Cradle-to-Career Data Governing Board Staff Report

Date Report Issued:	April 25, 2025
Attention:	Members of Cradle-to-Career Data Governing Board
Subject:	**Update on the Data and Tools Proposal Report and Feasibility Studies (Agenda Item 11)**
Staff Contact:	Ryan Estrellado, Director of Data Programs

At this meeting, the Governing Board (Board) will have the opportunity to hear a summary of the discussion about the 2024 Data and Tools Advisory Board (DTAB) proposals. This discussion took place at the March 19, 2025 DTAB meeting.

Requested Action:

There is no requested action for this item. This is an informational item only.

Background:

On February 28, 2025, the Board approved the Governance Manual Review Committee's proposed edits to the Governance Manual, which included a revision to the Advisory Board recommendation process. Rather than the Advisory Board taking a formal vote, the proposal cycle process will conclude with a report facilitated by the Board Liaison to the DTAB, accompanied by up to two DTAB members. This report will summarize all the proposals provided during that cycle and discussions that occurred at the Fall and Spring meetings.

Based on the newly revised proposal process, all proposals in the future will receive light-touch feasibility studies from the Office of Cradle-to-Career Data (Office). The goal of light-touch feasibility studies is to provide the Advisory Boards with proposed ways to operationalize the proposals.

Because the Office began the 2024 proposal process with the previous procedure and is ending the process with the revised procedure approved by the Board, this cycle will look different than in future iterations. Specifically, in

preparation for the Fall DTAB meeting, the Office performed light-touch feasibility studies on all six proposals received. At the Fall DTAB meeting, three proposals advanced, which received in-depth feasibility studies. Each proposal is listed below and are included in the attachments.

Digital Access to Learning Proposal:

The Digital Access to Learning proposal submitted by Member Borgen was advanced by DTAB in Fall 2024. The light-touch and in-depth feasibility studies, and the proposal can be found on Attachment B. This proposal asks the Office to explore what data is available on students' access to high-speed internet, digital devices, and other digital readiness supports.

Student Debt at Colleges and Universities Proposal:

The Student Debt at Colleges and Universities proposal submitted by Member Schak was advanced by DTAB in Fall 2024. The light-touch and in-depth feasibility studies, and proposal can be found on Attachment C. This proposal asks the Office to explore whether additional detail could be provided on the composition of student loans (subsidized versus unsubsidized) and parental loans.

Weaving Disaggregated Multilingual Learner Data Proposal:

The Weaving Disaggregated Multilingual Learner Data proposal submitted by Members Orlick and Owen was advanced by DTAB in Fall 2024. The light-touch and in-depth feasibility studies, and the proposal can be found on Attachment D. This proposal asks the Office to explore what additional information could be provided about students' duration as English Language Learners (ELL) and their newcomer status.

Graduate School Completion Rates by Field of Study Proposal:

The Graduate School Completion Rates by Field of Study proposal submitted by Member Schak did not advance in Fall 2024. The light-touch feasibility study and proposal can be found on Attachment E.

Including Internship and First-Destination Survey Data Proposal:

The Including Internship and First-Destination Survey Data proposal submitted by Member Phuong did not advance in Fall 2024. The light-touch feasibility study and proposal can be found on Attachment F.

Track Child Savings Accounts Data Proposal:

The Track Child Savings Accounts Data proposal submitted by Member Phuong did not advance in Fall 2024. The light-touch feasibility study and proposal can be found on Attachment G.

Board Liaison Report:

After the DTAB March 2025 meeting, Board Liaison Gonzalez-Vasquez worked with Member Borgen and Member Phuong to draft the Board Liaison Report, which can be found on Attachment A. This report includes a summary of the feasibility studies, proposals, and DTAB member discussion from the March 2025 meeting.

To facilitate Board member's deliberation and discussion, the Office has compiled some questions for members to consider:

- Which proposals are most feasible and have the highest potential impact?
- Are there any potential risks or drawbacks that need to be discussed?
- Does this proposal fit into the overall mission and vision of C2C?
- How would this proposal fit into the C2C workplan?

SECTION 1. PROPOSALS OVERVIEW

Several proposals were submitted for consideration, each seeking to expand the Cradle-to-Career (C2C) Data System:

1. Digital Access to Learning

- Proposed by Member Borgen, this focused on whether C2C could include data on students' access to internet-enabled devices in schools and home broadband access to devices (e.g., type of connectivity, usage of federal programs like eRate) to evaluate correlations to digital learning ot student success aligning to the state's Digital Equity Plan. Proponents saw this as an equity issue, noting that reliable broadband and digital tools can affect everything from K–12 performance to college readiness.

- Examples of potential data points included bandwidth speeds, the number of students receiving subsidized internet service, and device-distribution rates at the institutional level.

- While seen as valuable, the staff's deeper feasibility study highlighted that few existing administrative data streams capture this information, and new local-level data collection would be costly and burdensome.

2. Weaving Disaggregated Multilingual Learner Data

- Submitted by Members Orlick and Owen, this proposal aimed to refine how the system tracks different groups of English learners—for instance, distinguishing "Newcomers," "At-Risk English Learners" (ARLTEL), "Long-Term English Learners" (LTEL), and "Dually Identified" students.

- Members noted that grouping all English learners together hides important nuances and can limit targeted policy responses.

- Staff concluded that some of these variables could be derived from existing data—e.g., length of time a student is designated as EL—but others (such as special "newcomer" status) would require new data elements or improved statewide definitions provided by current data providers (i.e, California Department of Education).

3. Student Debt at California Colleges and Universities

- Proposed by Member Schak, this sought additional detail in C2C on federal and nonfederal (private) student loans, especially distinguishing amounts borrowed by parents versus amounts borrowed by students.

- Proponents underscored how loan type and cumulative debt can influence student outcomes and argued that state-level data would help families and

policymakers gauge affordability.

- The feasibility study showed that, for many segments, partial debt data is already being collected (e.g., certain community colleges and independent institutions), but it is inconsistent. UC and CSU do not yet upload all debt categories into C2C. Staff concluded that more consistent reporting might require additional authority or a revised data specification.

4. Other Proposals (Light-Touch Feasibility)

Three proposals did not receive a full feasibility study but were reviewed briefly:

- **Track Child Savings Accounts (CalKIDS)**: Member Phuong proposed that C2C integrate data on CalKIDS (state-funded college savings accounts) to show usage across different regions and student subgroups.

  The ScholarShare Investment Board (SIB) does have account-level data, collecting it at the individual level for C2C would require new data-sharing and matching processes. That being said, in the February 28th Governing Board meeting, SIB was added as a data-sharing partner.

- **Graduate School Completion Rates by Field of Study**: Member Schak proposed better tracking of graduate-level enrollments and degrees, including time-to-completion and field of study.

  Proponents sought to measure how students in master's, doctoral, or professional programs fare by major or Classification of Instructional Programs (CIP) code. Staff found that basic data on graduate enrollment and outcomes is already tracked in certain segment databases, suggesting the potential to create dashboards (like the existing Transfer dashboards) for graduate programs. However, cross-segment consistency and the level of detail on specific fields are uncertain.

- **Including Internship & First-Destination Survey Data**: Member Phuong advocated collecting information on whether undergraduates had internships or career-focused coursework, plus linking "first destination" survey data (e.g., whether graduates were employed or in graduate school within six months).

  Many colleges conduct first-destination surveys—tracking graduate outcomes like job placement or admission to further study—but response rates and data definitions vary widely. Similarly, data on internships or career-focused coursework is not uniformly reported to a central, statewide system.

  Staff concluded that linking these data (for instance, from the National Association of Colleges and Employers First-Destination Survey) into C2C might be feasible but would require a standardized reporting process and broad agreement from the higher education segments.

Overall, the light-touch reviews underscored the potential value of these proposals but also highlighted barriers such as data consistency, new data-sharing agreements, and the Office's capacity to handle additional data integration efforts.

SECTION 2. DISCUSSIONS KEY TAKEAWAYS

During the March 19 Data and Tools Advisory Board meeting, members engaged in active debate on the merits and complexities of each proposal:

- Digital Access
    - Some members expressed concern over data quality: K–12 addresses often change, and local surveys about connectivity may be underreported or unreliable. Others emphasized the strategic importance of bridging the digital divide and using the C2C platform to highlight digital inequities.
    - Several members pointed to potential "survey fatigue" and the risk of burdensome new data collections. Member Harlick brought up that additional data points (e.g., early education data) outside of the current data collection may also want to be added to a survey, which begs us to ask the question, "Which data is prioritized when adding to a survey?"
    - Member Borgen questioned whether we needed individual-level data versus institution-level data. This opened up the broader question of whether institutional-level data is sufficient, accurate, and/or valid enough for C2C system users to make informed decisions.
    - Despite these concerns, the group largely agreed that the concept of digital access as a factor in student success is "important" and that the Office could advocate for a consistent, statewide approach that may not be a survey.

- Multilingual Learners
    - Members spoke about the potential to glean new insights by distinguishing among diverse English learner populations (e.g., newcomers, long-term English learners). They noted how combining data about language status with educational outcomes could elevate resource decisions.
    - However, there was discussion about whether the necessary flags (such as "at-risk" or "dually identified") are routinely tracked and how feasible it would be for school systems to submit them.
    - The Board discussed using the terms that already exist and are used by the California Department of Education and California Schools Dashboard. These terms would be what families are already familiar with.

- Student Debt
    - The Board recognized debt data as a valuable lens on college affordability and student success. Legal concerns were discussed about displaying sensitive and potentially identifiable debt figures and how easily some segments (e.g., private

lenders) can share data. There are also strict regulations about how the data can be used according to the ISIR (Institutional Student Information Record) Guide.

- The conversation touched on parent loans and the possible mismatch between data privacy constraints and the goal of transparency. Some participants felt that robust feasibility and legal analysis would help clarify how to safely include these data.

- Other Proposals (Light-Touch Feasibility)

  - **Track Child Savings Accounts (CalKIDS)**: No discussion.

  - **Graduate School Completion Rates by Field of Study**:
    - The Board discussed that including this data may not accurately reflect the determinants that impact student success in Post-Baccalaureate education.

  - **Including Internship & First-Destination Survey Data**:
    - Some members acknowledged that collecting local data—e.g., internship experiences or child savings account usage—could fill crucial gaps, but it remains challenging because many existing systems either do not track these items or do so only sporadically.

  A few members stressed the importance of context surrounding student success and how those elements are currently missing from the data system. Participants stressed that the Office's role should not overextend into areas that require major new data collections that agencies are not already set up to manage.

Key Considerations

Across the March 19 discussions and feasibility reports, the Data and Tools Advisory Board recognized the value in each proposal while also acknowledging the practical hurdles of expanding data submissions. Members emphasized:

- **Prioritizing existing data**: Where relevant data are already tracked by state agencies (e.g., UC's graduate school records, K–12's EL timelines, some aspects of student debt), new additions to C2C might be more feasible.

- **Balancing new vs. existing mandates**: If a proposal demands novel local data collection (e.g., broadband speeds, home addresses for connectivity checks, or campus-by-campus internship tracking), feasibility diminishes without additional legislation or targeted funding.

- **Further Legal & Policy Analysis**: Especially for sensitive topics (like parent loan amounts) or newly minted programs (like CalKIDS), as well as the need for statewide identifiable data around device and broadband access for digital learning purposes.

Cradle-to-Career Data and Tools Advisory Board Proposal Form

Instructions:

Per the Governance Manual proposal forms submitted will address significant gaps regarding whether the data system is providing access to actionable information. Please note there should only be one proposal per form.

Name:

Jason Borgen

Proposal Title: *no more than 50 characters*

Digital Access to Learning

- By checking this box, I understand that if my recommendation is one of the top recommendations identified by the Advisory Board, I will be presenting my idea at the Fall Advisory Board meeting.

Type of Proposal[1]:

- Changes to practical tools for students *(Complete section one)*
- Adding data points not available through the P20W data set or adjusting (including the removal of) the existing P20W data points *(Complete section two)*

[1] The C2C Governance Manual notes that DTAB members can also submit recommendations related to the data request process and changes to tools such as dashboards. As the data request process and the dashboards are not yet live, those sections have been removed from the 2024 version of this form.

- Please note a proposal form(s) can recommend adding one data point or several clearly related data points to the data system.

Section One: Changes to Practical Tools

1.What is the nature of the gap regarding access to actionable information?

2. What type of tool should be developed?

3. How would a tool address the gap?

4. Who would be the likely user(s) of the tool?

5. How does the tool relate to the mission and vision of C2C?

Section Two: Adding Data Points Not Available Through the P20W Data Set or Adjusting the Existing P20W Data Points

1.Please state the research question of interest that cannot be fully addressed with the existing data elements in the P20W Data System.

How does access to digital tools and broadband Internet affect learning outcomes and succuss towards graduation and career-readiness?

2. How does this research question relate to the mission and vision of C2C?

In order to provide insights into student milestones, digital equity is essential to ensure students reach their full potential. The partnership with the state

department of technology and the Digital Equity Plan connects the goals of the state plan with data elements that can be provided by C2C through the following areas:

1. Identifying Gaps in Access

Research into digital equity can help identify which student populations lack adequate access to digital tools and broadband Internet. This is often linked to socioeconomic status, geographic location, and other factors. By understanding where these gaps exist, policymakers and educators can target resources and support to those who need them most.

2. Understanding Impact on Learning Outcomes

Access to digital tools and the internet is increasingly essential for modern education. Research can explore how limited access impacts learning outcomes, such as lower academic performance, decreased engagement, and reduced opportunities for remote learning. These insights can inform interventions that ensure all students have the tools needed for academic success.

3. Addressing the Digital Divide

The digital divide refers to the gap between those who have access to digital technology and those who do not. By researching digital equity, educators and policymakers can better understand the barriers that contribute to the digital divide, such as cost, lack of infrastructure, or limited digital literacy. This understanding is critical for developing strategies to bridge the divide, ensuring all students have equal opportunities.

4. Enhancing Digital Literacy

Digital literacy is a key component of student success in the 21st century. Research can highlight the importance of teaching digital skills and how lack of access to technology can hinder the development of these skills. This information can guide curriculum development and professional

development for educators, helping to integrate digital literacy into education at all levels.

5. Supporting Career Readiness

Digital tools and internet access are essential for preparing students for the workforce. Research can explore how digital equity affects students' readiness for careers, particularly in fields that require technological skills. By understanding these connections, educational programs can be tailored to better prepare students for the demands of the modern job market.

6. Informing Policy and Resource Allocation

Insights from research on digital equity can inform policy decisions and the allocation of resources. For example, they can justify investments in broadband infrastructure, device provision, and digital literacy programs. Policymakers can use this data to advocate for funding and initiatives that promote digital inclusion.

7. Promoting Social and Educational Equity

Digital equity research aligns with broader goals of social and educational equity. By ensuring all students have access to the same digital resources, we can work towards a more equitable education system where all students have the chance to succeed, regardless of their background.

