost relative to student growth, and which are most cost-effective.

Individual school leaders, of course, do (or should) know which students are receiving interventions. The problem is a data gap among schools, districts, and the state. Like most data gaps, this exists for a reason.

Many of the interventions students receive are school based, and almost all are district based. In other words, schools and districts will vary in whether they use, say, ST Math, DreamBox Learning, or another program for supplemental online math instruction. Likewise, districts that provide tutoring to students may use tutors from a national organization such as Saga Education, or they may rely on much smaller local mom-and-pop shops to provide this service. As a result, if one captured simply the names of the service providers, one could end up with a long list of providers with only a few locales using any given provider.

In addition, analyzing the effectiveness of a given intervention would require substantial information on what services are being provided. In other words, is a student being tutored in math or English? Is the program during the school day or after school? How many minutes a week does the student receive tutoring? Does the student even attend the sessions? Without answering these questions, the long list of providers will not provide much insight into what students received. Districts, if they have these data at all, do not have them standardized across schools.

So, for system-wide learning, two problems need to be solved. First, can the long list of individual providers be shortened into a tractable number of similar interventions? Second, can data collected about interventions be standardized such that characteristics of interventions aimed at similar problems can be identified? In short, there should be a typology that allows schools to select and track program and intervention participation for students. For ease of tracking and reporting, this typology should be easy to integrate into student information systems as much as possible.

Thus, a student receiving high-dosage tutoring might get tagged with “high-dosage tutoring, [Provider Name], math, two hours per week, in-school, in-person tutors.” This would allow policymakers to get a bird’s-eye view of what strategies are being used and allow researchers to begin to identify which

strategies are effective for which kids. The key is standardizing what data get collected across schools and districts.

Such standardization is unlikely to arise from bottom-up processes, but California could be a leader in creating simple, light touch reporting standards that could be integrated into SIS systems.

Social-Emotional Wellness and Learning Data

Social-emotional learning metrics such as self-management and self-efficacy have been shown to relate to academic and other outcomes (Robbins et al., 2004; Claro and Loeb, 2019), and social-emotional well-being metrics have been highlighted as important for tracking student success and opportunity on the pathway from education to career (Gonzalez et al., 2022). Particularly since the COVID emergency, many districts have been interested in building up the social-emotional skills and resilience of their students.

However, the measurement and capture of this data is quite inconsistent. The CORE districts developed a set of survey items designed to capture social-emotional wellness and belonging. Panorama Education offers its own set of student surveys, and the California Healthy Kids Survey also attempts to capture various SEL constructs. In a recent study, the author interviewed leadership at four districts about what data they collect on SEL and found that no two districts interviewed used the same instruments or constructs to measure non-PK SEL outcomes. In addition, because the surveys have a cost, two districts had recently changed or were about to change their student survey providers. As a result, even while districts may be able to track progress internally (so long as they do not change data collection measures over time), they will not be able to compare outcomes across systems to understand relative performance across the network.

Here is a place where the state could play a role, both to subsidize the measurement of these constructs and to encourage LEAs to use the same instruments. The California Healthy Kids survey might be a place to start; however, this survey really focuses on many more topics (e.g., drug and alcohol use), and it is not clear the constructs are as robust as those on the CORE and Panorama surveys.

Conclusion and Recommendations

Accurate data linked across silos that is provided in a timely manner is essential to educational improvement. The good news is that California has developed a robust education data infrastructure over the past fifteen years with CALPADS, the California College Guidance Initiative, and other reporting tools. California has made substantial progress since the work in *Getting Down to Facts II* (2018).

The advent of the California Cradle-to-Career Data System, connecting disparate information from preschool to K-12 to college to workforce participation, is a major opportunity. Policy makers and researchers will now have much more ability to understand better the longer-term impacts of educational interventions and how different parts of the PK-workforce system impact one another.

The thoughtful and methodical process of constructing the C2C system and the care taken in ensuring privacy concerns are met in the C2C system is to be commended. C2C is also quickly evolving, with more data expected to come online soon (e.g., National Student Clearinghouse data).

However, challenges remain for California.

