

# California Cradle-to-Career Workgroup Meeting Summary

December 1, 2020

The California Cradle-to-Career Data System Workgroup, which is comprised of partner entities named in the authorizing legislation, provides recommendations to the Governor's Office regarding data system development.

This document provides a summary of the key points that emerged from substantive discussion over the course of the December 1, 2020 Workgroup meeting. More information about the meeting, including support materials, a recording of the meeting, and the PowerPoint, are available at <https://cadatasystem.wested.org/meeting-information/Workgroup> (click on "Meeting Materials").

The following Workgroup representatives attended the meeting:

Thomas Vu, Association of Independent California Colleges and Universities; Barney Gomez, California Community College Chancellor's Office; Mary Nicely, Sarah Neville-Morgan, and Cindy Kazanis, California Department of Education; Brenda Bridges Cruz, California Department of Technology; Akhtar Khan (for Natasha Nicolai) California Department of Social Services; Elaine Scordakis, California Health and Human Services Agency; Amy Fong, California School Information Services; Ed Sullivan, California State University; Patrick Perry, California Student Aid Commission; Michele Perrault, Commission on Teacher Credentialing; Muhammad Akhtar (for Amy Faulkner), Employment Development Department; Joy Bonaguro, GovOps; Jeanne Wolfe, Labor and Workforce Development Agency; Sara Pietrowski, State Board of Education; Chris Furgiuele, University of California Office of the President

## Public Comment

Anna Alvarado of the EDGE Coalition praised the inclusion of apprenticeship data, given the governor's goal of implementing 500,000 apprenticeships by 2029. She also expressed support for adding a seat to the governing board that will foster the inclusion of workforce training in the data system. In order to recover from the pandemic, California will need to support adult learners to retrain and upskill.

Samantha Tran of Children Now noted her organization's support for the proposed scope, scale, and vision for the data system. She observed that a longitudinal data system is long overdue and will provide strong returns for a small investment. She shared her appreciation for the planning process and urged workgroup members to continue to ensure the data system will provide information that is accessible and actionable.

Liz Guillen of Public Advocates stressed the importance of including the proposed operational tools in the first phase because they will ensure the data system will be used by students, families, and educators. The way to build ongoing support will be to demonstrate the value of the system, particularly for promoting equity and closing achievement gaps. She echoed earlier comments in appreciation of the planning process and inclusion of apprenticeship data.

Angela Perry of The Institute for College Access and Success commended the inclusion of data on financial aid and private postsecondary institutions. Given the pandemic, it is critical that data be provided that allows the state to support the neediest students. She suggested that information on private postsecondary be included earlier in the timeline, given that many students are electing to enroll in private institutions, the history of predatory behavior by some private colleges, and the federal government deregulation of this sector.

## Pilot Project

Kathy Booth of WestEd described the opportunity for partner entities to propose pilot projects to test the legal and technical structure that has been proposed for the data system. Joy Bonaguro of GovOps clarified that projects will be implemented in the spring and will provide valuable information for the procurement process. The technological implementation will be managed by GovOps, using information that has been tagged and loaded by the partner entities. If the test is successful, the pilot will also result in a publicly available dashboard that will provide valuable intersegmental information to the public. The pilot would be funded by accessing remaining dollars associated with the legislation that authorized the planning process.

## Timeline and Priorities

Kathy Booth of WestEd clarified how the timeline was developed and then the workgroup considered the annual priorities for each of the first five years of the data system.

## Overall Comments

Chris Furgiuele of UC suggested that work begin in the first half of 2021 to engage agency leadership and implement the governance structure. The Governor's Office should help to lead this engagement due to the formal protocols that exist at some of the partner entities.

Ed Sullivan of CSU noted that the timeline will be impacted by how quickly work begins in the next fiscal year, particularly related to purchasing the technology necessary to implement the analytical tools.

Cindy Kazanis of CDE emphasized the importance of establishing a budget line item for National Student Clearinghouse data. Currently the segments are purchasing individual licenses. For example, CDE spent more than \$380,000 in the last fiscal year to secure this information.

## Evaluation

Workgroup members noted that more attention should be paid to the evaluation process in the timeline. For example, there should be a specific item for the first year related to developing an annual evaluation plan.

Joy Bonaguro of GovOps noted that in addition to independent evaluations, the managing entity should be setting up feedback mechanisms that support ongoing evaluation and improvement efforts. For example, use of digital services or the data request process could be analyzed in accordance with ongoing performance metrics such as how many people are accessing the tools and what aspects they are using the most. The user centered design process could also provide early feedback and identify key marker to track over time. Customer satisfaction surveys could examine the accessibility and usability of the tools. Finally, the governing board should examine how the data gets used, such as if they inform policy decisions. She shared an example of an evaluation plan and underlying logic model that was implemented for technology tools in San Francisco: <https://datasf.org/resources/open-data-metrics/>

Akhtar Khan of CDSS noted that his agency has examples of evaluation processes used for the Open Data Portal dashboards that could support annual reviews, leading up to a more formal evaluation of the Cradle-to-Career data system at the end of phase one.