3. Please propose additional data element(s) needed in order to successfully address the research question of interest.

In School Access:

- K12HSN school connection speeds
- Community College and UC calREN Connection speeds
- eRate adoptions rates by district
- Devices provided to students in class ratio

At Home Access:

- ACP/ECF/Internet for All adoptions
- Devices provided to student for home use
- Device access at home meeting minimum requirements for distance learning and/or access to online curricula
- The California Public Utilities Commission (CPUC)Broadband Map of connected areas/schools

4. Please verify that the proposed data element(s) do not already exist in the P20W Data System.

There is no connection to digital technologies, connectivity, nor broadband listed in the P20W data sets.

5. If the P20W Data System does not currently include the proposed data element(s), can the proposed data element(s) be derived from the existing data element(s) in the P20W Data System?

N/A

6. If the P20W Data System does not currently include the proposed data element(s), are there existing data element(s) closely related to the proposed data element(s)? If so, please list them and why they are not sufficient to answer the proposed research question.

No.

7. C2C keeps a repository of previous data elements that were either a) considered during the planning process or b) proposed through the mechanisms as outlined in the Governance Manual, but were ultimately not included in the P20W, along with related feasibility studies. Do any of the data element(s) being proposed overlap with the data elements in this repository? (New proposals can build on or duplicate prior proposals. It is helpful for proposers to share that context, including prior related feasibility studies.)

N/A

8. Are the proposed data element(s) already collected by a state-level entity? For data element(s) already collected, please answer question nine. For data element(s) not collected, please answer question ten.

Some are, i.e. CPUC

9. [For proposed data element(s) already collected] To the extent possible, please share details pertaining to the proposed data element(s). Relevant details may include but are not limited to: a) corresponding entity that collects and houses the data element(s); b) specific variable name(s) used in the originating data system; and c) timeframe available.

- K12HSN school connection speeds – bandwidth speeds, collected annually
- Community College and UC calREN Connection speeds – bandwidth speeds
- eRate adoptions rates by district – Universal Service Administrative Company – collected annually
- Devices provided to students in class ratio (Tech budgets)
- ACP/ECF/Internet for All adoptions – CETF – ongoing collection
- Devices provided to student for home use – Survey/Registration survey Data – varied collection
- Device access at home meeting minimum requirements for distance learning and/or access to online curricula -Survey
- The California Public Utilities Commission (CPUC)Broadband Map of connected areas/schools – ongoing collection

10. [For proposed data element(s) not collected] Please propose institution(s) that would be most suited for the new data collection effort.

N/A

11. Please explain the desired level(s) of grain size for each data element proposed. (i.e., individual-level, institution-level, or other aggregated levels)? Multiple grain sizes may be requested for each proposed data element.

Some individual levels some institutional level. Disaggregated by region, demographics, type (urban, suburban, rural, etc). Correlated to student/institution academics, engagement, behavior, and attendance, college entry, etc.

12. Please explain the intended use case(s) for the proposed data elements (i.e., dashboards, query builder, or the research request tool)? Multiple use cases may be requested for each proposed data element.

These elements would support the alignment of the California Digital Equity Plan around Education and workforce development in a dashboard. Using these data sources they also would fit well in query builders to support research and justify funding allocations and priority areas to support areas that show correlation between digital access and student success as we well as college/career-readiness.

Digital Data

Voting Option for Fall 2024 Proposals from the Data and Tool Advisory Board

This document provides background information to support prioritization of potential new data points for the Cradle-to-Career (C2C) Data System P20W data set. P20W data points include data uploaded directly by data providers, as well as metrics that are calculated from the information shared by data providers. P20W data points will be used to populate dashboards and a query builder tool. When the data request process becomes available, users may request other formulations of metrics that are generated from the information shared by data providers.

Staff from the Office of Cradle to Career Data (Office) worked with Advisory Board members who submitted proposals to provide greater specificity about the suggested item.

Proposal Name

Digital Access to Learning

Type of Data Point

Based on an initial review of the proposal, the Office notes that these items are data points that are not currently in the P20W data set.

Proposed Data Point Construction

In consultation with the Advisory Board member who proposed the data point, the Office operationalized the proposal to provide information on digital data by creating 11 metrics. The Office could utilize information from the California Public Utilities Commission (CPUC) California Interactive Broadband Map to generate some of these metrics or to create visualizations.

1) **K12 School Internet Access Connection Speeds** - Average internet connection speeds at a school site during the academic year:

- Download speed is 0-99 Mbps
- Download speed is 100-999 Mbps
- Download speed is 1-9 Gbps
- Download speed is more than 10 Gbps

2) Public Postsecondary Connection speeds - Average internet connection speeds at a college site during the academic year:

- Download speed is 0-99 Gbps
- Download speed is 100-399 Gbps
- Download speed is 400G or more

3) **K12 district participation in federal eRate universal service broadband support program** - Whether the district receives federal funds to support digital access:

- Receives support for telecommunications, telecommunications services and Internet access
- Receives support for internal connections, basic maintenance of internal connections, and managed internal broadband services
- Does not receive eRate support

4) **Proportion of students provided with digital devices for in-class use** - Number of students who have devices that are available in the classroom, compared to number of students in the district:

- 0-20%
- 21-40%
- 41-60%
- 61-80%
- 81-100%

5) **Type of digital devices available to a student for at in-class use -** Type of devices that are available to students in the classroom:

- Smart phone
- Tablet
- Windows, Apple, or Chromebook computer less than 5 years old
- Windows, Apple, or Chromebook computer 5 years old or more

- No digital devices available

6) **Proportion of students who received federal support for digital access -** Number of students who participated in a federal program that subsidized the costs of digital connections or equipment, compared to number of students in the district or the postsecondary institution:
- 0-20%
- 21-40%
- 41-60%
- 61-80%
- 81-100%

7) **Proportion of students provided with digital devices for at home use -** Number of students who have devices that are available at home, compared to number of students in the district:
- 0-20%
- 21-40%
- 41-60%
- 61-80%
- 81-100%

8) **Type of digital devices available to a student for at home use -** Type of devices that students have at home:
- Smart phone
- Tablet
- Windows, Apple, or Chromebook computer less than 5 years old
- Windows, Apple, or Chromebook computer 5 years old or more
- No digital devices available

9) **K12 Student Cellular Service Access** - Type of internet access available at the student's home address during the academic year:
- 5G
- 4G
- Below LTE
- No cellular service available

10) **K12 Student Home Internet Access Connection Speeds** - Average internet connection speeds available at the student's home address during the academic year:

- Download speed is at least 100Mbps
- Download speed is less than 100Mbps

11) **Staff Capacity to Support Digital Literacy:** Number of staff supporting digital literacy and technology, compared to number of students in the district or the postsecondary institution:

- Less than 1:300
- 1:300 or more

Feasibility Study: Digital Access

At the October 1, 2024 meeting, the Data and Tools Advisory Board advanced a proposal from Jason Borgen to include information on digital access in the Cradle to Career Data System (C2C) analytical data set. Specifically, the California Department of Education (CDE), California Community Colleges (CCC), California State University (CSU), and University of California (UC) were requested to provide information on:

- download and upload speeds at each local educational agency (LEA) or postsecondary institution
- whether each LEA or postsecondary institution received federal funds to support broadband access
- the proportion of students benefiting from federal broadband aid at each LEA or postsecondary institution
- the proportion of students receiving digital devices for in-class and at-home use at each LEA or postsecondary institution
- the devices students have access to for in-class and at-home use that were provided by the LEA or postsecondary institution, at each LEA or postsecondary institution
- the type of digital devices that students have at home
- the type of internet access available at each K-12 student's home address
- the average internet connection speeds available at each K-12 student's home address
- the number of staff supporting digital literacy and technology at each LEA or postsecondary institution

The full text of the proposal can be found at this link.

When conducting feasibility studies, the Office of Cradle-to-Career Data (Office) considers four factors: data availability, data reliability and data validity, cost, and compliance. Information on each aspect is included below.

Data Availability

Data Sources and Historical Range: *Where the data originates from and the span of years for which data are available*.

Currently, none of the agencies that have signed the participation agreement for the C2C analytical data system collect the proposed data points. While the Office has identified other entities that collect related information, none of the institutions identified below collect this information in a manner that can be connected to the individual-level data within the P20W analytical data set.

Education Institution Broadband and Digital Device Access

CENIC's California Research and Education Network (CalREN) provides broadband infrastructure for public educational institutions and has information that could be used to identify download and upload speeds for LEAs and postsecondary institutions.

For LEAs, information on upload and download speeds can be downloaded from the DataLink website. However, this information is self-reported on a voluntary basis. Information may be outdated or missing.

CENIC maintains information on broadband speeds for all community colleges, CSUs, and UCs, which could be provided for the analytical data set, but does not have information on independent colleges.

Federal Funds

The Federal Communications Commission (FCC) eRate program helps pay for first-time and recurring broadband costs for K-12 schools and libraries. Institutions that take advantage of this funding are connected to the CalREN system. Information on grant recipients is available on a public portal. However,

because funding may be provided for consortia, information is available at the level of the county offices of education, not the LEA level.

There are a variety of federal programs that could conceivably be used to support broadband access, but would require analysis to identify whether funds went to specific LEAs or postsecondary institutions. For example, the federal pandemic-era Broadband Equity, Access, and Deployment (BEAD) program provided one-time funds for multiple types of technology-related investments. Initial funding went to residences and business. If funds remain, they could be spent on educational institutions, but it is unlikely that there will be sufficient remaining resources. Information on these grants is tracked by the California Public Utilities Commission and could potentially be provided for the analytical data set.

The U.S. Department of Agriculture also provides funding to support broadband in rural regions, which could potentially be provided to educational institutions.

Finally, the FCC Emergency Connectivity Fund, which provided resources from 2021 - 2024, supported education institutions to purchase hotspots, devices, and laptops.

Home Digital Devices

The Household Pulse Survey is an online survey conducted by the U.S. Census that measures how emergent social and economic issues are impacting households across the country.[1] Information is available at the federal and state levels, as well as the nation's 15 largest metropolitan statistical areas. The survey includes questions on internet and digital device access, but the topics covered by the survey change in response to current events, so the specific information gathered may not be consistent from year to year. However, information is not

[1] Recent studies on digital access in California, such as by the California Budget & Policy Center and the Public Policy Institute of California, leveraged the Household Pulse Survey to document reliable access to the internet and digital devices.

available at the LEA or individual level in a way that could be combined with existing records in the P20W analytical data set.

The California Department of Technology partnered with the nonprofit California Emerging Technology Fund to produce a Statewide Digital Equity Survey in both 2021 and 2023. Information on access to broadband and digital devices was gathered using phone and online surveys from between 1,650 and 3,200 households. As is the case with the Census study, information is not available at the LEA or individual level in a way that could be combined with records in the P20W analytical data set.

K-12 Student Internet Access Type and Speeds

The California Interactive Broadband Map provides information at the level of individual addresses on whether fixed broadband is available as well as upload and download speeds for specific mobile plans. However, information on fixed broadband speeds is available for only half of the state and access to upload and download speeds may not be the same for every plan that a family could potentially purchase. This map also does not include information on whether households have a digital plan or which plans they use.

Staff Supporting Digital Literacy

Information on the number of staff who support digital literacy and technology cannot be discerned from currently available data. Job classifications do not identify these specific roles. For example, in K-12, the only categories that are documented are: administrator, pupil services, and teacher.

Future Relevance and Data Availability: *Evaluate whether the new data will remain useful as the analytical dataset evolves and if the information is likely to be collected in the future.*

California supports the Broadband for All Action Plan, which includes several initiatives focused on closing the digital divide through support for the State Digital Equity Plan, middle and last-mile network and infrastructure funding, and the former Affordable Connectivity Program. As the need for digital access and

literacy continues to grow, the proposed data points align with the direction of these initiatives.

However, other than information from CalREN on institutional broadband access, it is not clear that other proposed data points will be available.

Data Reliability and Data Validity

Institutional Variability: *Whether there are variations in administrative practices and data recording across institutions at the local level.*

Some communities have created local surveys to better understand the digital divide.[2] Because these surveys have been developed locally, they would not provide consistent information for the C2C analytical data set.

CENIC staff noted it has been challenging for LEAs to track the location of the digital devices they have purchased, which would mean that data collection on which students have access to specific devices would likely be of low quality.

Asking families to report on their upload and download speeds and on the digital devices they own would likely produce low-quality data. For example, parents and guardians would need to be instructed on how to evaluate upload and download speeds.

Agency Variability: *Data and metric definitions across various agencies follow a uniform format and standard.*

CCC noted that it might be possible to extrapolate information on institutional provision of digital devices using the new Vision Aligned Reporting (VAR) metrics. However, data collection began in fall 2024 and will not be due until fall 2025. The quality of the VAR data will need to be evaluated before determining whether they provide reliable information on the proposed data points that could be integrated into the C2C analytical data set.

Similar data are not available from CDE, CSU, or UC.

[2] See for example, a study commissioned in Long Beach.

Data Integration Across Agencies: *The Office's ability to consolidate data from multiple agencies.*

Because information is not gathered by state agencies, one possibility would be to integrate survey information in the C2C analytical data set. However, survey data are partial – the responses of some Californians are extrapolated to represent other people with similar characteristics. Unless surveys are sent to tens of thousands of Californians, the information on digital devices used at home, broadband access, and upload and download speeds would be based on very small samples of students.

For example, the Household Pulse Survey cannot be used to show differences between all parts of the state. While information is available at the state level, regional data is only available for Los Angeles, Riverside, and the Bay Area. Even where there is regional information available, because the identity of the people who responded to the survey is not shared by the U.S. Census, it would be impossible to map those results to specific children or the education institutions that respondents attended. Similarly, because the Statewide Digital Equity Survey went to such a small sample of Californians and their identities would not be provided to C2C, results could not be shown at the individual or institutional level.

Broadband and Digital Device Access
Although it is possible to map CalREN data to different parts of the state, there are significant challenges in using address information to track broadband and digital device access at the student level and thus assess the impact of the digital divide on factors like retention or graduation.

Students' address information is often unreliable, particularly for students facing food and housing insecurity, who may not have stable addresses. In addition, K-12 children may live with more than one family member. CCC, CSU, and UC highlighted that address information is collected at the point when students apply to college and may not reflect where a student subsequently moves.

In addition, address information associated with an institution may not be an effective proxy for broadband and digital device access. This is particularly an issue for community colleges. Many students take classes online and may not

benefit from higher broadband speeds or digital devices that are available on campus. Furthermore, students can take courses at community colleges in other parts of the state through California Virtual Campus, further diluting institutional address as a way to evaluate individual students' broadband access.

Finally, access to digital devices is just one of many factors that might shape outcomes, particularly given that lack of access may be an indicator of economic insecurity. Other factors, like inadequate food or housing, might also inform differences in outcomes like graduation rates. In the context of dashboards or query builders, it is not possible to convey this level of nuance.