First, uneven analytical capacity within LEAs and uneven data and research supports to LEAs means that, due to the vagaries of district size and geographic region, some districts will be able to take greater advantage of the data available to better support student outcomes than others.

Second, while C2C represents a clear step forward for the state, many districts (especially smaller districts) may not be able to take advantage of the insights the data could offer due to how the data suppression rules play out in the current dashboards.

Third, there are significant gaps in the data the state collects, meaning that the effectiveness of billions of dollars of state investments is going substantially unmonitored.

Below are some specific recommendations that could help to solve these issues.

For C2C:

- The publicly available data in the *current* C2C dashboard has limited usefulness to small- and medium-sized individual districts, with many data points being masked due to complementary

cell-size concerns. As noted above, there are various options to reduce the number of cells masked without compromising privacy through, for instance, stacking years and/or providing districts willing to enter into data privacy agreements with restricted access views that have lower cell-size requirements.

- Some of the data points provided to C2C by participating agencies discard important underlying information. For instance, it would be helpful for the C2C data to capture scaled scores on state assessments rather than proficiency levels. Proficiency levels can be reconstructed from scaled scores, but not the other way round. Scaled scores provide substantially more information about performance and can be essential for researchers looking at program impacts.
- While C2C has made important progress on linking wage data to college attendees, the C2C and the state should continue to seek ways to link wage data to individuals who do not attend college while protecting privacy. Understanding these individuals' workforce journeys is critical to evaluating alternative pathways into the workforce and the costs of not attending college.
- The processes for gaining access to the data enclave are still under development. Two things are essential for this to create a successful policy research asset for California. 1) A process for getting external data fields into the enclave. 2) A research approval process that is transparent and does not allow departments veto power for reasons beyond the legal use of data and protection of privacy.
- C2C should continue with its focus on improving the operational data available to guide and support students through their post-secondary transition. Connecting districts to the eTranscript system and in general reducing frictions for students transferring between systems is critical.

As C2C is a rapidly evolving system, some of these may resolve naturally over the next few years as progress continues; other changes would need to go through the C2C's Governing Board and public decision-making processes.

Outside of C2C, there are several areas where the state's data infrastructure could be improved.

- Early childhood data is fragmented across departments, incomplete, and integrates poorly with CALPADS and other TK-12 systems. The state should continue to work towards an integrated

system that pulls in data from across the different systems – including offerings not controlled by the CDE.

- To this point, the state has collected almost no data on who receives expanded learning. This is surprising given the large per student amounts being funded. While the state is beginning to collect attendance data, it still will not have information about daily dosage or types of programs students were exposed to. If we want to know what works in this field and whether the dollars spent are appropriately targeted and effective, more data will need to be collected. The state knows even less about community-based programs – an important source of services in this sector.
- In general, the state should encourage “light” data collection (enrollment and dosage data) for major interventions that would allow an understanding of whether they are appropriately targeted and whether certain types of interventions are effective. The Local Educational Agency Program File component of CALPADS (developed for reporting ELO-P attendance) is designed to be flexible and used across multiple programs. The amount of data captured will need to be extended, however, as the current file only has a field for program code and days attended.
- Given their importance, California should consider incentives for school districts to adopt standard Social-Emotional Wellness measures. Districts currently use a wide array of measures, some well-supported, others not. Even when districts use well-supported instruments, if instruments differ across LEAs, comparison will not be possible. California has successfully required the use of the DRDP in state preschools. Alternatively, Illinois provides a related example in which it balances state mandates with local control. There the state requires districts to administer school climate surveys. Districts can use one of several alternatives, and the state makes one alternative (5Essentials) available to districts without charge.

While the data needs of LEAs, researchers, and families will never be fully met, the above suggestions are intended to help California continue to improve its data ecosystem to provide more relevant and actionable data to policymakers, the public, and our students.

