



Cradle-to-Career Data System

Priorities Policy Paper

September 2026

C2C Jobs and Opportunity Hub

THE OPPORTUNITY

For the first time in California, C2C can connect data on employment and earnings to many different pathways including TK-12 education, higher education (public, non-profit, and for-profit), short-term programs, job training, and apprenticeships. This is possible because C2C has linked and brought together (and deidentified) person-level data from 16 state-level data providers.

Through a new C2C Jobs and Opportunity Hub, the state could create an efficient and scalable capability to create nimble, customized data reports on jobs and opportunity. Specifically, for people in Sacramento and across the state to make evidence-informed, data-driven decisions, they need not only the state-wide picture on jobs and opportunity but also regional analyses and customized data reports related to specific programs or initiatives. These customized data reports can be built only by C2C, which has built out its statutorily required function of connecting data on the state's comprehensive education and training paths to jobs and opportunity. The capacity created through the C2C Jobs and Opportunity Hub would ensure that relevant, timely data reports are responsive to current trends and policy changes in California.

WHAT WOULD BE CREATED?

The C2C Jobs and Opportunity Hub would release data reports and responsive custom data cuts that would address questions like these:

- How are Californians doing in the labor market, and how does this vary based on the degrees they received or the region where they live and work?
- How are these outcomes affected by the different education or job training paths they took?
- How much do Californians earn after getting specific degrees, including breaking this down by major and institution?
- What do outcomes look like for people who did not attend college? How do those paths vary over time?
- How does earnings growth and opportunity vary based on whether people stay with the same employer for a long time or work for many different companies over time?

This would meaningfully change the data and information Californians can use to inform their decisions about which education and career paths to pursue.

HOW DOES THIS IMPROVE GOVERNMENT EFFICIENCY AND EFFECTIVENESS?

Through C2C, the state can avoid reinventing the wheel of data sharing agreements and one-off data linkages. Typically, state entities share data with each other for specific, narrowly defined purposes in response to reporting requirements or legislation. This can lead to many layers of bi-lateral data sharing agreements that are burdensome to set up and where the resulting information can be used only for limited purposes.

C2C's charge in the Cradle-to-Career Data System Act envisioned C2C as a "shared utility" for the state, where C2C connects data comprehensively, so that it can be used in a flexible, nimble way for many purposes. Through C2C, many state departments and agencies share data with

C2C via one unifying legal agreement, which enables stable, long-term data products across previously siloed data sources. C2C's strict privacy protections also avoid proliferating the most sensitive individual level data.

HOW WOULD THIS BE BUILT IN PRACTICE?

C2C's work over the past four years has already overcome the most challenging obstacles to building this capability: C2C's data partners have all signed the same legal agreements, and C2C built the data system infrastructure to securely ingest, link, and deidentify data in a secure, privacy-first environment. C2C can connect the State of California's official records on jobs and earnings to the comprehensive longitudinal TK-12, higher education, and workforce training data C2C has already linked. Modest additional investments in C2C and resources for compute costs would enable the build of the C2C Jobs and Opportunity Hub.

C2C Data Insights and Research Gateway

THE OPPORTUNITY

C2C can serve as a flashlight to find wins for the people of California. C2C has already connected data from TK-12 and all of higher education, financial aid data, comprehensive social services take up data,

workforce training and apprenticeship data, teacher training data, and job and earnings outcomes. This has created a rich treasure chest of potential data insights that need to be unlocked so that California can put that information to use.

WHAT WOULD BE CREATED?

The C2C Data Insights and Research Gateway would have six connected components:

- **Open Data Files for Public Analysis**

C2C will take the data C2C has linked together across silos and make those available for the public. C2C has already piloted this approach by making the 9,000 cuts of data that power the C2C Student Pathways Dashboard available on the C2C website as aggregated, privacy protecting data flat files. These have already been used by PPIC and others to analyze policy questions. C2C expects that future files would be widely used by researchers at universities, policy research organizations, state entities such as the Legislative Analysts Office, and advocacy organizations.

- **Responsive Data Reports Based on Requests**

C2C can be responsive to requests from Californians (via C2C's governance structure) to create public C2C data reports based on the data needed to address timely, current issues. C2C is piloting this approach in the current fiscal year by creating College and Career Data Files for more than 1,000 Local Education Agencies across the state. These were designed based on a data report proposal that advanced through C2C's Advisory Board process.

- **Impact Evaluation at Scale Via the C2C Secure Data Enclave**

While public dashboards provide descriptive

information, often the most pressing questions are causal: did a program or policy have its intended impact? Why did (or did it not) work? Who benefited the most?

Often, the key reason impact evaluations are not done is because the entity that knows who received a particular program or who was affected by the policy cannot access the long-term outcomes data to measure the impact. With its overarching legal agreement and partnerships with its data providers, C2C can solve that problem at scale, while preserving privacy enabling analysis within C2C's secure environment instead of sharing data externally with researchers.