Federal Funds

Information on whether LEA or postsecondary institutions received federal funds for technology could be identified by having an Office staff person review government websites to document which institutions won specific awards. This would have to be compiled manually by first identifying each federal funding opportunity and then comparing awardees to lists of California's LEAs and postsecondary institutions. However, it may not be possible to readily identify what proportion of students benefited from those awards.

Staff Supporting Digital Literacy

Many LEAs and postsecondary institutions prioritize digital literacy and have embedded these skills across the curriculum. It would theoretically be possible for an Office staff person to review the learning outcomes for each of the thousands of courses taught at each LEA and postsecondary institution in the state to determine which classes teach digital literacy and technology skills, but this effort requires knowledge of the related curriculum and would need to be repeated each year. In addition, information is not readily available on how many different faculty teach each course. This approach would also miss other staff who teach digital literacy. For example, at many community colleges, librarians teach courses on research skills that include a digital literacy component, which would not necessarily be flagged in learning outcomes for a course.

Cost

Startup Costs: *Costs to begin collection including both direct costs and associated staff time for the Office and its data providers.*

There would be a significant cost and burden to begin collecting these data points. If LEAs and postsecondary institutions were asked to report this information, they would have to devote time and resources to documenting receipt of federal funds and what proportion of students benefit from those resources, staffing allocated to digital literacy and technology, how many students have access to specific digital devices, the types of broadband access students have at home, and what devices they own. They would also need to amend their local data systems to allow them to track this information. Finally, they would need to devote time to uploading the information to the state level data systems maintained by CDE, CCC, CSU, or UC.

Startup processes for the data providers would also be significant. They would need to amend the data structure of their information systems, develop guidance documents regarding how to report this information, host training for LEAs and postsecondary institutions regarding new requirements, and validate the new information. Generally, focused attention is required for three years to develop, explain, collect, and ensure the quality of new data points.

The primary purpose of the P20W data system is to connect pre-existing administrative data. The Office does not currently have the staff or funding necessary to conduct surveys as a data point contribution to the P20W data system. To document the upload and download speeds and available devices at each K-12 student's home, the Office would need substantial support. If the Office is asked to manually compile information, such as which educational institutions received specific grants, this would reduce the amount of time available to do other tasks related to the data system.

Ongoing Costs: *Costs necessary to maintain collection of the stated data elements for the Office and data providers*.

LEAs and postsecondary institutions would need to devote time each year to tracking the data points on funding, staffing, and distribution of technology tools.

LEAs and postsecondary institutions would need to devote time each year to tracking the data points on funding, staffing, and distribution of technology tools

as provided through statute. Ongoing costs for staff and collection processes and systems for local and state agencies are expected to be less than initial start-up costs but would still be significant once the data collection systems and processes are established and the quality and amount of data is determined to be useful. The cost of establishing these processes may be difficult to quantify and could be substantial both for state agencies, schools, and LEAs.

Annual costs for the Office to compile information could shift unpredictably. For example, several of the federal programs named in the proposal were one-time funds associated with the COVID 19 pandemic. In addition, funding for digital access and infrastructure may come from various sources and grants, so time would need to be spent each year identifying those programs. For example, the National Telecommunications and Information Administration, Federal Communications Commission, the U.S. Department of Agriculture, and the California Public Utilities Commission administer digital access programs.

Compliance

Legal Requirements: *Compliance with privacy laws, intellectual property rights, and any other relevant regulations.*

To change data provider reporting requirements, legislation may be needed to create the authority to collect the requested information.

Scope: *Whether the proposal is consistent with the scope of work that is described in the Cradle-to-Career Act.*

Discussions held during the planning process addressed the need to provide contextual information to support more nuanced understanding of equity gaps among populations. For example, the initial list of data points for the analytical data set include institutional characteristics such as school suspension rates, proportion of high school graduates attaining the College/Career Readiness indicator, or whether the institution participates in a College Promise program. However, specific data points that should be included in the analytical data set are not spelled out in the Cradle-to-Career Act.

Neutrality of the Office: *Whether the proposal might jeopardize the Office's neutral stance.*

The C2C analytical data set does not currently capture information on funding. Including information on supplemental federal funds for digital resources could lead to more requests for including information on grants, which could be used to make arguments about how state resources are allocated or to use the C2C data set to audit state investments.

Suppression and Regulatory Feasibility: *The impact of compliance policies on access to data based on the Office's data suppression policy and other regulatory concerns such as reidentification risks.*

The proposed data points focus primarily on institutional characteristics, rather than individual students, so suppression will not be applied to those data points.

Stewardship and Participation Agreement: *Concerns regarding the participation agreement and stewardship on the proposed data element and its applications.*

The primary concern expressed by data providers is that the information requested is not currently available and would be difficult to document. They noted that using the data request process to force legislative changes to agency reporting and accountability requirements would not be consistent with the governance structure of C2C.

Furthermore, data providers were concerned about the implications of combining survey data with administrative data. While discrepancies between administrative and survey data are often addressed in research studies, attempting to combine these two different types of information in tools like dashboards and query builders may lead those with less training in research methods to draw faulty conclusions. For example, surveys of former students conducted by data providers may have response rates of 30 percent or lower. When researchers examine these results, they evaluate whether the people who responded are representative of the entire student body. Someone who does not have training in survey methods may not realize that broadband access

information could be overstated if people with fewer resources participate in the survey at lower rates.

Cradle-to-Career Data and Tools Advisory Board Proposal Form

Instructions:

Per the Governance Manual proposal forms submitted will address significant gaps regarding whether the data system is providing access to actionable information. Please note there should only be one proposal per form.

Name:

J. Oliver Schak

Proposal Title: *no more than 50 characters*

Student Debt at California Colleges and Universities

- By checking this box, I understand that if my recommendation is one of the top recommendations identified by the Advisory Board, I will be presenting my idea at the Fall Advisory Board meeting.

Type of Proposal[1]:

- Changes to practical tools for students *(Complete section one)*
- Adding data points not available through the P20W data set or adjusting (including the removal of) the existing P20W data points *(Complete section two)*

[1] The C2C Governance Manual notes that DTAB members can also submit recommendations related to the data request process and changes to tools such as dashboards. As the data request process and the dashboards are not yet live, those sections have been removed from the 2024 version of this form.

- Please note a proposal form(s) can recommend adding one data point or several clearly related data points to the data system.

Section One: Changes to Practical Tools

1.What is the nature of the gap regarding access to actionable information?

2. What type of tool should be developed?

3. How would a tool address the gap?

4. Who would be the likely user(s) of the tool?

5. How does the tool relate to the mission and vision of C2C?

Section Two: Adding Data Points Not Available Through the P20W Data Set or Adjusting the Existing P20W Data Points

1.Please state the research question of interest that cannot be fully addressed with the existing data elements in the P20W Data System.

These data points would be used to provide a more complete picture of how students are financing their education and how much cumulative debt they are taking. These data will enable stakeholders to make statements such as "among undergraduate students who received financial aid at the UCs, they borrow $X in federal student loans and $X in non-federal student loans, on average, to help pay for college expenses." This data point could also be used to analyze whether some students are borrowing non-federal, private

loans – which carry higher risks for the borrower – at disproportionate rates. Stakeholders can then find a trend between borrowing specific types of loans by the types of institutions a student attends, and what that means for college affordability and equity statewide. Additionally, this data point can be helpful in informing institutions' internal decision-making (i.e., tuition policies, institutional financial aid allocations, etc.).

2. How does this research question relate to the mission and vision of C2C?

These data will be helpful for students, their families, researchers, as well as advocates. It is critical for students and their families to make sure that they don't need to be overly reliant on loans to finance their education at the institution that they plan on attending. Advocates and researchers can also leverage this data to point out any equity gaps in borrowing rates and identify concerns where borrowing rates are high for especially risky and expensive nonfederal loans.

3. Please propose additional data element(s) needed in order to successfully address the research question of interest.

Type of debt (e.g., federal, non-federal (private loans), undergrad/grad/parent) from all post-secondary segments; Amount of debt from UC and CSU.

4. Please verify that the proposed data element(s) do not already exist in the P20W Data System.

Type of debt does not exist in the current P20W Data System;
Amount of debt exists in the P20W through reporting from CCC, BPPE, and Independents, but is not reported from UC and CSU.

5. If the P20W Data System does not currently include the proposed data element(s), can the proposed data element(s) be derived from the existing data element(s) in the P20W Data System?

They cannot be derived.

6. If the P20W Data System does not currently include the proposed data element(s), are there existing data element(s) closely related to the proposed data element(s)? If so, please list them and why they are not sufficient to answer the proposed research question.

Not closely related to existing data elements.

7. C2C keeps a repository of previous data elements that were either a) considered during the planning process or b) proposed through the mechanisms as outlined in the Governance Manual, but were ultimately not included in the P20W, along with related feasibility studies. Do any of the data element(s) being proposed overlap with the data elements in this repository? (New proposals can build on or duplicate prior proposals. It is helpful for proposers to share that context, including prior related feasibility studies.)

This was proposed by Marshal Anthony Jr. in 2023 but did not advance to the top three to be heard at the Fall 2023 meeting.

8. Are the proposed data element(s) already collected by a state-level entity? For data element(s) already collected, please answer question nine. For data element(s) not collected, please answer question ten.

Yes, this data already exists at the state level because the amount of debt is already being collected by state entities including CCC, BPPE, and independent institutions of higher education. TICAS has also worked with the President's/Chancellor's Offices at both UC and CSU and has requested student debt data from them before. Although TICAS did not ask for data disaggregation by debt type in the past, the data provided by the two systems does include all the debt that the students have that the systems are aware of.

9. [For proposed data element(s) already collected] To the extent possible, please share details pertaining to the proposed data element(s). Relevant details may include but are not limited to: a) corresponding entity that collects

and houses the data element(s); b) specific variable name(s) used in the originating data system; and c) timeframe available.

CCC, BPPE, AICCU, UC, and CSU collect and maintain data on debt amounts. In addition to administering federal student loans, institutions must certify non-federal loans in most cases. Some institutions report amounts for non-federal (e.g., private) loans to Peterson's through the Common Dataset. Peterson's breaks down amount and borrowing rates by loans type, including federal, private, state, and institutional. These data are proprietary and provided to consumers in aggregate.

10. [For proposed data element(s) not collected] Please propose institution(s) that would be most suited for the new data collection effort.

N/A

11. Please explain the desired level(s) of grain size for each data element proposed. (i.e., individual-level, institution-level, or other aggregated levels)? Multiple grain sizes may be requested for each proposed data element.

Individual level for both debt amount and type. If individual level is not available for some elements related to non-federal loans, institution-level reporting should be considered as well.

12. Please explain the intended use case(s) for the proposed data elements (i.e., dashboards, query builder, or the research request tool)? Multiple use cases may be requested for each proposed data element.

More complete information on student debt would help power dashboards and research queries focused on college affordability. A dashboard on affordability could show the typical amount of debt among recent graduates, and how likely they were to leave school with any debt. Such a dashboard would assist students and families in navigating educational pathways and financial decisions. And the tool would lift up less well-known information, since non-federal debt amounts are not readily available from existing tools like the College Scorecard.

With individual level data, debt amounts could be broken down by race and indicators poverty and wealth. Leaders could use this information to elevate and refine efforts to address inequities in college opportunity and affordability. Researchers could use queries and data requests to ask questions about how debt influences student success and experiences through college, and how better policies and practices could improve the mix of financial aid that students use to pay for school.

Type of Debt Data

Voting Option for Fall 2024 Proposals from the Data and Tool Advisory Board

This document provides background information to support prioritization of potential new data points for the Cradle-to-Career (C2C) Data System P20W data set. P20W data points include data uploaded directly by data providers, as well as metrics that are calculated from the information shared by data providers. P20W data points will be used to populate dashboards and a query builder tool. When the data request process becomes available, users may request other formulations of metrics that are generated from the information shared by data providers.

Staff from the Office of Cradle to Career Data (Office) worked with Advisory Board members who submitted proposals to provide greater specificity about the suggested item.

Proposal Name

Student Debt at California Colleges and Universities

Type of Data Point

Based on an initial review of the proposal, the Office notes that the amounts of several types of debt for CCC, CSU, and UC students are currently in the P20W data set. However, there is not detailed information on the types of federal loans. Therefore, the proposal provides recommendations on new federal loan data elements and on how to create dashboard visualizations.

Proposed Data Point Construction

In consultation with the Advisory Board member who proposed the data point, the Office operationalized the proposal to provide information on types of debt by creating 2 metrics.

1) **Federal Loans to Students** - The postsecondary student received a direct federal loan
   Example: A 28 year old student received a $4,000 federal direct loan each year to help cover the cost of earning a bachelor's degree in accounting at CSU Dominguez Hills. He graduates after five years. The amount of federal direct loans would be $20,000.

2) **Federal Loans to Parents -** The parents of the postsecondary student received a federal loan
   Example: The parents of an 18 year old student receive $10,500 per year to cover the costs of her attendance at UC Berkeley to earn a bachelor's in political science. She graduates in four years. The amount of federal loan to her parents would be $42,000.

Data Points for the Proposed Visualization

C2C would develop a dashboard that shows the amount of different types of debt held by college graduates and students who have exited higher education, based on metrics already in the P20W data set. The data points that would be utilized are:

- Amount of institutional loans
- Amount of state loans
- Amount of federal loans to students
- Amount of federal loans to parents
- Amount of private loans

Proposed Visualization

The visualization would be included in the Financial Aid Dashboard.

Users would be able to see:

- The amount of loans in each of the five loan categories

The dashboard would allow users to disaggregate results by the following characteristics:

- Institution attended
- Age bracket
- Gender
- Race/ethnicity
- K-12 socioeconomically disadvantaged status
- Program of study
- Graduation status
- Debt at the point of bachelor's degree attainment
- Debt incurred between completing a bachelor's degree and exiting a graduate degree program

Feasibility Study:
Expanding Financial Aid Data

At the October 1, 2024 meeting, the Data and Tools Advisory Board advanced a proposal from J. Oliver Schak to expand financial aid data in the Cradle to Career Data System (C2C) analytical data set. Specifically, the California Community Colleges (CCC), California State University (CSU), University of California (UC), and the California Student Aid Commission (CSAC) were requested to provide information on:

- for each student, the amount of debt from Subsidized loans, Unsubsidized loans, and Grad PLUS loans[1]
- for each student, the amount of federal loans given to parents through a Parent PLUS loan versus loans made to students
- for each student, the amount of debt they take out each year (particularly to distinguish debt before and after students transition from undergraduate to graduate education)

The full text of the proposal can be found at this link.

When conducting feasibility studies, the Office of Cradle-to-Career Data (Office) considers four factors: data availability, data reliability and data validity, cost, and compliance. Information on each aspect is included below.