Amy Fong of CSIS supported the idea of establishing an evaluation strategy early and recommended that the technology and the program/services pieces be examined separately.

Ed Sullivan of CSU felt that the evaluation should be linked to some form of accountability, such as an annual report to the legislature that highlights what is working, where there is room for improvement, and opportunities for action. An ongoing process will ensure there are no surprises at the five-year evaluation mark. He suggested that annual evaluations should be completed in the fall, so that information can be provided to the legislature in advance of the annual budgeting process.

### Synthetic Data

Kathy Booth of WestEd explained that a synthetic data set would provide information that mirrors the patterns found in individual-level data, which would allow for more sophisticated research projects, but would not provide information on actual students. Baron Rodriguez of WestEd noted that Maryland and Texas have developed synthetic data sets to allow for research projects that the data providers felt would not be possible without violating FERPA. However, synthetic data is expensive to develop and maintain, and is a relatively new practice.

Chris Furgiuele of UC stated that synthetic data should only be accessed in the secure data enclave and research requests should still be subject to a review process. Therefore, he saw the development of the secure data enclave and the synthetic data set as linked. Ed Sullivan agreed that the synthetic data set should be kept in the secure data enclave.

Amy Fong of CSIS questioned whether the secure data enclave and synthetic data set decision should be considered separately. Baron Rodriguez of WestEd affirmed given examples from other states, the two do not need to be linked. Given this perspective, the workgroup discussed the appropriate timeframe for evaluating the creation of a synthetic data set.

Amy Fong raised a concern about the ongoing costs of maintaining a synthetic data set, and Cindy Kazanis concurred.

Sara Pietrowski of SBE thought the initial development of the secure enclave should be undertaken first, and that the synthetic data set should be addressed later.

The group agreed that the evaluation of the synthetic data set should be moved up one year (to year three), at which point it should be examined in the context of demand for information outside of the query builder and data request process. If synthetic data proves both valuable and feasible, it should be built in the fourth year.

### **VOTE**

The workgroup unanimously approved the timeline.

### Fiscal Impact

Given the lack of clarity about specific technical requirements for the data system, the workgroup was not able to develop a comprehensive budget in time for the legislative report deadline. Therefore, the workgroup developed ballpark costs for 2021-22, based on budget categories, known figures for the operational tools, and costs from states with similar systems.

### Cost Drivers

The workgroup discussed the impact of scale on the budget, both in terms of the number of people in California and the proposed number of records in the data system. After reviewing comparisons to other states, Patrick Perry of CSAC noted that scale has a minimal impact on cloud storage costs. Instead, costs

are more likely to vary based on the number of agencies that the managing entity will be supporting and the number of technology linkages that will be managed. Amy Fong of CSIS concurred and noted that data quality is also directly related to staffing.

Cindy Kazanis of CDE noted that her agency partners with CSIS to provide support for data collection and quality management. It takes more than 50 people to work with school districts and charter schools, plus over 100 people at CDE.

Kathy Booth of WestEd reminded the group that, with the exception of the independent colleges, information would be provided by state agencies and therefore will be a less complex task. The managing entity will be working with data that has already gone through this extensive data collection, review, and certification process.

Kathy Booth also noted that the Cradle-to-Career data system will require upgrades to underlying agency data systems, particularly for CDE and CCCCCO related to the implementation of the operational tools. Other states have received significant federal funding to modernize their data systems, so they could provide more real-time and flexible data access.

Patrick Perry of CSAC recommended pursuing federal funding to support development of the data system. Baron Rodriguez of WestEd shared the website that describes federal funding opportunities: <https://nces.ed.gov/programs/slds/stateinfo.asp>

Barney Gomez of CCCCCO wondered if the federal government will continue to support disconnected state systems, or if they will pursue the development of a single, national data system.

### *Budget Categories*

The group identified the following key budget categories, and considerations for each:

#### *Staffing*

Patrick Perry of CSAC noted that decisions about the technical infrastructure—such as whether information is hosted inhouse or outsourced—will have a significant impact on staffing costs. Barney Gomez concurred.

This insight was born out by staffing figures for Kentucky—which has 24 staff for a home-grown system—compared to Minnesota, which has 13 staff for an outsourced system and relies on in-kind staffing from data providers.