- **Secure Environments for State Departments to Do Their Own Data Analysis**

C2C can create custom gateways within the C2C Secure Data Enclave where departments and agencies can conduct their own data analysis for their own analysis and reporting.

- **Linking Data to Unlock Resources for Californians**

C2C has already connected data to unlock tens of millions of dollars in unclaimed CalKIDS scholarship accounts, by connecting data from CalKIDS to data from the California Community Colleges. Similar data linkages can reduce the administrative burden Californians face in sending official transcripts or submitting information to prove eligibility for specific services.

- **C2C Research Library Transparently Sharing Results**

In its new 5-year strategic plan, the Board committed to Open Science best practices. This means that C2C will build a public research library so that anyone can learn from the studies that are done using the underlying C2C data. The general principle is that C2C will support openness and transparency around results from studies, regardless of whether the results were positive, negative, or null.

HOW WOULD THIS BE BUILT IN PRACTICE?

C2C has the linked data, technical infrastructure, privacy policies, and legal

agreements in place to create all six components of the C2C Data Insights and Research Gateway. The only limiting factor is C2C's bandwidth of data analysis staff and marginal increases in technology and compute costs.

Critically, C2C would continue and expand its partnerships with California's universities to leverage the bandwidth of researchers to conduct detailed data analysis and research. Such collaborations would preserve C2C's identity as a neutral source of trusted information and it would also avoid an inefficient use of state resources to try to replicate already existing research capabilities at universities.

C2C Data Privacy and Security Accelerator

THE OPPORTUNITY

C2C is positioned to pioneer more advanced privacy-protecting technologies so they can be used more broadly across state government.

As the state's one-stop shop for cradle-to-career data, C2C has a statutory charge to ensure that individual privacy is protected. Via the initial build of the data system, C2C has already collaboratively advanced the use of privacy protecting approaches. This includes C2C implementing high data security and privacy standards at scale via the data

ingestion process for all 16 of its current state-level data providers. (This includes strong procedures for encrypting sensitive information in transit and at rest.) C2C has also implemented strict privacy-protecting protocols for all of the data products it releases. This means C2C's data products display data only for sufficiently large groups of people. This prevents possible reverse engineering of personal information in C2C's data products.

The outcome of C2C's strict data privacy standards has enabled an unprecedented level of information to be connected via C2C.

WHAT WOULD BE CREATED?

C2C can serve as a state-wide data privacy and security accelerator that enables more entities across the state to implement privacy enhancing approaches.

Specifically, C2C can bring the following approaches to scale in California:

- **C2C Secure Data Enclave**

Typically, when researchers receive state approval to use state administrative data for research, the state transmits the data to the researchers. Because the data physically leaves the control of the state, the data providing department must trust that the researcher will keep the data secure on their own university or institution's servers, and that they will eventually securely destroy the data. Because of the risks inherent in that process, it is understandable that state departments typically give unitary data access only to a small number of trusted researchers who have access to significant resources to keep data secure. However, this status quo significantly limits the vast number of useful data analysis, studies, and impact evaluations that could be done to better serve Californians.

C2C is currently piloting a new C2C Secure Data Enclave. This is a controlled, secure environment where approved researchers will be able to do analysis on the underlying, deidentified data in the C2C data system. With the secure enclave, researchers come to the data, instead of having data shipped out to researchers. When their study is complete, the analysis and insights can leave the enclave (following careful privacy review), but the data itself stays put—it cannot be removed from the controlled environment.

C2C can scale and build out the Secure Enclave infrastructure it is piloting to serve as the one-stop shop for research data access at scale. This will have usefulness not only for the TK-12 and higher education policy community but also for many other policy areas connected to the data C2C has linked, including economic development and job opportunities, labor and workforce development, take up and use of social services, and services provided to people with developmental disabilities. More broadly, the enclave could enable studies and impact evaluation across any policy area where long-term outcomes such as educational attainment or labor market outcomes are relevant.

- **Test and Scale Privacy-Protecting Data Release Approaches**

C2C can pioneer more advanced privacy-protecting technologies so they can be used more broadly across government. Examples include:

- Data suppression protocols (greying out data when there are too few people represented in a cell of a data table.)
- Noise infusion (injecting some “random noise” into public data reports so that information can be shared even for smaller groups of people while protecting privacy. This is particularly useful for rural areas of the state or for groups of people including tribal communities.)
- Synthetic data sets (public use data sets that protect individual privacy by being based on constructed information.)

- **Pilot and Scale Data Linkage Technology that Reduces Reliance on Sensitive Personal Information**

C2C is piloting technologies such as tokenization that can enable data linkages without proliferating or sharing sensitive personally identifiable information.

California Cradle-to-Career Data System
400 R Street, Suite 140
Sacramento, CA 95811