Specific Data Points

If approved, the data points that would be added to the C2C analytical data set would be:

[1] The proposal discussed by the Data & Tools Advisory Board referenced Subsidized and Unsubsidized Stafford Loans. Stafford loans were provided under the Federal Family Education Loan Program, which ended in July 2010. All federal loans are now made through the Federal Direct Loan Program. However, many people still use the term "Stafford loan" to refer to these programs. Therefore, for the purpose of this feasibility study, the Office examined the feasibility of collecting information on Direct Subsidized and Unsubsidized Loans.

- Amount of Subsidized loans received by a student within an academic year
- Amount of Unsubsidized loans received by a student within an academic year
- Amount of Grad PLUS loans received by a student within an academic year
- Amount of Parent PLUS loans received by family member for a student's undergraduate education within an academic year

Note on Terms

Subsidized and Unsubsidized loans: Subsidized loans refer to loans that are based on financial need, where the federal government pays the interest on a loan while a student is still in school and during a grace period after leaving school. Unsubsidized loans mean that students do not demonstrate financial need and are responsible for all interest payment.

Parent and Grad PLUS loans: Parent PLUS loans are given to the parents of students to help cover the cost of college. Grad PLUS loans are given directly to graduate students.

Data Availability

Data Sources and Historical Range: *Where the data originates from and the span of years for which data are available.*

Individual postsecondary institutions receive information on federal loans from the National Student Loan Data System (NSLDS). Access to NSLDS is strictly limited to people with specific job responsibilities related to administering financial aid. Local institutions report this information to state agencies for the purpose of improving aid distribution.

CSAC receives information about federal loans through the Institutional Student Information Record (ISIR). There are strict regulations about how state financial aid agencies may use this information, including specifying that information must pertain to the administration of state financial aid.

CCC currently collects high-quality information on Subsidized, Unsubsidized, and Parent PLUS loans. Information on the amount of debt per loan type is available, so it would be possible to distinguish between amounts given for Subsidized, Unsubsidized, and Parent PLUS loans. It is also able to provide information on the total amount of loans that a student takes out each year, which could be used to calculate the amount of debt each student had taken out at the end of their time at community college (although this would not include how that debt might have compounded over time).

Both CSU and UC currently collect information on Subsidized, Unsubsidized, and PLUS loans. They do not distinguish between Parent PLUS and Grad PLUS loans, but this could be inferred by determining the program in which the student is enrolled. Data is of high quality starting in 2017-18[2].

CSAC receives information on students' aggregate Subsidized and Unsubsidized loan amounts to support decisions about providing state financial aid, so it does not have definitive information on the amount of loans students receive each year. For example, CSAC has access to the principal balance and pending disbursements but not annual awards. In addition, they do not have information on cumulative debt, which is impacted by factors such as forbearance policies that allow people to defer or reduce payments when changing jobs or facing financial difficulties.[3]

CSAC does not receive information on Parent PLUS and Grad PLUS loans.

Information on student loan status is only provided to CSAC when a student submits an application for additional financial aid. For example, if a student took out a loan each year while pursuing a bachelor's degree and graduated after five years, CSAC would not have information on the loan amount for the final year. In addition, if a student has consolidated their loans, it may not be easy to discern the amount of the loans that were Subsidized or Unsubsidized.

CSAC does not currently store information on students' loan amounts. If CSAC were to begin storing this information, it would not be available retrospectively.

[2] UC's data is of high quality starting in 2002-03.

[3] See https://studentaid.gov/manage-loans/lower-payments/get-temporary-relief/forbearance

Future Relevance and Data Availability: *Evaluate whether the new data will remain useful as the analytical dataset evolves and if the information is likely to be collected in the future.*

Policy makers, advocates, and families all express concern about the cost of college and research shows that significant debt burdens erode the economic mobility associated with postsecondary education. As a result, many researchers, states, and the federal government are developing Return on Investment indicators that take debt into account when calculating the value of education. It is likely that information on debt will continue to be collected and will be of value to numerous interest holders in the future.

Data Reliability and Data Validity

Institutional Variability: *Whether there are variations in administrative practices and data recording across institutions at the local level.*

Because information is derived from a federal file that is provided in a consistent format to postsecondary institutions, there should be little variability across individual CCC, CSU, and UC institutions.

Agency Variability: *Data and metric definitions across various agencies follow a uniform format and standard.*

Although source data are similar for CCC, CSU, and UC, it will be challenging to calculate debt by year. State agencies differ in whether they count summers as a leading term or a trailing term. As a result, debt incurred to pay for summer enrollment may not be reported consistently. CCC and UC treat summers as a leading term. CSU is currently working to address this challenge by shifting its collection practices to providing information by term, but this change has not yet been implemented.

As noted above, the source file and format of information on loans that CSAC receives is different from the data received by postsecondary institutions and agencies, because it looks at aggregate loan amounts rather than annual loan distribution, and does not include Parent PLUS and Grad PLUS loans.

Data Integration Across Agencies: *The Office's ability to consolidate data from multiple agencies.*

CCC, CSU, and UC already provide a number of data points on financial aid for the P20W analytical data set, including the amount and type of financial aid received.

Currently, the format that CCC, CSU, and UC use to provide data on financial aid to the Office is modeled after the format used by CCC. This format could be extended for CSU and UC to report on Subsidized, Unsubsidized, Parent PLUS, and Grad PLUS loans. Because information would be an extension of existing files, it should be feasible for the Office to integrate additional information from CCC, CSU, and UC.

Integrating information from CSAC would be significantly more difficult because new file formats would be needed and methodological changes would be required to account for differences in how loan information is structured.

Cost

Startup Costs: *Costs to begin collection including both direct costs and associated staff time for the Office and its data providers.*

Costs for adapting the file structures to include information on the additional variables should be minimal for CCC, CSU, UC, and the Office.

Costs for CSAC to collect additional information would be significant. CSAC would need to adapt their underlying database to include additional information and change the structure of their import from the federal ISIR file. Because of the way that ISIR data is provided, which breaks out the relevant information into multiple data points, millions of additional rows of data would need to be stored, which would require upgrades to CSAC's infrastructure.

Ongoing Costs: *Costs necessary to maintain collection of the stated data elements for the Office and data providers*.

Once the new file formats and data upload mechanisms are established, there should not be significant ongoing costs.

Compliance

Legal Requirements: *Compliance with privacy laws, intellectual property rights, and any other relevant regulations.*

Guidance provided in the Higher Education Act places limitations on when financial aid data originating from the Free Application for Federal Student Aid (FAFSA) can be shared. As a result, concerns about including financial aid data in the C2C analytical data set were noted during the planning process. In November 2024, the federal government provided guidance that indicates financial aid data may be used for research conducted by or on behalf of education institutions, higher education agencies, and state financial aid agencies to promote college attendance, persistence, and completion.

Affected data providers highlighted the potential risk to federal student aid availability if additional data elements related to financial aid are provided to the P20W system. Data providers cited that they cannot provide additional information regarding aid related elements requested as part of the feasibility study.

Multiple data providers also highlighted that federal guidance continues to evolve on this subject, noting the risk of loss of federal aid outweighs any possible benefit derived from providing the more granular data to C2C.

Scope: *Whether the proposal is consistent with the scope of work that is described in the Cradle-to-Career Act.*

The initial legislation prioritized providing information on "college access, completion, and long-term effects of access to state financial aid." Providing additional information on federal loans would be consistent with the discussions held during the planning process regarding better understanding the impact of debt and the content for the proposed financial aid dashboard that was included in the reports to the legislature during the C2C planning process. However, specific data points that should be included in the analytical data set are not spelled out in the Cradle-to-Career Act.

Neutrality of the Office: *Whether the proposal might jeopardize the Office's neutral stance.*

Given that examination of financial aid was part of the original purview of C2C, providing additional financial aid information is not likely to impact its neutrality.

UC notes the P20W system currently has all the data it needs to track student affordability and student loan debt per prevailing institutional and federal reporting rules that consider federal direct subsidized and unsubsidized loans. UC expressed that tracking the borrowing of parents falls outside existing reporting norms and may cause confusion for consumers of this information. The Office notes these concerns and highlights it as a matter for the Advisory Board's and Governing Board's consideration.

Suppression and Regulatory Feasibility: *The impact of compliance policies on access to data based on the Office's data suppression policy and other regulatory concerns such as reidentification risks.*

Given that only a subset of students take out Subsidized, Unsubsidized, Parent PLUS, and Grad PLUS loans, it is likely that information will be suppressed at the institutional level, particularly when disaggregating results for specific types of students. For example, in the UC system, only 5% of students have Parent PLUS loans.

Stewardship and Participation Agreement: *Concerns regarding the participation agreement and stewardship on the proposed data element and its applications.*

CSAC has indicated its primary concern is the legality of sharing the requested information. As per the Participation Agreement, agencies may not share information in violation of federal rules. CSAC further underscored they do not own this data and are unable to report on the proposed data points.

UC has expressed concern that including data on Parent PLUS loans may cause confusion because this information is excluded from its financial aid dashboards and from federal reporting, per directions from the Integrated Postsecondary Education Data System. The Office notes that as with many of its other planned

dashboards, its reporting may not align exactly with the federal reporting requirements of its data providers. If these data points are provisioned, their uses would include not just potential dashboards but also help facilitate research in the domains of financial aid.

Cradle-to-Career Data and Tools Advisory Board Proposal Form

Instructions:

Per the Governance Manual proposal forms submitted will address significant gaps regarding whether the data system is providing access to actionable information. Please note there should only be one proposal per form.

Name:

Jennifer Orlick and Laura Owen

Proposal Title: *no more than 50 characters*

Weaving Disaggregated Multilingual Learner Data into C2C

Type of Proposal[1]:

- Changes to practical tools for students *(Complete section one)*
- Adding data points not available through the P20W data set or adjusting (including the removal of) the existing P20W data points *(Complete section two)*
    - Please note a proposal form(s) can recommend adding one data point or several clearly related data points to the data system.

Section One: Changes to Practical Tools

1.What is the nature of the gap regarding access to actionable information?

[1] The C2C Governance Manual notes that DTAB members can also submit recommendations related to the data request process and changes to tools such as dashboards. As the data request process and the dashboards are not yet live, those sections have been removed from the 2024 version of this form.

2. What type of tool should be developed?

3. How would a tool address the gap?

4. Who would be the likely user(s) of the tool?

5. How does the tool relate to the mission and vision of C2C?

Section Two: Adding Data Points Not Available Through the P20W Data Set or Adjusting the Existing P20W Data Points

1.Please state the research question of interest that cannot be fully addressed with the existing data elements in the P20W Data System.

- How does Dual Language Learner (DLL), long-term English learner (LTEL) status, Newcomer, Dually identified English Learners, and At-risk English learner (ARLTEL) status correlate with academic performance throughout a student's educational journey, from early learning and care to high school graduation?
- How does long-term English learner (LTEL) status impact post-secondary education enrollment completion rates, employment prospects, and overall socio-economic outcomes?
- What are the educational trajectories of at-risk, dually identified English Learners, and long-term English learners (LTELs) compared to other student populations, such as reclassified fluent English proficient (RFEP) or non-English learners? Specifically, what are the differential outcomes

between different EL types, including LTELs in K-12, postsecondary, and workforce?

- Where are there concentrations of DLLs, ARELs, Newcomers, and LTELs across the state?
- For dually identified ARLTELs and LTELs what disabilities trigger special education identification?
- What services and differentiated instruction are provided to ARLTELs and LTELs?
- For English learners who are identified within Specific Learning Disabilities, what disability led to the identification?
- What are the enrollment procedures for older newcomer students?
- What language approaches are provided to Dual Language Learners?
- In what program models are ELs enrolled?

2. How does this research question relate to the mission and vision of C2C?

The educational experiences and outcomes among multilingual learners in California's educational systems differ, with some students being supported to develop English proficiency and others remaining as English learners for extended periods of time. Data on outcome metrics like math and English Language arts proficiency reveal disparities between students who are designated as Long-Term English Learners and their peers who are reclassified as English proficient, and newly disaggregated data on students at-risk of becoming long-term ELs will likely shed additional light on nuanced educational needs and experiences. With this in mind, these research questions included in this proposal relating to including additional information on students' EL status beyond what is currently planned for inclusion in the P20W data system) directly support C2C's mission to serve as a source of actionable data on educational outcomes. These questions and classifications also support C2C's vision of providing insights into critical milestones on students' educational trajectory. The ability to trace the educational trajectory of students designated as any of the above classification, both preceding and following their assumption of said designations, can equip educators and policymakers with information needed to better support multilingual learners to reach their full potential.

3. Please propose additional data element(s) needed in order to successfully address the research question of interest.

- Long-term English learner (LTEL) classification
- At-risk English learner classification (ARLTEL)
- Newcomers
- Dual Language Learners (DLLs)
- Dually Identified (ELs with IEPs)

4. Please verify that the proposed data element(s) do not already exist in the P20W Data System.

The data system currently only includes:
(see website: Childhood English Language Learner - California Cradle-to-Career Data System | C2C)
Display Options

- Never an English language learner
- Assessed and determined to be proficient in English
- English language learner
- English language learner who became proficient
- Not Reported

Available Years

- Early learning and care data: 2009-10
- K-12 data: N/A
- Postsecondary data: N/A

Time Frame

- If ever: Any time up to and including the selected academic year
- School Year: July 1 – June 30

Geography/Locale

- Display level selected (such as institution, region, statewide)

5. If the P20W Data System does not currently include the proposed data element(s), can the proposed data element(s) be derived from the existing data element(s) in the P20W Data System?

Yes, LTEL status and At-risk EL status can be derived and integrated into C2C in two ways:

- Calculated by the number of years a student has been reported as an English learner in the school system (LTEL=6 or more years; AREL=4-5 years) and adding two classifications for these two subgroups
- Integrated directly from CDE's available EL-status data and adding two additional classifications for these two subgroups

Newcomer, Dual Language Learner, and Dually Identified ELs:

- Newcomers and Dual Language Learners now have been defined and data will be available in 2025
- Dually identified for IEPs has been collected by CDE as well.

6. If the P20W Data System does not currently include the proposed data element(s), are there existing data element(s) closely related to the proposed data element(s)? If so, please list them and why they are not sufficient to answer the proposed research question.

Currently, the P20W data point included is "English language learner" followed by a "Time Frame" that is defined as "If ever: Any time up to and including the selected academic year" which may include the number of years that a student has been designated as an English learner. However, these two elements may not be sufficient in answering the proposed research questions if the two elements are not included in the dashboards and tools to fully answer LTEL, ARLTEL, DLL, etc. specific questions.

7. C2C keeps a repository of previous data elements that were either a) considered during the planning process or b) proposed through the mechanisms as outlined in the Governance Manual, but were ultimately not included in the P20W, along with related feasibility studies. Do any of the data element(s) being proposed overlap with the data elements in this repository? (New proposals can build on or duplicate prior proposals. It is helpful for proposers to share that context, including prior related feasibility studies.)

No

8. Are the proposed data element(s) already collected by a state-level entity? For data element(s) already collected, please answer question nine. For data element(s) not collected, please answer question ten.