References

- Bates, Mary Ann. (2026). "Update from the Office of Cradle-to-Career Data Executive Director." Retrieved from <https://c2c.ca.gov/meetings/governing-board-meeting-14/>.
- Burland, Elizabeth, Susan Dynarski, Katherine Michelmore, Stephanie Owen, and Shwetha Raghuraman. (2023). "The Power of Certainty: Experimental Evidence on the Effective Design of Free Tuition Programs." *American Economic Review: Insights* 5 (3): 293–310.
- California Cradle-to-Career Data System Act. (2021). AB 132, Statutes of 2021, Chapter 144, Section 8.
- California Cradle-to-Career Data System. (2024). *eTranscript California & Career Passport Task Force Recommendations*. Retrieved from <https://c2c.ca.gov/wp-content/uploads/2024/07/Final-eTranscript-California-Career-Passport-Task-Force-Recommendations.pdf>
- California Cradle-to-Career Data System. (2024, May 13). *Data Elements by Provider*. Retrieved from <https://c2c.ca.gov/resources/data-elements-by-provider/>.
- California Cradle-to-Career Data System. (2025, September) *Report to the Legislature FY 2024-25*. Retrieved from <https://c2c.ca.gov/wp-content/uploads/2025/08/C2C-2024-25-Report-to-Legislature.pdf>
- California Cradle-to-Career Data System. (2026a, February 9). *Roadmap & Progress*. Retrieved from <https://c2c.ca.gov/roadmap-progress/>.
- California Cradle-to-Career Data System. (2026b, February 17). *FAQ: Student Data Pathways Story*. Retrieved from <https://c2c.ca.gov/faq-student-pathways-data-story/>.
- California Cradle-to-Career Data System. (2026c). *Request Access to Data*. <https://c2c.ca.gov/request-access-to-data/>
- California Department of Education. (2025, October 3). *California College Guidance Initiative Update*. Retrieved from <https://www.cde.ca.gov/ds/sp/cl/ccgiupdate20251003.asp>
- California Department of Education. (2025, November 12). *CALPADS Update Flash #305*. Retrieved from <https://www.cde.ca.gov/ds/sp/cl/calpadsupdfflash305.asp>.
- California Governor's Council for Career Education. (2025). *California's Master Plan for Career Education*. Retrieved from <https://www.gov.ca.gov/wp-content/uploads/2025/04/2025-CA-Master-Plan-for-Career-Education.pdf>

- Claro, S., & Loeb, S. (2019). *Self-management skills and student achievement gains: Evidence from California's CORE districts*. Policy Analysis for California Education, Stanford University. <https://eric.ed.gov/?id=ED600478>
- Data Quality Campaign. (2022, June). *Investing in California's Data Future: How California's P-20W Cradle-to-Career Data System Could Take the State from Last to First*. <https://dataqualitycampaign.org/wp-content/uploads/2022/06/CA-P-20W-case-study.pdf>.
- Data Quality Campaign.(2023). *What Now? A Vision to Transform State Data Systems to Inform People's Pathways through Education and the Workforce*. https://dataqualitycampaign.org/wp-content/uploads/2023/04/DQC-What-Now_A-Vision-to-Transform-State-Data-Systems.pdf
- Data Quality Campaign. (2025, March 11). *Driver's License Data Can Help States Better Understand Education and Workforce Pathways*. https://dataqualitycampaign.org/wp-content/uploads/2025/03/DQC_Drivers-license-data.pdf
- Dee, T. S., & Huffaker, E. (2026). Accelerating Opportunity: The Effects of Instructionally Supported Detracking. *American Educational Research Journal*.
- Dynarski, Susan, CJ Libassi, Katherine Michelmore, and Stephanie Owen. (2021). "Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students." *American Economic Review* 111 (6): 1721–56.
- Fensterwald, John. (2023, August 18). "California Ed Department back down on punishing education researchers who testify against it." *EdSource*. <https://edsources.org/2023/california-backs-down-on-punishing-education-researchers-who-testify-against-it/696031>
- Fullerton, Jon. (2021, October). *Bridging the Gaps in Education Data*. American Enterprise Institute. <https://www.aei.org/wp-content/uploads/2021/11/Bridging-the-gaps-in-education-data.pdf?x97961>
- Fullerton, J., Burch, P., Garland, M., Makori, A., Rapaport, A., Saavedra, A., Sarmiento, K., Silver, D., and Torres, M. (2025). *Measuring Supply of and Demand for Expanded Learning Programs in Lynwood*. Los Angeles, CA: USC EdPolicy Hub. <https://rossier.usc.edu/documents/expanded-learning-details>
- Gonzalez, Christina, Agnew, Chris, and Loeb, Susanna. (2026). "The Learning Experiences that Matter and AI's Role." *Getting Down to Facts III*. Stanford, CA: Stanford University.