Yes, this data is housed within CDE. Additionally, classifications for DLLs will be collected and housed by DSS.

9. [For proposed data element(s) already collected] To the extent possible, please share details pertaining to the proposed data element(s). Relevant details may include but are not limited to: a) corresponding entity that collects and houses the data element(s); b) specific variable name(s) used in the originating data system; and c) timeframe available.

- CDE collects and houses At-Risk and LTEL data. In 2022, Governor Newsome also signed AB 1868 (2022) which required CDE to report disaggregated standardized test scores in English language arts, math, and science for different subgroups of ELs, including long-term English learners (LTELs), those at risk of becoming long-term English learners, current ELs, and ELs that have been reclassified. It also required the department to report how many English learners have been dual-identified as having a disability.
- "At-Risk" and Long-Term English Learners (LTEL)
- 2015-16 school year
- AB 714 (2023) requires the identification of newcomers
- AB 393 (2023) requires the identification of DLLs

See: At-Risk and Long-Term English Learners (LTEL) by Grade (ca.gov)

10. [For proposed data element(s) not collected] Please propose institution(s) that would be most suited for the new data collection effort.

For dual language learners, DSS may best positioned to collect this data during the built out of the Early Childhood Integrated Data System (ECIDS).

11. Please explain the desired level(s) of grain size for each data element proposed. (i.e., individual-level, institution-level, or other aggregated levels)? Multiple grain sizes may be requested for each proposed data element.

- Individual-level (Student characteristic)
- Individual School, District, County, State

12. Please explain the intended use case(s) for the proposed data elements (i.e., dashboards, query builder, or the research request tool)? Multiple use cases may be requested for each proposed data element.

These data elements (LTEL, At-risk LTEL, newcomer, DLL and dually identified status) could be displayed as data points and disaggregation options in the Student Pathways Diagram, dashboards, query builder, and the research request tool.

Multilingual Learner Data

Voting Option for Fall 2024 Proposals from the Data and Tool Advisory Board

This document provides background information to support prioritization of potential new data points for the Cradle-to-Career (C2C) Data System P20W data set. P20W data points include data uploaded directly by data providers, as well as metrics that are calculated from the information shared by data providers. P20W data points will be used to populate dashboards and a query builder tool. When the data request process becomes available, users may request other formulations of metrics that are generated from the information shared by data providers.

Staff from the Office of Cradle to Career Data (Office) worked with Advisory Board members who submitted proposals to provide greater specificity about the suggested item.

Proposal Name

Weaving Disaggregated Multilingual Learner Data into C2C

Type of Data Point

Based on an initial review of the proposal, the Office notes that these items are data points that are not currently in the P20W data set.

Proposed Data Point Construction

In consultation with the Advisory Board member who proposed the data point, the Office operationalized the proposal to expand information on English language learners by creating 5 metrics that are in alignment with definitions currently used by the California Department of Education:

1) **Long-term English Learner (LTEL)** - Students in grades 6-12 who have been enrolled for 7 or more years and are still designated as English language learners.
   For example: A student who moved to California from Vietnam was flagged as an English language learner in 4th grade and has not attained English proficiency by 11th grade. This student would be defined as a Long-term English learner.

2) **At-Risk English Learner (ARLTEL)** - Students in grades 3-12 who have been enrolled in a U.S. school and are still designated as English language learners for 4 to 6 years.
   For example: A student who moved to California from Syria was flagged as an English language learner in 1st grade and has not attained English proficiency by 5th grade. This student would be defined as an at-risk English learner.

3) **Newcomer** - Students in their first three years of enrollment at a K-12 school in the U.S. who are designated as English language learners.
   For example: A student moves to California from Guatemala and enrolls in 4th grade. At that time, the student was flagged as an English language learner. This student would be defined as a Newcomer.

4) **Dual Language Learner (DLLs)** - Children ages birth to 5 who are learning two or more languages at the same time, where this data is reported.
   For example: A preschool student lives in a home where both Tagalog and English are spoken. This student would be defined as a Dual Language Learner.

5) **Dually Identified -** Students who are English language learners who also have an Individualized Education Program to address a disability
   For example: A student moves to California from Malawi and enrolls in 10th grade. The student is classified as an English language learner and has an IEP to provide materials in Braille because the student is visually impaired. This student would be defined as dually identified.

Feasibility Study: Multilingual Learners

At the October 1, 2024 meeting, the Data and Tools Advisory Board advanced a proposal from Jennifer Orlick and Laura Owen to include information on multilingual learners in the Cradle to Career Data System (C2C) analytical data set. Specifically, the California Department of Education (CDE) and the Department of Social Services (CDSS) were requested to provide information on:

- Students with Long-term English Learner status (LTEL)
- Students with At-Risk English Learner status (ARLTEL)
- Students with newcomer status
- Students who were in an immersion K-12 classroom
- Students with Dually Identified status (in both preschool and K-12)
- Students with Dual Language Learner (DLLs) status for children in Title V subsidized early learning programs prior to 2025
- Variables from the California Preschool Data Collection (CAPSDAC) system, including:
    - Language the Lead Teacher Uses
    - Languages the Lead Teacher is Proficient in
    - Languages Other Staff Use
    - Languages Other Staff are Proficient In
    - Language Program Type

The full text of the proposal can be found at this link.

When conducting feasibility studies, the Office of Cradle-to-Career Data (Office) considers four factors: data availability, data reliability and data validity, cost, and compliance. Information on each aspect is included below.

Note on Terms

The data points being considered are defined as follows:

- Students with Long-term English Learner status (LTEL)
    - Students in grades 6-12 who have been enrolled for 7 or more years and are still designated as English language learners.
- Students with At-Risk English Learner status (ARLTEL)
    - Students in grades 3-12 who have been enrolled in a U.S. school and are still designated as English language learners for 4 to 6 years.
- Students with newcomer status
    - Students in their first three years of enrollment at a K-12 school in the U.S. who are designated as English language learners.
- Students who were in an immersion K-12 classroom
    - 50/50 program or Other Language Allotment majority non-English program
- Students with Dually Identified status (in both preschool and K-12)
    - Students who are English language learners who also have an Individualized Education Program to address a disability
- Students with Dual Language Learner (DLLs) status for children in Title V subsidized early learning programs prior to 2025
    - Children ages birth to 5 who are learning two or more languages at the same time, where this data is reported.

While the proposal used the term "dual language learner" for children in early learning programs, children are usually described as being multi-language learners before starting kindergarten. This designation acknowledges that in their first five years, people are still gaining fluency in all languages to which they are exposed.

Data Availability

Data Sources and Historical Range: *Where the data originates from and the span of years for which data are available.*

Long-Term English Learners and At-Risk English Learners
CDE does not have information on students' long-term English language status from birth to age 5, given that children of all backgrounds are still gaining English proficiency during their early years.

Starting in kindergarten, the agency tracks Students with Long-term English Learner (LTEL) and At-Risk English Learner (ARTEL) status, but the definitions do not align with the ones proposed for the C2C analytical data set. For example, when CDE calculates LTEL and ARTEL status, they include variables such as prior test scores, in addition to the number of years the student has been classified as an English learner.

However, CDE provides information for the C2C analytical data set that indicates the year in which a student was identified as an English language learner. The Office staff could calculate LTEL and ARTEL status according to the proposed definitions and clarify that this varies from ways these variables are calculated by CDE.

Alternatively, CDE could provide flags on students who meet its definition of LTEL[1] students as part of its submission to the C2C analytical data set.

Newcomers
The concept of newcomer status is defined by California Assembly Bill 714 in the same way as the U.S. code for immigrant children and youth:

- aged three through twenty-one;
- not born in any U.S. state (each of the 50 states, the District of Columbia, and the Commonwealth of Puerto Rico); and
- not attending one or more schools in the United States for more than three full academic years.[2]

Subcategories of newcomers include English learners, immigrant children and youth, migratory children, refugees, students who may have experienced

[1] See https://www.cde.ca.gov/ds/ad/lteldef.asp
[2] Section 7011(5) of Title 20

limited or interrupted formal education, and unaccompanied children who are refugees or undocumented.[3]

Currently, CDE only collects information on English language learners and children who have participated in migrant education programs. It already provides information on English language learner status and children in migrant education programs for the C2C analytical data set. The data CDE provides to the P20W cannot be used to identify Newcomer students as defined in Section 7011(5) of Title 20 of the United States Code. CDE is currently reviewing existing data elements related to identification of these students. Currently,CDE is unable to provide further information on the feasibility of adding new data to the P20W for this purpose.

Given the proposed definition, the Office staff could calculate students designated as English language learners who have been enrolled in California public schools for three years or less.

Immersion K-12 Programs

This information is not available. CDE provides information on children who are receiving English language services and language of instruction, but it does not track information for immersion classrooms that contain both native-English language and English-language learner students. CDE does not collect data from LEAs that lend themselves to meeting the definition as included in the proposal.

Dually Identified

CDSS does not track information that could be used to calculate this metric.

Currently, CDE does not identify preschoolers who are flagged both as dual language learners and having an Individualized Education Plan (IEP) to address a disability at the student level. While work is underway to collect this information, it will not be available for several years.

Starting with children of kindergarten age, the Office staff could identify dually enrolled students using the data points that CDE provides on English language learners and children with IEPs, but these figures may not align with data posted

[3] See the CDE website at https://www.cde.ca.gov/sp/ml/newcomerstudnts.asp

by CDE. CDE currently makes their reporting on this information available via DataQuest.

Language Learners in Title V Subsidized Programs

CDSS has comprehensive information on the language spoken at home by people who receive public benefits for childcare, foster care, Medicaid, CalWORKS, CalFresh, and In-Home Supportive Services, which can be used to identify English language learners. This information is already included in the C2C analytical data set. The data points that CDSS already provides include:

- Early learning and care primary language
- Race/ethnicity
- Sex/gender
- Foster youth
- Child with a disability

CDE's California Preschool Data Collection System

CDE is currently building the second phase of the preschool data collection system and will not have responsive data by the end of 2025. Until the system is fully operational, LEAs have adapted to reporting in the new system, and the data have been deemed valid and reliable, CDE will not be able to provide preschool data.

Future Relevance and Data Availability: *Evaluate whether the new data will remain useful as the analytical dataset evolves and if the information is likely to be collected in the future.*

The issue of language skills is likely to continue to be important, but with a growing emphasis on how language skills relate to immigration status. The definitions proposed do not expressly identify factors like documentation or refugee status, but language skills may become conflated with immigrant status.

Many of the requested data points on English language learner status and disability status are already available in the C2C analytical data set or could be calculated by the Office. This information is likely to remain available.

CDE is unlikely to collect information on dual-immersion classrooms, as this is not an area where they have authority to require reporting from LEAs.

Data Reliability and Data Validity

Institutional Variability: *Whether there are variations in administrative practices and data recording across institutions at the local level.*

Many of the requested data points could be calculated using data that are already collected at the state level. Therefore information is likely to be reasonably consistent across local institutions. This includes data points from CDE on Long-Term English Learners, At-Risk English Learners, Newcomers, Dually Enrolled Children (starting in kindergarten), and from CDSS on English Language Learners.

However, information on newcomers may be incorrect at the local level due to lack of data sharing agreements across state lines. For example, a child who enrolls in a California school for the first time in grade 3 and is not fluent in English but who was born in and previously attended school in Arizona might be flagged as a newcomer.

Similarly, if the Office calculates LTEL and ARTEL status, the measures would only be based on enrollment in public California K-12 institutions. If, for example, a student in the 11th grade is an English language learner, but had moved back and forth between California and Arizona multiple times, they would not be included in the LTEL or ARTEL metrics.

Agency Variability: *Data and metric definitions across various agencies follow a uniform format and standard.*
Data definitions for CDE and CDSS may not always align. For example, CDE uses a combination of surveys and standardized assessments to determine English language proficiency. CDSS evaluates language skills by household declaration[4][5].

---

4 https://www.cdss.ca.gov/Portals/9/Additional-Resources/Forms-and-Brochures/2020/A-D/CCD26.pdf?ver=2023-10-06-141046-190

5 https://cdss.ca.gov/inforesources/child-care-and-development/contractor-resources/cdss-cdmis-support/cdmis-manual/appendix-a-data-definitions/child-is-english-learner

For CDSS, the term "newcomer" is not based solely on language skills because it is used in the context of immigration and social services and may take into account a child's refugee status or the date of entry into the country.

Data Integration Across Agencies: *The Office's ability to consolidate data from multiple agencies.*

The Office is already combining information from CDE and CDSS to more fully document the experiences of individuals. The primary challenge will be implementing definitions that differ from data produced by CDE and CDSS. For example, there may be some confusion about long-term English learners because state and federal statutes differ in the number of years a student should be classified as an English learner--some say 5 and some say 6, while others do not set a timeframe. There will also be differences based on how CDE and CDSS identify newcomer students.

Cost

Startup Costs: *Costs to begin collection including both direct costs and associated staff time for the Office and its data providers.*

Costs for the Office would focus on identifying how to calculate new metrics and validating the approach with data providers and experts. The staff would also need to spend time developing clarifying language regarding why data shown in the C2C analytical data set does not align with similar information produced by CDE and CDSS. Time spent on these tasks would need to be balanced by reducing time spent on other aspects of managing the C2C analytical data set.

The Office notes that provisioning LTEL and ARTEL data as defined by CDE's definitions would create an increased workload for both the CDE and the C2C. Additionally, it may increase the cost to the P20W system to ingest new data, re-ingest old data, and potentially update all the reports and tables that have been built off the previously submitted files. CDE requested that the C2C office estimate a cost for this work which should include the estimated staff FTE hours to update the file specifications document and update the system to ingest new data. After this, the CDE will calculate the cost for the work of creating flags across all years of data, recompiling and submitting these data, and working

with the C2C team to update the file specifications document as part of determining feasibility.

Ongoing Costs: *Costs necessary to maintain collection of the stated data elements for the Office and data providers*.

Once the new metrics are calculated and validated, ongoing costs should not be significant unless there were changes in the data collection process or elements of the data providers.

Compliance

Legal Requirements: *Compliance with privacy laws, intellectual property rights, and any other relevant regulations.*

All of the data in the C2C P20W dataset are subject to relevant state and federal privacy laws.

Scope: *Whether the proposal is consistent with the scope of work that is described in the Cradle-to-Career Act.*

The initial legislation prioritized providing information on the "impact of early education on student success and achievement as a student progresses through education segments and the workforce." Providing information on English language fluency in pre-K and K-12 would be consistent with the dashboards that were proposed to the legislature, including on early education (which lists outcomes like whether students are likely to be classified as English language learners once in K-12) and on college and career readiness (which proposes disaggregating outcomes based on whether a student was a childhood English language learner). However, specific data points that should be included in the analytical data set are not spelled out in the Cradle-to-Career Act.

In 2020-21, during the planning process for C2C, the working group discussed whether information on immigration status should be included in the P20W analytical data set. There was a conscious choice to not include this data point. For example, the Research Agenda Subcommittee elected to not distinguish between students who applied for federal financial aid and those who applied

for state financial aid for undocumented students who were brought to California as children.