- Gonzalez, Naihobe, Elizabeth Alberty, Stacey Brockman, Tutrang Nguyen, Matthew Johnson, Sheldon Bond, Krista O’Connell, Adrianna Corriveau, Megan Shoji, Megan Streeter, Jennifer Engle, Chelsea Goodly, Adrian N. Neely, Mary Aleta White, Mindelyn Anderson, Channing Matthews, Leana Mason, and Sheryl Felecia Mean. (2022). “Education-to-Workforce indicator framework: Using data to promote equity and economic security for all.” Seattle: Mathematica. Available at https://educationtoworkforce.org/sites/default/files/2023-04/E-W-Indicator-Framework_Final.pdf
- Haderlein, Shira and Polikoff, Morgan. (2026). "Public Accountability in California: Evaluating the SARCs and the California School Dashboard." *Getting Down to Facts III*. Stanford, CA: Stanford University.
- Hurt, Alexandria, Reed, Sherrie, and Duyan, Casey. (2026). "K-16 Partnerships: UC Davis Aligns Education and Workforce." *Policy Analysis for California Education*.
https://edpolicyinca.org/sites/default/files/2026-03/pb_hurt-mar2026.pdf
- Johnson, Hans and Mejia, Marisol. (2025, September 4). New Data on College Completion Shows Room for Improvement. *Public Policy Institute of California*.
<https://www.ppic.org/blog/new-data-on-college-completion-shows-room-for-improvement/>
- Krupnick, Matt. (2024, Dec. 9). "Washington to Guarantee College Tuition for Low-Income Families." *The Seattle Times*.
<https://www.seattletimes.com/education-lab/state-to-guarantee-college-aid-for-low-income-families/#:~:text=Now%20some%20states%2C%20including%20Washington,%20>
- Legislative Analyst's Office. (2026). "The 2026-27 Budget: K-12 Proposals."
https://lao.ca.gov/reports/2026/5131/2026-27_K-12_Proposals_021926.pdf
- Miller, C. and Phillips, M. (2025). “Long-Term Consequences of Early Access to Educational Opportunity.” *American Educational Research Journal*. Volume 62, Issue 3, Pages 651-686.
- Miller, C., Phillips, M., and Ahearn, C. (2024). “Leaks in the College Access Pipeline: Examining Summer Melt in a Large Urban School District.” *AERA Open*, vol. 10. Sage Publications.
<https://doi.org/10.1177/23328584241278314>
- N.A. (2020). "Cradle-to-Career Data System Fact Sheet." Retrieved from
<https://cadatasystem.wested.org/wp-content/uploads/2023/03/C2C-Fact-Sheet-FINAL.pdf>
- National Academies of Sciences, Engineering, and Medicine. (2025). *The Future of Youth Development: Building Systems and Strengthening Programs*. Washington, DC: The National Academies Press.

- The National Network of Research-Practice Partnerships. (2025). *California RPP Subnetwork: 2025 Yearbook*. Retrieved from <https://rice.app.box.com/s/5zrl7xd3uwoovf1vd6thp24m3ipuphij>.
- Petek, Gabriel. (2026). *The 2026-27 Budget: K-12 Proposals*. The Legislative Analyst's Office. https://lao.ca.gov/reports/2026/5131/2026-27_K-12_Proposals_021926.pdf
- Phillips, M., Reber, S., and Rothstein, J. (2018). "Making California Data More Useful for Educational Improvement." *Getting Down to Facts II*. Stanford, CA: Stanford University and Policy Analysis for California Education. Retrieved from https://gettingdowntofacts.com/sites/default/files/GDTFII_Report_Phillips-Reber-Rothstein.pdf
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 120(2), 261–288. <https://doi.org/10.1037/0033-2909.130.2.261>
- Stipek, Deborah and Meloy, Beth. (2026). "Early Childhood Education – Data." *Getting Down to Facts III*. Stanford, CA: Stanford University.
- Trinidad, Jose Eos. (2026). "Re-Envisioning California County Offices of Education: Descriptive, Historical, and Comparative Perspectives." *Getting Down to Facts III*. Stanford, CA: Stanford University.
- Walters, Dan. (2022, January 2022). "Will California's New School Data System Really Work?" *CalMatters*. <https://calmatters.org/commentary/2022/01/will-californias-new-school-data-system-really-work/>

Figure 1. C2C Domain Coverage and Agencies Providing Data

C2C Data Domains

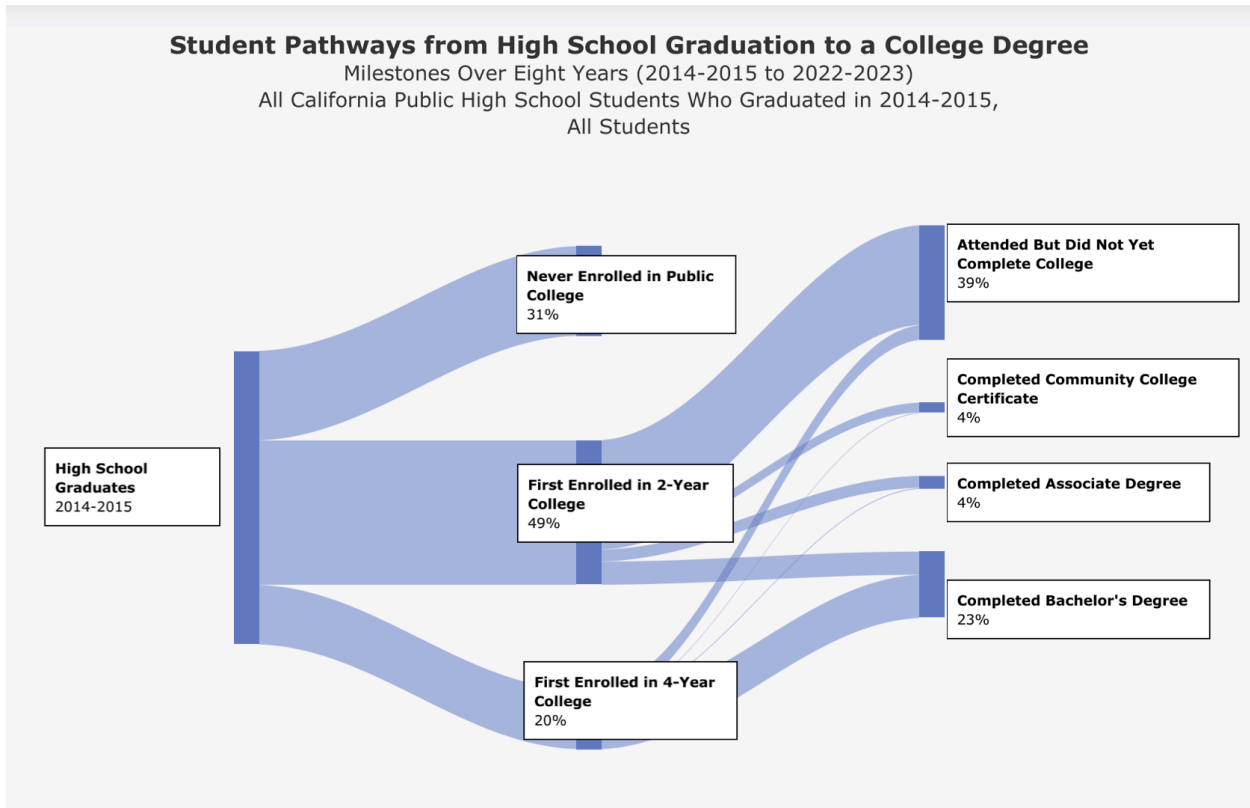
- Early Education Experiences
- K-12 Experiences
- Postsecondary Experiences
- Workforce Participant Characteristics
- Workforce Training Experiences
- Postsecondary Financial Aid Experiences
- Social Services Experiences
- Employment Variables
- Student Characteristics
- Teacher Variables
- K-12 Institution
- Postsecondary Institutions