Neutrality of the Office: *Whether the proposal might jeopardize the Office's neutral stance.*

Given that data points regarding English language learner status were part of the original purview of C2C, providing additional information on this subject is not likely to impact its neutrality.

Suppression and Regulatory Feasibility: *The impact of compliance policies on access to data based on the Office's data suppression policy and other regulatory concerns such as reidentification risks.*

Suppression protocols implemented by the Office should be sufficient to protect the privacy of individuals in public-facing tools and exceed suppression protocols used by both CDE and CDSS. However, given the sensitivity of data on immigration status, there is likely to be extra scrutiny on whether people who are flagged as English language learners could be reidentified.

Stewardship and Participation Agreement: *Concerns regarding the participation agreement and stewardship on the proposed data element and its applications.*

Given the sensitivity of data elements pertaining to immigration status, data providers often constrain the information shared within their agencies and so may choose to not provide this information to other state entities like the Office.

Cradle-to-Career Data and Tools Advisory Board Proposal Form

Instructions:

Per the Governance Manual proposal forms submitted will address significant gaps regarding whether the data system is providing access to actionable information. Please note there should only be one proposal per form.

Name:

J. Oliver Schak

Proposal Title: *no more than 50 characters*

Graduate School Completion Rates by Field of Study

- By checking this box, I understand that if my recommendation is one of the top recommendations identified by the Advisory Board, I will be presenting my idea at the Fall Advisory Board meeting.

Type of Proposal[1]:

- Changes to practical tools for students *(Complete section one)*
- Adding data points not available through the P20W data set or adjusting (including the removal of) the existing P20W data points *(Complete section two)*

[1] The C2C Governance Manual notes that DTAB members can also submit recommendations related to the data request process and changes to tools such as dashboards. As the data request process and the dashboards are not yet live, those sections have been removed from the 2024 version of this form.

- Please note a proposal form(s) can recommend adding one data point or several clearly related data points to the data system.

Section One: Changes to Practical Tools

1.What is the nature of the gap regarding access to actionable information?

2. What type of tool should be developed?

3. How would a tool address the gap?

4. Who would be the likely user(s) of the tool?

5. How does the tool relate to the mission and vision of C2C?

Section Two: Adding Data Points Not Available Through the P20W Data Set or Adjusting the Existing P20W Data Points

1.Please state the research question of interest that cannot be fully addressed with the existing data elements in the P20W Data System.

Little is known about the share of graduate-level students who successfully complete their degree. Federal data on persistence and degree completion is in most cases limited to undergraduates. Students, leaders, and community members deserve to know more about how well graduate schools see students through to graduation, and how quickly students complete their intended degree.

Research has documented inequities in graduate school success for women and students of color, particularly at programs that focus on STEM (see Posselt, 2020, "Equity in Science"). Outcome data that's broken down by school, field of study, gender, and race and ethnicity are critical elements to informing this knowledge base.

2. How does this research question relate to the mission and vision of C2C?

The availability of graduate school success rates would empower students to better navigate educational options beyond a four-year degree. Removing gaps in completion data would spotlight programs that best support the success of graduate students of all backgrounds and facilitate conversations about evidence-based improvements at graduate schools, spanning the state and nation.

3. Please propose additional data element(s) needed in order to successfully address the research question of interest.

The P20 Data System should add detailed enrollment information on graduate students by postsecondary school and field of study (e.g., 5-digit CIP or major), including indicators on when and where students started their graduate-level studies. Collected data points should be sufficient to enable the construction of program-level outcomes cohorts and student persistence and success metrics (e.g., share of students who earn their degree within six years of program entry).

4. Please verify that the proposed data element(s) do not already exist in the P20W Data System.

These do not exist, as detailed information about enrollment by field of study is only available for undergraduate students (based on website documentation).

5. If the P20W Data System does not currently include the proposed data element(s), can the proposed data element(s) be derived from the existing data element(s) in the P20W Data System?

No, derivation requires the construction of graduate-level enrolment cohorts by school and field of study.

6. If the P20W Data System does not currently include the proposed data element(s), are there existing data element(s) closely related to the proposed data element(s)? If so, please list them and why they are not sufficient to answer the proposed research question.

There are similar elements for undergraduate students only. The system also has information on who earns a graduate each year, but this does not tell us who did and did not complete their program, who remains enrolled in their program after starting their studies, and how much time it typically takes to complete a specific masters or Ph.D. program.

7. C2C keeps a repository of previous data elements that were either a) considered during the planning process or b) proposed through the mechanisms as outlined in the Governance Manual, but were ultimately not included in the P20W, along with related feasibility studies. Do any of the data element(s) being proposed overlap with the data elements in this repository? (New proposals can build on or duplicate prior proposals. It is helpful for proposers to share that context, including prior related feasibility studies.)

This indirectly overlaps with a prior proposal to track retention and persistence by major among undergraduate students.

8. Are the proposed data element(s) already collected by a state-level entity? For data element(s) already collected, please answer question nine. For data element(s) not collected, please answer question ten.

I would recommend that a feasibility study examine how to collect these elements from the UC and CSU segments, as an initial starting point.

9. [For proposed data element(s) already collected] To the extent possible, please share details pertaining to the proposed data element(s). Relevant details may include but are not limited to: a) corresponding entity that collects

and houses the data element(s); b) specific variable name(s) used in the originating data system; and c) timeframe available.

IPEDS provides a framework on collecting and disseminating data on graduate school completions by field of study (i.e., CIP). Nearly all postsecondary schools report these data. This proposal calls for analogous data to be collected on initial and ongoing enrollment at graduate school programs.

Typically, students apply and enroll at specific graduate schools and programs, allowing school data systems to track enrollment and persistence from entry to exit. This proposal asks C2C to conduct a study on how to collect these data from schools or segments so that the researchers can create graduate-level cohorts and student success rates by field of study.

10. [For proposed data element(s) not collected] Please propose institution(s) that would be most suited for the new data collection effort.

11. Please explain the desired level(s) of grain size for each data element proposed. (i.e., individual-level, institution-level, or other aggregated levels)? Multiple grain sizes may be requested for each proposed data element.

Individual-level for enrolment by school and field of study

12. Please explain the intended use case(s) for the proposed data elements (i.e., dashboards, query builder, or the research request tool)? Multiple use cases may be requested for each proposed data element.

With these new data points, a data story could tell the public the fraction of students who earn their graduation degree within six years by school and program. The story could additionally track how many students finish with their desired degree, with a strong focus on opportunity gaps by gender and race. Additionally, a time-to-degree metric on the dashboard could highlight the average number of years students take to earn their degree in each program. Such a metric would inform affordability considerations for students about the

true financial and opportunity costs of different educational pathways that go beyond four-year college.

Graduate School Data

Voting Option for Fall 2024 Proposals from the Data and Tool Advisory Board

This document provides background information to support prioritization of potential new data points for the Cradle-to-Career (C2C) Data System P20W data set. P20W data points include data uploaded directly by data providers, as well as metrics that are calculated from the information shared by data providers. P20W data points will be used to populate dashboards and a query builder tool. When the data request process becomes available, users may request other formulations of metrics that are generated from the information shared by data providers.

Staff from the Office of Cradle to Career Data (Office) worked with Advisory Board members who submitted proposals to provide greater specificity about the suggested item.

Proposal Name

Graduate School Completion Rates by Field of Study

Type of Data Point

Based on an initial review of the proposal, the Office notes that these data points could be derived based on information that is currently in the P20W data set.

The proposal includes recommendations for how to use the derived data points to create dashboard visualizations.

Data Points for the Proposed Visualization

C2C would develop a dashboard that shows progress of individuals through graduate school at CSU and UC, based on the metrics already in the P20W data set. By using the "postsecondary graduate school status" metric in combination with the other data points provided by CSU and UC, it is possible to view progress for graduate school students only. The data points that would be utilized are:

- Applied to graduate school
- Accepted at graduate school
- Institution attended
- Enrolled in college
- Term of enrollment
- Postsecondary graduate school status
- Declared four-year institution major
- Retained from fall to spring
- Returned for a second year
- Earned a postsecondary award
- Postsecondary award discipline
- Type of four-year award

Proposed Visualization

The visualization would mirror specifications for the Transfer Dashboard.

Using a cohort of people who applied to attend graduate school, the dashboard would show:

- How many were accepted
- How many enrolled
- How many were retained in their first year
- How many returned for a second year
- How many graduated in the same program that they originally applied to and enrolled in
- How many graduated within six years from any graduate program
- Median time to degree

The dashboard would allow users to disaggregate results by the following characteristics:

- Age bracket
- Gender
- Race/ethnicity
- Received financial aid

Users would also be able to see:

- Institution attended
- Academic discipline
- Type of graduate program (Master's Degree, Doctor's Degree-Research, Doctor's Degree-Professional practice, Doctor's Degree-Other, Postbaccalaureate and Post-Master's Certificate)
- Time to graduation

Cradle-to-Career Data and Tools Advisory Board Proposal Form

Instructions:
Per the Governance Manual proposal forms submitted will address significant gaps regarding whether the data system is providing access to actionable information. Please note there should only be one proposal per form.

Name:

Diana Phuong

Proposal Title: *no more than 50 characters*

Including Internship and First-Destination Survey Data in the Cradle-to-Career Data System

- By checking this box, I understand that if my recommendation is one of the top recommendations identified by the Advisory Board, I will be presenting my idea at the Fall Advisory Board meeting.

Type of Proposal[1]:

- Changes to practical tools for students *(Complete section one)*
- Adding data points not available through the P20W data set or adjusting (including the removal of) the existing P20W data points *(Complete section two)*
    - Please note a proposal form(s) can recommend adding one data point or several clearly related data points to the data system.

[1] The C2C Governance Manual notes that DTAB members can also submit recommendations related to the data request process and changes to tools such as dashboards. As the data request process and the dashboards are not yet live, those sections have been removed from the 2024 version of this form.

Section One: Changes to Practical Tools

1.What is the nature of the gap regarding access to actionable information?

2. What type of tool should be developed?

3. How would a tool address the gap?

4. Who would be the likely user(s) of the tool?

5. How does the tool relate to the mission and vision of C2C?

Section Two: Adding Data Points Not Available Through the P20W Data Set or Adjusting the Existing P20W Data Points

1.Please state the research question of interest that cannot be fully addressed with the existing data elements in the P20W Data System.

Does a student's ability to participate in career-accelerating opportunities during college impact their ability to graduate and land employment and/or graduate school?

2. How does this research question relate to the mission and vision of C2C?

This would contain data providing important insight into milestones that influence student success outcomes, allowing students and families to make the most informed decisions about their pathways and what opportunities to seek out during the college experience. This aligns with the C2C mission to expand access to data and tools that help Californians navigate the education to employment pipeline.

3. Please propose additional data element(s) needed in order to successfully address the research question of interest.

1.) Internship data (i.e. did a student have one or more internships during college)?
2.) Credit-bearing career course data (i.e. did a student participate in a career course during college)?
3.) First-destination data (i.e. what is a student's first role upon leaving their postsecondary program)?

4. Please verify that the proposed data element(s) do not already exist in the P20W Data System.

These data elements are not currently included.

5. If the P20W Data System does not currently include the proposed data element(s), can the proposed data element(s) be derived from the existing data element(s) in the P20W Data System?

While the P20W data system includes some related elements (for example, whether a student completed a career and technical education course while enrolled in K-12 and whether a student participated in an apprenticeship program) there are no similar data elements included for postsecondary data. There is currently nothing included about student internship data or credit-bearing career courses in college, and the postsecondary data elements lack any focus on career-accelerating experiences a student may have had access to during their higher education journey.

Similarly, the post-college data elements look at wages, industry of employment, and high-demand occupation status, but lack a real indicator of whether a student is landing in employment and/or graduate school within six months of graduation. Including the National Association of Colleges and Employer's (NACE) first destination survey data in the system would provide a more complete picture of students' post-college outcomes. This data set looks at each year's graduating class and determines what percent are employed full-time and/or enrolled in graduate school six months after graduation. It also breaks the data down by program/major and across all levels of postsecondary (Associate, Bachelor's, Master's, and Doctorate). The data set

also includes earnings and bonus data.

6. If the P20W Data System does not currently include the proposed data element(s), are there existing data element(s) closely related to the proposed data element(s)? If so, please list them and why they are not sufficient to answer the proposed research question.

As described above, the most similar data look at completion of a career and technical education course in K-12 or an apprenticeship program post-K-12. While these are career-accelerating experiences, they are different experiences and occur at a distinctly separate part of the education-to-employment pipeline. They do not allow consumers or researchers to fully address the above research question. Similarly, wage data does not necessarily provide a complete picture of whether a student's postsecondary experience bettered their economic outcomes or allowed them to obtain a strong first opportunity post-degree attainment. NACE's First-Destination Survey data would allow for a more complete picture of post-college outcomes.

7. C2C keeps a repository of previous data elements that were either a) considered during the planning process or b) proposed through the mechanisms as outlined in the Governance Manual, but were ultimately not included in the P20W, along with related feasibility studies. Do any of the data element(s) being proposed overlap with the data elements in this repository? (New proposals can build on or duplicate prior proposals. It is helpful for proposers to share that context, including prior related feasibility studies.)

The data elements being proposed do not overlap with data elements in the current repository or included in previous proposals.

8. Are the proposed data element(s) already collected by a state-level entity? For data element(s) already collected, please answer question nine. For data element(s) not collected, please answer question ten.

Some of these data elements are currently being collected by various institutional systems in the state. We have the listed the entities below, but

propose that the Cradle-to-Career system would eventually become the central repository for this data.

9. [For proposed data element(s) already collected] To the extent possible, please share details pertaining to the proposed data element(s). Relevant details may include but are not limited to: a) corresponding entity that collects and houses the data element(s); b) specific variable name(s) used in the originating data system; and c) timeframe available.

The first entity that collects information on internship and career outcomes are the California Community Colleges. The Chancellor's Office collects and reports information on student enrollment, progress, success, employment outcomes and earnings for all community college pathways. The information is presented on their LaunchBoard, which is hosted by Cal-PASS Plus. The source for many of these metrics is the Chancellor's Office Management Information System, but they also draw from sources such as the CSU/UC Cohort Match, the National Student Clearinghouse, and the Employment Development Department Unemployment Insurance Dataset. This data includes whether or not a student become employed, transferred to a four-year institution, has a job closely related to their field of study, and a program's median annual earnings.

At the four-year institution level, the University of California system collects first job outcomes and internship experience through their UC Undergraduate Experience Survey (UCUES). UCUES is administered to all nine UC campus and all information is self-reported by students. In 2022, the UCUES was administered to 220,000 students with 52,000 respondents after five months. Student responses can also be disaggregated by student-level characteristics such as Pell Grant status, gender, first-generation status, ethnicity, etc. Of note, the UC survey includes questions about internship and service learning, asking "have you completed or are you completing a credit bearing or non-credit bearing internship, practicum, or field experience." This data is able to be disaggregated by student characteristic as well, and has been reported out every other year since 2016 (see here).