C2C Data Providers

- Bureau for Private Postsecondary Education
- California Community Colleges
- California Department of Education
- California Student Aid Commission
- California State University
- Commission on Teacher Credentialing
- California Workforce Development Board
- Department of Industrial Relations
- California Department of Health Care Services
- Department of Developmental Services
- Department of Social Services
- Employment Development Department
- Early Learning and Care
- Employment Training Panel
- Independent Institutions of Higher Education
- Scholar Share Investment Board
- University of California

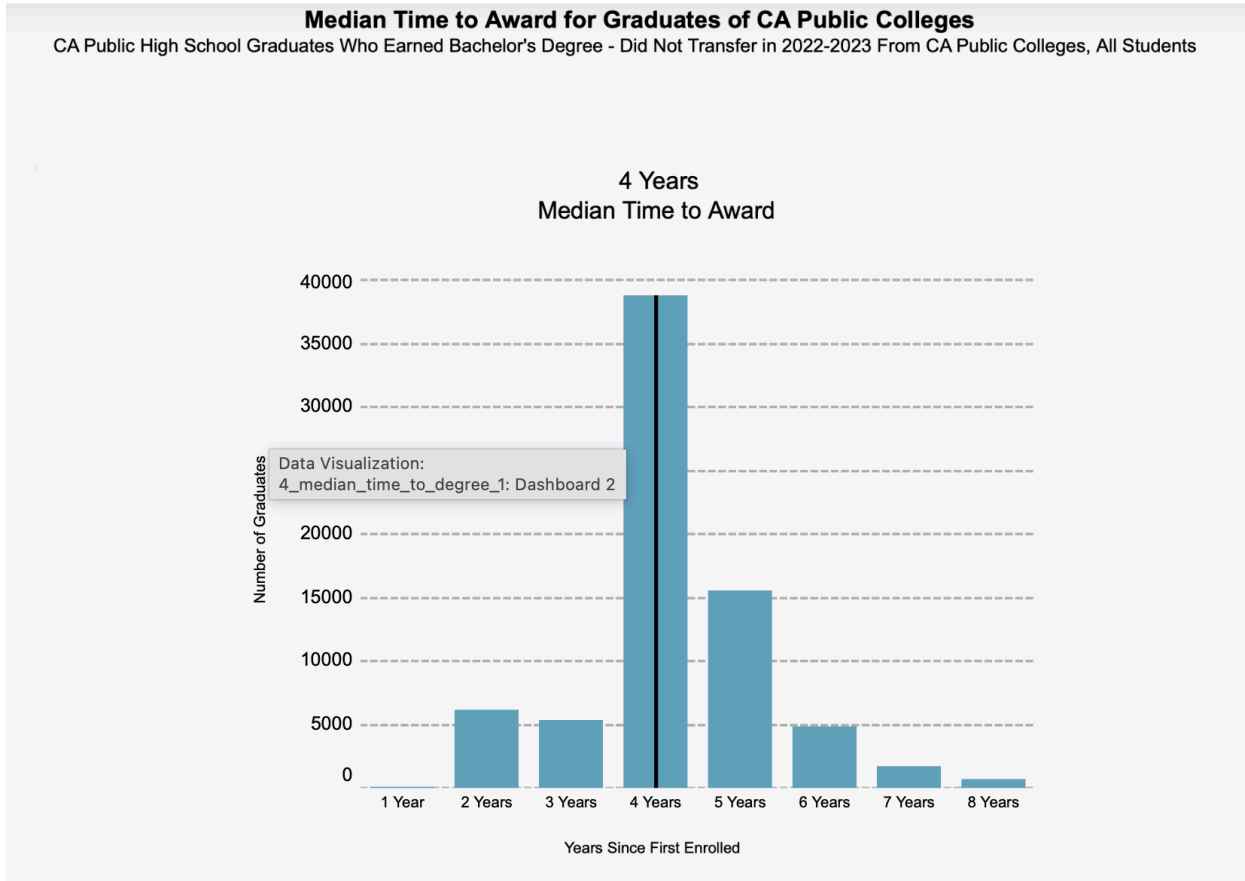
Source: California Cradle-to-Career Data System, 2024