10. [For proposed data element(s) not collected] Please propose institution(s) that would be most suited for the new data collection effort.

Outside of the current UC survey above, we were not able to locate data collection efforts on internship data for any CCs or CSUs. We were also not able to locate any existing metrics or data collection as to whether a student completed an academic, credit-bearing career course while enrolled in their postsecondary program. We would propose adding existing internship survey data into the C2C system, and studying the feasibility of CCs, CSUs, and UCs adding questions about internship completion and career course completion into their existing student surveys so that this data could be included in the state system. It is possible that some of this data is being collected but not publicly reported, which could be identified by a feasibility study.

We would also propose that NACE's First-Destination Survey (or whatever current collection system each institution uses) be included in the Cradle-to-Career Data System. 344 schools annually submit their data into this system, including private universities (Gold Gate, Claremont McKenna College, Pitzer College, Scripps College, University of San Diego, University of Southern California, Azusa Pacific University, University of La Verne, and Westmont College), four UC campuses (University of California-Berkeley; University of California- Riverside; University of California- Santa Barbara; and University of California- Santa Cruz), and four CSUs (San Jose State University, CSU- Stanislaus, CSU-Fullerton, and California Polytechnic State University- San Luis Obispo). At a minimum, the data that these eight institutions are already collecting and reporting to NACE could be included in the system.

11. Please explain the desired level(s) of grain size for each data element proposed. (i.e., individual-level, institution-level, or other aggregated levels)? Multiple grain sizes may be requested for each proposed data element.

Internship data, career course data, and first-destination survey data would ideally be collected individually through student survey but reported out in the aggregate. We would propose this data be institution level, system level, and ideally program level, with the ability to disaggregate by key student characteristics.

12. Please explain the intended use case(s) for the proposed data elements (i.e., dashboards, query builder, or the research request tool)? Multiple use cases may be requested for each proposed data element.

In an ideal world, this data would be included in dashboards, query builders, and in research request tools. This would help students/families, institutions, community organizations, and funders (public and private) make smart choices and investments into the inputs that are leading to strong outputs.

At Braven, our internal data suggests that internships are crucial to landing a strong first job post college, and that they support higher earnings. We also know that low-income and first-generation college students are far less likely to have internships in college, and particularly to have paid, career-aligned internship experiences, which impacts their ability to see a strong return-on-investment of their postsecondary degree. To demonstrate the importance of this metric, we have seen that Braven Fellows with internship experiences in undergrad show a 21 percentage point uplift in quality job attainment over students with no internship experience during college (70 percent vs 49 percent). We also see large gains in strong job attainment and earnings as a result of undergraduate academic career courses taken during a students' college coursework. Including these data elements would 1. Demonstrate whether access to these career-accelerating experiences move the needle on outcomes and 2. Incentivize institutions to prioritize these experiences for students if so, closing equity gaps over time.

Internships

Voting Option for Fall 2024 Proposals from the Data and Tool Advisory Board

This document provides background information to support prioritization of potential new data points for the Cradle-to-Career (C2C) Data System P20W data set. P20W data points include data uploaded directly by data providers, as well as metrics that are calculated from the information shared by data providers. P20W data points will be used to populate dashboards and a query builder tool. When the data request process becomes available, users may request other formulations of metrics that are generated from the information shared by data providers.

Staff from the Office of Cradle to Career Data (Office) worked with Advisory Board members who submitted proposals to provide greater specificity about the suggested item.

Proposal Name

Including Internship and First-Destination Survey Data in the Cradle-to-Career Data System

Type of Data Point

Based on an initial review of the proposal, the Office notes that these items are data points that are not currently in the P20W data set. However, there is information available on whether students secured a job and if they went to graduate school from a different data source than the one suggested.

Proposed Data Point Construction

In consultation with the Advisory Board member who proposed the data point, the Office operationalized the proposal to provide information on career opportunities by creating 3 metrics.

1) **Postsecondary student who participated in an internship** - Student had an internship while enrolled at a postsecondary institution
   Example: A student at College of San Mateo participated in a paid internship at Stanford to conduct research on nanotechnology applications. This student would be counted as participating in an internship.

2) **Postsecondary student who took a career course** - Student took a credit-bearing course in a technical discipline or a credit-bearing course in career preparation while enrolled at a postsecondary institution
   Example: A student at Sacramento State took a course in Project Management. The student would be counted as having taken a career course.

3) **Postsecondary student who entered employment or graduate school within six months of graduation** - Students who responded to the National Association of Colleges and Employers (NACE) First Destination survey and reported they were either employed or enrolled in graduate school six months after completing a bachelor's degree.
   Example: A student graduated from CSU Fullerton with a business major in May and got a job as a restaurant manager in July. This student did not respond to the First Destination survey. The student would not be counted as having entered employment or graduate school within six months.

Cradle-to-Career Data and Tools Advisory Board Proposal Form

Instructions:

Per the Governance Manual proposal forms submitted will address significant gaps regarding whether the data system is providing access to actionable information. Please note there should only be one proposal per form.

Name:

Diana Phuong

Proposal Title: *no more than 50 characters*

Track Child Savings Accounts Data

- By checking this box, I understand that if my recommendation is one of the top recommendations identified by the Advisory Board, I will be presenting my idea at the Fall Advisory Board meeting.

Type of Proposal[1]:

- Changes to practical tools for students *(Complete section one)*
- Adding data points not available through the P20W data set or adjusting (including the removal of) the existing P20W data points *(Complete section two)*

[1] The C2C Governance Manual notes that DTAB members can also submit recommendations related to the data request process and changes to tools such as dashboards. As the data request process and the dashboards are not yet live, those sections have been removed from the 2024 version of this form.

- Please note a proposal form(s) can recommend adding one data point or several clearly related data points to the data system.

Section One: Changes to Practical Tools

1.What is the nature of the gap regarding access to actionable information?

2. What type of tool should be developed?

3. How would a tool address the gap?

4. Who would be the likely user(s) of the tool?

5. How does the tool relate to the mission and vision of C2C?

Section Two: Adding Data Points Not Available Through the P20W Data Set or Adjusting the Existing P20W Data Points

1.Please state the research question of interest that cannot be fully addressed with the existing data elements in the P20W Data System.

This proposal aims to include child savings account (CSA) data in C2C, data that is currently not included. The objective is to ensure we can track the utilization of CSAs across California, starting with the state's CSA program, CalKIDS.

CalKIDS is a state-wide program operated by the ScholarShare Investment Board (SIB) within the Office of the State Treasurer which establishes and funds

college savings accounts (CSAs) for eligible California children. This data is centrally collected and managed by SIB which has a small team of less than five staff to administer the statewide program across 58 counties.

- As of March 31, 2024, over 4.3 million California children & youth have a CalKIDS account; and California has invested more than $2 billion in the CalKIDS program. The annual allocation to fund CalKIDS accounts in the FY23-24 state budget is $185M, this does not include the cost to administer or promote the program.
- The State Treasurer's Office automatically enrolls eligible children into CalKIDS, but families must "register" their child's CalKIDS account through an online portal to see account balances and request disbursement of the funds for post secondary expenses.
- CalKIDS eligibility includes:
    - all babies born in California on or after July 1, 2022 ($25 or $100 if born on or after July 1, 2023 to establish the account and up to $75 in participation incentives);
    - all 1st through 12th grade low-income public school students from the 2021-22 academic year ($500 to seed their accounts and an additional contribution for homeless students ($500) and students in the foster care system ($500); and
    - all eligible low-income 1st grade public school students from school year 2022-23 forward (contingent on ongoing state funding).
    - Ongoing outreach and engagement activities throughout the state and in partnership with local communities raises awareness of and engagement with this resource for post-secondary education.

The objective of this proposal is to close a gap in actionable information around 1) the utilization of the CalKIDS program, and 2) its impact on college access and affordability. The following are examples of questions that stakeholders hope this data element will answer which currently no other data element can answer:

- *Which California Students have a CalKIDS account - and what other*

resources are CalKIDS account holders also receiving (per data in C2C)?

- *What percentage of eligible students registered for CalKIDS accounts? Are registration targets being met?*
- *Which schools and communities need additional support to increase awareness & engagement?*
- *Are there any patterns of non-engagement (by language group or geography) that may point to needed shifts in CalKIDS program operations or outreach?*
- *What percentage of registered CalKIDS accounts were linked to a ScholarShare 529 plan?*
- *Does this data show any correlation between CalKIDS account engagement and student's academic achievement in schools or communities throughout the state?*
- *What percentage of CalKIDS accounts have been accessed to pay for college?*
- *Do we see any correlation between use of CalKIDS funds to pay for college expenses and college persistence?*
- *How much CalKIDS money (by county, city) is disbursed to pay for college expenses each year?*

2. How does this research question relate to the mission and vision of C2C?

Including CalKIDS utilization and impact data in the Cradle to Career Data System will help ensure better educational outcomes for our students.

- Access to CalKIDS data will raise awareness of this resource to pay for post-secondary education among enrolled **students and families**.
- **Educators** (including School District leaders, principals, K-12 teachers, early education teachers, and other school staff) will be better positioned to promote awareness of CalKIDS among their families and provide input to CalKIDS on effective outreach and implementation appropriate to their communities. For example, they can enhance their enrollment and registration procedures to include a step where students who have not claimed can get immediate support claiming/registering their account. "Student level data" is a very frequent request of school principals and superintendents. They want to do outreach, but do not

want to blanket parents/students that don't have accounts with information because it has the potential to build false hope. If the C2C system has a level of access that is just for schools, then student-level data would be very helpful. We recognize that this cannot be available to the general community. Last, it would be helpful to gather/report claim and disbursement rates by colleges or trade programs to answer the question "*How many students at Evergreen Community College have a CalKIDS account, but have not actually requested a funds distribution?*" There are large numbers of students that are currently in college that haven't accessed their funds but right now we know this anecdotally, we need the data to understand how large this group is.

- **Policymakers, Researchers & Advocates** will have access to the data necessary to influence effective implementation of CalKIDS, shape related public policy, facilitate the integration of this new program with other college readiness and access initiatives and ensure CalKIDS positively impacts on the goal of educational equity. For example, communities with low CalKIDS registration can use this information to advocate for resources to improve outreach and increase signups. Additionally, communities who are exploring launching local children's savings account programs can use information about who has and has not registered their CalKIDS account to inform the design and implementation.

The inclusion of CalKIDS data in the C2C system is one step towards accountability for this statewide program. CalKIDS data available in dashboards, the query builder, and research requests can be used to increase awareness of and engagement with this significant state investment in building an expectation of post-secondary education into our families and school system from birth onward.

As described on the C2C website and various publications, "the C2C Data system seeks to foster evidence-based decision-making to help Californians build more equitable futures and empower individuals to reach their full potential. The C2C data system will provide tools to help students reach their goals and deliver information on education and workforce outcomes….It will provide insights into critical milestones in the pipeline from early care to K–12 to higher education, skills training, and employment."

CalKIDS is a new and substantial state investment in our children's post-secondary success and economic mobility. As a cradle to career platform, it is critical that C2C expand its data elements to include resources provided to youth at birth and during their K-12 schooling, and to measure its impact on improving college access, affordability, persistence, and attainment, especially for those that are under-resourced and underserved.

A report by the National College Attainment Network (NCAN) found that CSA enrollment tripled nationally in 2022, thus there is momentum that we can capitalize on here in California [1]. "CalKIDS is the largest program and showcases how critical state funding is to ensure enrollment at scale, enrolling 3.4 million children across California." What we are able to accomplish here sets an example for the rest of the country. What we do to grow utilization now in California will impact the trajectories of millions of students, many of whom are still grappling with the return on investment of a college education.

Lastly, as we expand the Postsecondary Financial Aid Experiences data to include CSAs, we must be sure to disaggregate where possible. Echoing the recommendation of California Competes, "C2C should present data for small geographic service areas, such as neighborhoods, cities, K–12 school districts, and community college districts, to the greatest extent possible" [2].

Sources:
[1] Report Finds CSAs Tripled Nationally in 2022, August 2023, www.ncan.org/news/news.asp?id=649219
[2] People-Powered Data: Designing a Data System for and with All Californians, August 2023, https://californiacompetes.org/wp-content/uploads/2023/08/People-Powered-Data-C2C-Brief-FINAL.pdf

3. Please propose additional data element(s) needed in order to successfully address the research question of interest.

This proposal also recommends that C2C provide the following data points **once available**:

- CalKIDS data disaggregated by city, zip code, and county
- Everything currently available in the SIB Data Dictionary (CDE + Public Health)
    - Student & Parent demographics
        - Can't uplift equity issues otherwise
        - Can't demonstrate asset building success otherwise
    - Unhoused
    - Foster care
    - Paternal/maternal education levels for newborns
    - Visibility — not for the public to see but for relevant stakeholders (e.g. researchers, CalKIDS Institute)
- Demographic data for newborns which I understand is imperfect because race is based on the parent giving birth
- Individual level data instead of the SSID for those not yet enrolled in school (de identified but with an identifier to follow them)
- When the student was enrolled/funds set aside by SIB including which Academic year and what Grade the student was in
- When the student registered/claimed their account and each time they sign in to their portal including the Month and Year
- Total number of CalKIDS accounts established since there is an automatic setup step by the state for eligible students
- If the student connected to a ScholarShare 529 account and when
- The value of people's CalKIDS accounts and, where applicable, the linked ScholarShare 529 accounts
- Disaggregated by City

The following are data points that will need to be collected in the future:

- Number and demographics of **Eligible Accounts versus Accounts Claimed/Unclaimed in locally administered CSA programs** (see the list of programs at https://norcalpromisecoalition.org/ca-csa-coalition/).
- Total amount of **Funds Dedicated versus Funds Claimed/Unclaimed in locally administered CSA programs**.
- Number and demographics of children who have **multiple asset building accounts** (CalKIDS, local CSA, ScholarShare 529, HOPE, etc.).
- **Statewide Student Identifier (SSID)** to track student-level data.

4. Please verify that the proposed data element(s) do not already exist in the P20W Data System.

I have checked and there are no data points related to "saving" in the current data elements listed as of July 24, 2024.

5. If the P20W Data System does not currently include the proposed data element(s), can the proposed data element(s) be derived from the existing data element(s) in the P20W Data System?

I have had the privilege to meet with leaders at SIB specifically around this proposal, including the former Executive Director Julio Martinez and CSA Initiatives Manager Noah Lightman. They have confirmed that much of the data that they have is already provided by CDE to C2C (e.g. demographics, SSID). However, specific data around CalKIDS accounts, claims, and funds are not available in C2C.

6. If the P20W Data System does not currently include the proposed data element(s), are there existing data element(s) closely related to the proposed data element(s)? If so, please list them and why they are not sufficient to answer the proposed research question.