Figure 2. C2C Student Pathways Data Story



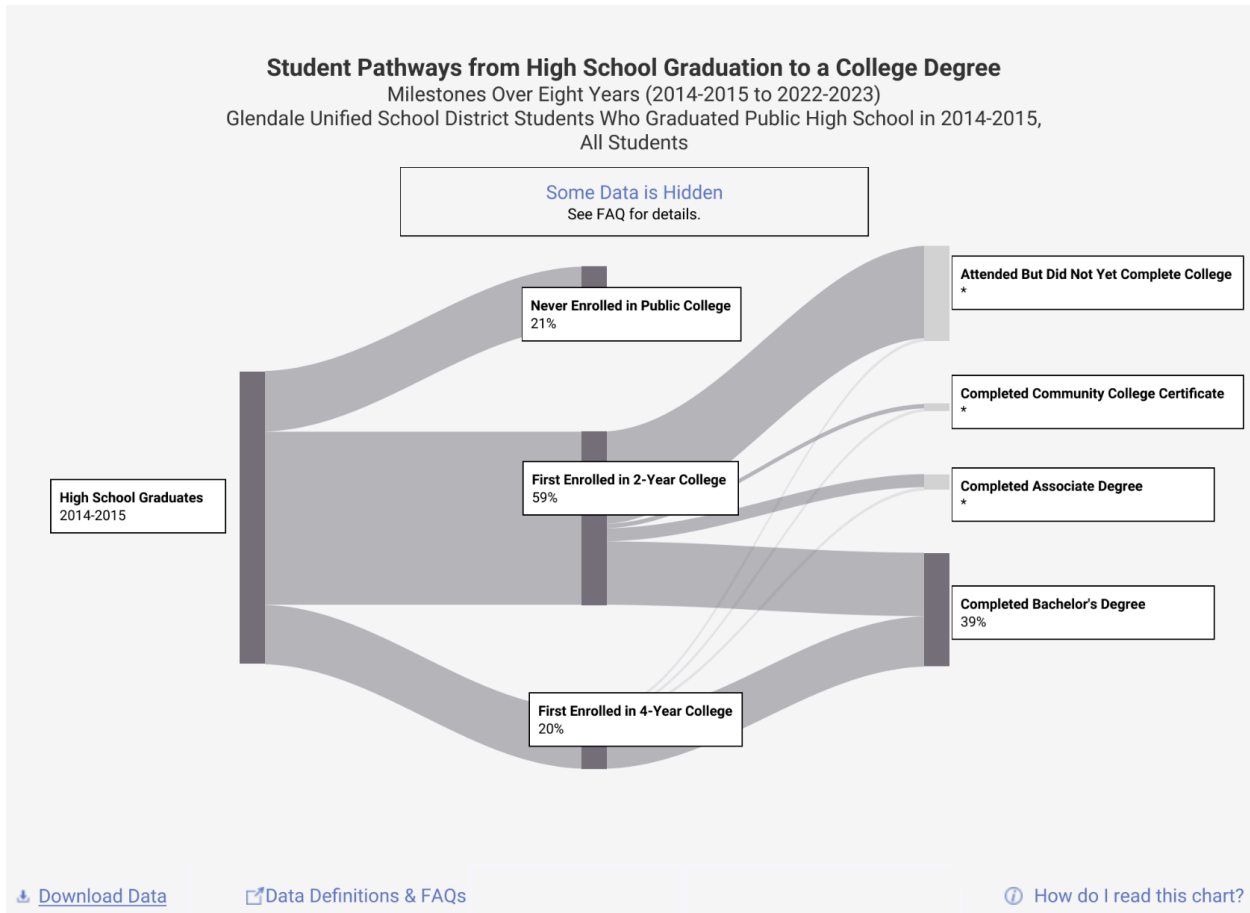
Source: California Cradle-to-Career Data System, 2025

Figure 3. C2C Student Pathways Data Story



Source: California Cradle-to-Career Data System, 2025

Figure 4. C2C Student Pathways Data Story



Source: California Cradle-to-Career Data System, 2025