There are none, this is specific to child savings accounts and the **Postsecondary Financial Aid Experiences** are all based on activities that do not happen until later in a student's educational journey, such as high school, when they are college-bound. The CSA data is much earlier than that, starting as early as birth.

7. C2C keeps a repository of previous data elements that were either a) considered during the planning process or b) proposed through the mechanisms as outlined in the Governance Manual, but were ultimately not included in the P20W, along with related feasibility studies. Do any of the data element(s) being proposed overlap with the data elements in this repository? (New proposals can build on or duplicate prior proposals. It is helpful for proposers to share that context, including prior related feasibility studies.)

There are no previously submitted data elements in the repository that align or overlap with this request as of July 24, 2024.

8. Are the proposed data element(s) already collected by a state-level entity? For data element(s) already collected, please answer question nine. For data element(s) not collected, please answer question ten.

Yes, the SIB is housed under the CA State Treasurer's Office. The team can be found at https://www.treasurer.ca.gov/scholarshare/contacts.asp.

9. [For proposed data element(s) already collected] To the extent possible, please share details pertaining to the proposed data element(s). Relevant details may include but are not limited to: a) corresponding entity that collects and houses the data element(s); b) specific variable name(s) used in the originating data system; and c) timeframe available.

Data would be collected from the California State Treasurer's Office through SIB, which manages the CalKIDS program.

As of March 31, 2024, available data includes:

- County
- Funds dedicated
- Funds unclaimed
- Eligible accounts
- Already claimed

From my knowledge, data is available from October 2022 to the present.

The feasibility study should help determine whether data can pass through SIB directly to C2C or through another partner, such as CSAC. It would also need to determine what support is needed to collect more than the data currently shared publicly. If not, additional capacity and investment is likely needed to:

1. Build CalKIDS reports with an enhanced data dictionary,
2. Run those reports at a determined frequency (ideally monthly, pie-in-the-sky is daily),

3. Share that data with C2C, either directly or through another entity

10. [For proposed data element(s) not collected] Please propose institution(s) that would be most suited for the new data collection effort.

N/A

11. Please explain the desired level(s) of grain size for each data element proposed. (i.e., individual-level, institution-level, or other aggregated levels)? Multiple grain sizes may be requested for each proposed data element.

Publicly published data is currently only available at the county level. More granular, district-level has been accessed by outreach grantee partners. I recommend that publicly accessible grain sizes be considered for:

1. **School district** - to support them in launching district-wide, multi-school campaigns to increase CalKIDS claims and help families use the funds.
2. **School** - to support school-wide campaigns.
3. **City** - to support city-level campaigns that can be supported by departments that have high-touch points with families such as public libraries, parks, and social services.
4. **Student** - to facilitate targeted outreach and promotion activities by teachers, schools, and community partners to ensure that the students are aware of and taking full advantage of their CalKIDS account.

There is broad interest in CalKIDS data. For researchers, community advocates and policy-makers having access to CalKIDS data within the C2C system will enable a level of impact analysis that would otherwise not be possible.

The CalKIDS program is the largest "child savings account" initiative in the nation. Awareness of and utilization of CalKIDS by enrolled children and youth is foundational to its success. Research suggests that CSAs can have beneficial effects for students and families across the life course, including improved early child socioemotional development, child health, maternal mental health, educational expectations, and academic performance. Having savings set aside for education is also associated with improved postsecondary educational enrollment and completion. Many of these benefits are strongest for children from low-income families. In order for these benefits to manifest, students and families must be aware of the program. Including CalKIDS data in the C2C will foster more transparency and

enable more collaborations and direct outreach to dramatically increase families' awareness of CalKIDS. California is well positioned to build on the success of CalKIDS to advance a national CSA policy. Further, utilization data can be used to advocate for increased investments in CalKIDS in the future.

For this reason, there are multi-sector stakeholders that would benefit from this tool at the local, regional, and state levels, including those who work directly with families and those who train/support practitioners.

Braven Bay Area is a member of the Northern California College Promise Coalition (NCCPC). NCCPC is a network of nearly 70 college promises, college access, workforce, and other programs and other organizations who are committed to supporting CalKIDS outreach in their communities. This includes cities, colleges, K-12 and charter districts, and more. NCCPC is also the backbone function for the CA Child Savings Account Coalition. The CA CSA Coalition is a network of 15 locally managed CSA programs such as Brilliant Baby out of Oakland Promise in Alameda County, and College in My Future by Excite Credit Union in Santa Clara County. Many of these programs are also grantees of the outreach funds distributed by the Treasurer's Office to support promotion of CalKIDS.

Braven Bay Area has a strong presence in the City of San Jose through our partnership with San Jose State University. San Jose has a locally managed CSA program. Per **John Hogan** at Excite Credit Union who manages the local program College In My Future and is a CalKIDS outreach grantee, "When we have real-time visibility to school level registration data, we are able to make specific actions plans to increase registrations and/or validate that recent activities have been successful (or not successful). This is very much appreciated by school and community leaders."

As you can see, when these programs have access to data, they can do powerful things with that information to activate the community into action. As an example, the following chart was developed by Excite Credit Union and shared with K-12 and community leaders across the City of San José. This data enticed the local community to share information about CalKIDS to families who attend the schools and districts listed, quoting the registration rates for a stronger, data-informed call to action to increase claims across the city.

As of October 16, 2023 -- Selected High School districts/groups			
	Students with CalKIDS accounts		
	Total	**Registered**	
Campbell Union High	3360	253	7.5%
East Side Union High	10400	780	7.5%
Fremont Union High	1715	149	8.7%
Gilroy & Christopher	1802	91	5.0%
Los Gatos Saratoga HSD	231	21	9.1%
ACE Charter High	372	20	5.4%
Alpha Cindy Avitia High	442	40	9.0%
Latino College Preparatory Aca	369	40	10.8%
KIPP (2 San Jose High Schools)	655	44	6.7%

Important Note: 2022 and 2023 high school graduates also have CalKIDS accounts (if low income).

All stakeholders in California can benefit in some way from the collection and reporting of CSA data. This information can help inform the utilization of CalKIDS, support more targeted outreach across the state in areas where there is low utilization, and ensure equitable access to and benefit from the program. After interviewing several local CSA programs who received funding from the Treasurer's Office to support outreach, we learned that some data needed to inform outreach and follow-up to communities is available upon request. Due to limited capacity and staffing, data at different grain sizes (such as city) and at more regular frequencies (like monthly), are under exploration with the Treasurer's Office.

12. Please explain the intended use case(s) for the proposed data elements (i.e., dashboards, query builder, or the research request tool)? Multiple use cases may be requested for each proposed data element.

Currently, there are no tools with data points in the C2C data system to track information about saving for college, including through programs like CalKIDS. The tools that should be developed include:

- Several new DATA POINTS under the category **Postsecondary Financial Aid Experiences**.
    - Currently, the data points listed do not include Children's Savings Account (CSA) programs, which are typically long-range investments starting at a child's birth or as they enter elementary school.
    - The California Kids Investment and Development Savings Program (CalKIDS) is a CSA program, administered by the ScholarShare Investment Board (SIB), an agency of the California State Treasurer's

Office. CalKIDS sets aside money, including seed deposits and other potential incentives, to help students pay for qualified higher education expenses. To learn more about CalKIDS, visit CalKIDS.org.
    - CalKIDS participants may also establish individual accounts with ScholarShare 529, California's 529 College Savings Plan. This program is managed by the State Treasurer's Office.
- A new or enhanced DASHBOARD under the category **Financial Aid**.
    - I explain more in question #12.

Given the tremendous financial investments of the state to fuel access to CSAs, it is imperative that C2C develop both the ability to track CalKIDS utilization statewide, and produce the information in an intuitive dashboard for stakeholders.

Dashboards

- The **Goals** would include:
    - For Students and Families: clarify how financial aid enables students to enroll in and complete postsecondary education, including the impact of asset building investments like CalKIDS (note that HOPE is another type of account).
    - For Faculty and Counselors: clarify whether students are securing financial aid like CalKIDS and ScholarShare 529 College Savings Plans, and how this relates to postsecondary education applications, enrollment, and success.
    - For Administrators and Policymakers: specify how access to sufficient financial aid, such as asset building investments like CSAs which can begin as early as birth, relates to postsecondary access and success to inform budget decisions.
    - For Advocates and Researchers: clarify how financial aid, such as asset building investments like CSAs which can begin as early as birth, relates to postsecondary applications, enrollment, and success, and if all students are benefiting from aid.
- The **Dashboard** would include:
    - Infographic showing proportion of students who receive state aid like CalKIDS and ScholarShare 529 College Savings Plans who apply to and enroll in college, compared to similar students who don't receive aid, plus a separate graph showing the demographics of families who claim CSA accounts compared to those who do not, plus a separate graph showing the total CSA funds students have access to when they link to a ScholarShare529 plan compared to those who do not.
    - Disaggregation by a) year, b) county, c) student characteristics

(including race/ethnicity, gender, age bracket, parental education level, military status, foster status, and homelessness status), d) link to 529 plan, plus one of the following 3) school district, school, and/or city.
    - Export with the ability to provide summary files or charts based on the variables selected.

Query builder

- Local, regional, and state leaders can run specific queries leveraging information about CalKIDS to better understand how student outcomes are impacted by asset building accounts managed by the state. They can research answers to questions like those listed in Section One question #3 and Section Two question #1. For example, this data would be valuable to the 15 locally managed CSA programs supporting their families in claiming both local and statewide accounts.

Research request tool

- Researchers can conduct complex analysis and evaluate the CalKIDS program for their communities and the state. For example, this data would be valuable to the UCLA CalKIDS Institute.

Currently, the SIB holds and controls CalKIDS data. This data is not available in the public realm. All CalKIDS information requests (from school districts, school leaders, community organizations, researchers, etc.), must be made to the SIB. SIB has the discretion to provide CalKIDS data (or not) as frequently (or infrequently) as its staff have the bandwidth to share it.

Those who are leading local grassroots campaign to increase CalKIDS claims are limited to data from the SIB that is:

- At the county level only for non-grantee requests
    - School level data available for outreach grantees like Excite Credit Union
- Provides the following *aggregate* data points:
    - Funds dedicated
    - Funds unclaimed
    - Number of eligible children who have CalKIDS accounts
    - Number of eligible children (or their parents) who have gone online to "claim" their CalKIDS accounts

By adding CalKIDS data to C2C and making it available at scale through dashboards

and data queries, local, regional, and statewide practitioners and CalKIDS champions can leverage the following data to answer critical questions:

1. Optimally, CalKIDS data in the C2C system would be linked to the specific eligible students. This will enable parents, teachers, school leaders, and researchers to know if a student has funds set aside by CalKIDS, alongside the other public resources and data embedded in C2C.
2. Data by county* - answers the question, *How many families are impacted in this area out of the total number of residents?*
3. The number of ELIGIBLE accounts* - answers the question, *How many families in this area have funds set aside by CalKIDS and have not yet registered?*
4. The number of CLAIMED (registered)* accounts - answers the question, *How many families in this area have already registered accounts?*
5. The number of high school graduates who have received a DISBURSEMENT from the CalKIDS account to pay for post secondary education expenses.
6. Data by school or school district - answers the question, *How many students have accounts so that our educators can support outreach and CalKIDS claims?*
7. The number of UNCLAIMED accounts - answers the question, *How many families in my county have not yet registered accounts? Where are they in the process of registering?*

a. Currently, we can infer this number by subtracting the ALREADY CLAIMED number from the ELIGIBLE ACCOUNTS number. However, this does not take into consideration other possible status/benchmarks for a family going through the process of claiming an account.

a. Timeline for families - There are three possible pathways for students and families to claim their accounts:
    - **#1 At Birth** - families need the Local Registration Number (LRN) that appears on the child's birth certificate OR the unique CalKIDS ID found on a mailed letter.
    - **#2 While enrolled in a public school for 1st-12th grade** - students need the Statewide Student Identifier (SSID) provided by the school or school district OR the unique CalKIDS ID found on a mailed letter.
    - **#3 After graduating from high school** (funds can be used starting at age 17 and must be spent by age 26) - students need the SSID.

a. There are a few steps that students and families need to take, regardless of which of the three pathways they take:
    - Confirm Eligibility using the **Eligibility Tool**
    - **Claim Your Account**

a. For each of these steps, there are at least a few possible benchmarks for

families besides 1) having funds set aside for them by SIB, and 2) claiming those funds:

1. Land on the appropriate webpage
    1. Eligibility Tool: https://calkids.org/get-started/?form=eligibility-tool#eligibility-tool
    2. Claim Your Account: https://calkids.outcometracker.app/p/CalKids/calkids/registration
1. Respond to the first set of questions:
    1. Eligibility Tool: *Which are you asking about? (Newborn, student, student in 1st grade or younger)*
    2. Claim Your Account: *County, birthdate, Registration code (LRN, SSID, or CalKIDS ID*
1. Respond to the next set of questions, which depends on the first set of responses.

* Data already available and reported by ScholarShare Investment Board since October 2023.

Last, we recommend thoughtful consideration of the order in which data is displayed to users and. The order of data could influence how the data serves the user's intentions.

College Savings Accounts

Voting Option for Fall 2024 Proposals from the Data and Tool Advisory Board

This document provides background information to support prioritization of potential new data points for the Cradle-to-Career (C2C) Data System P20W data set. P20W data points include data uploaded directly by data providers, as well as metrics that are calculated from the information shared by data providers. P20W data points will be used to populate dashboards and a query builder tool. When the data request process becomes available, users may request other formulations of metrics that are generated from the information shared by data providers.

Staff from the Office of Cradle to Career Data (Office) worked with Advisory Board members who submitted proposals to provide greater specificity about the suggested item.

Proposal Name

Track Child Savings Accounts Data

Type of Data Point

Based on an initial review of the proposal, the Office notes that these items are data points that are not currently in the P20W data set.

Proposed Data Point Construction

In consultation with the Advisory Board member who proposed the data point, the Office operationalized the proposal to provide information on college savings accounts by creating 3 metrics. This information could be linked to existing P20W data points to answer research questions about who is participating in college savings programs and how that relates to educational outcomes.

1) **Student who is eligible for a CalKIDS college savings account** - Student is eligible to participate in the California Kids Investment and Development Savings Program
   Example: A person was born in Redding on June 5, 2005. He would not be eligible because children must have been born after July 1, 2022 or be a current low-income K-12 student to participate in the program.

2) **Student with an active CalKIDS college savings account** - Student has an active account in the California Kids Investment and Development Savings Program
   Example: A child was born in Riverside on September 18, 2023. Her parents activated her account but did not put additional funds into the account. The child would be shown as having an active CalKIDS account.

3) **Student with an active ScholarShare 529 college savings account** - Student has an active account with California's ScholarShare 529 program
   Example: A child was born in Mammoth Lakes on September 18, 2023. Her parents created a ScholarShare 529 account after being prompted as part of claiming her CalKIDS savings account. The child would be shown as having an active ScholarShare 529 account.