Cindy Kazanis of CDE noted that there should be sufficient staff to support the governance structure as well. To provide a comparison point, she noted that the State Board of Education has 12 staff.

Tom Vu of AICCU indicated that there should be staffing at the managing entity dedicated to supporting the data providers and the operational tools.

Chris Furguele of UC agreed, indicating that it would be important for the managing entity to support activities such as following-up on deadlines and troubleshooting tasks like data submissions. This will include developing strong documentation and providing training when there is staff turnover. He worried that funding levels shown for other state agencies as a comparison (\$2-2.5 million) would not be sufficient.

### *Operations*

Amy Fong of CSIS suggested greater specificity for this category, such as office space and furniture. She also wondered whether the pandemic could lead to more virtual offices, which would reduce operational costs. Barney Gomez supported the idea of using a virtual office to keep costs down.

Elaine Scordakis of CHHS noted that the state has standardized list of expenses for overhead that could be used to estimate costs.

Chris Furguele of UC noted that community engagement outreach expenses should be included in operations, such as materials, conferences, and meeting costs.

Brenda Bridges Cruz of CDT suggested that operations costs include licenses for new tools and the standard Microsoft Office Suite for the staff.

### *Technical infrastructure*

Chris Furguele of UC reflected on the variation in the tools that have been produced by other states. He urged the group to prioritize high quality public-facing resources, as Kentucky has done.

### *Support to data providers*

Elaine Scordakis of CHHS, Akhtar Khan of CDSS, and Ed Sullivan of CSU emphasized the importance of underwriting the work that will be done by data providers to populate the data system.

Ed Sullivan of CSU estimated that the Chancellor's Office will need an additional 1 FTE, first to load data into the system and norm the data, and then to provide ongoing support, such as reviewing and finalizing data requests. However, he expressed concern about the ability to hire staff in the next year, particularly if budgets constrict due to the economic downturn.

Barney Gomez of CCC worried that one position would not be sufficient and estimated that the volume of community college data would require 2 FTE.

Chris Furguele of UC noted that it is difficult to estimate the workload that the data request process will generate.

Tom Vu of AICCU wondered where those staff would be allocated, particularly for independent colleges where information will be provided by individual institutions. He recommended that when funding is provided to data providers, there should be flexibility for how those dollars would be spent.

### *Additional expenses*

Elaine Scordakis of CHHS noted there will need to be support for increased workload for the Committee for the Protection of Human Subjects related to serving as the Institutional Review Board for unitary data requests.

Cindy Kazanis of CDE flagged the need to cover subscriptions and licensing for products such as the National Student Clearinghouse information.

### *Governor's Office Perspective*

Ben Chida of the Governor's Office echoed the reality that it is not possible to develop a detailed budget request at this time and reminded the group that figures can be refined over the spring as part of the normal budgeting process.

Sarah Neville-Morgan of CDE suggested that it would be helpful to align data infrastructure needs identified in the newly released Master Plan for Early Learning and Care with the Cradle-to-Career recommendations. Ben Chida agreed that the master plan recommendations should be taken into account in coming years.

In response to a question from Tom Vu of AICCU about whether it would be appropriate for the workgroup to keep its recommendations broad and aspirational, Ben Chida reflected that the process to date has been very helpful. The workgroup defined the “why” for the system and built a coherent vision for what should be available to address those needs. The workgroup has also clarified the “how” by providing detailed recommendations about implementation, such as creating the legal templates.

Ed Sullivan of CSU and Amy Fong of CSIS expressed concern that if a funding proposal is put forward as a one-time expense, it would not be sustained. Ben Chida clarified that the Governor’s Office does not view funding as a single-year endeavor. The project is part of their vision for effective government stewardship, which entails building infrastructure over a longer time horizon.

### Recommendation Language

The workgroup agreed that at this point, it is most prudent to propose a budget range for 2021-22, with a goal of providing more detailed figures in the second legislative report. The group also agreed that funding for the data system should be provided in an ongoing manner. While some costs will be one-time, most will be ongoing. Sustained funding is particularly important for hiring, to ensure that the state could recruit appropriate talent. It would also prevent a scenario where the significant time and effort put into planning for and standing up the system is wasted.

Ed Sullivan of CSU expressed concern about going any lower than \$15 million, which would be a skeletal budget for a project of this scale. He noted that a lower funding level could compromise core functionality or risk the removal of one of the tool sets, such as the operational tools that would be of most use to students and families. He further emphasized the need for champions for the proposal, given all the competing needs in the state.

Amy Fong of CSIS suggested that any request emphasize how the funding will benefit the public. For example, clarify that it would be worthwhile to invest in a user-centered design process and focus on data quality to yield reliable, actionable information. Chris Furgiuele of UC concurred.

Patrick Perry of CSAC wondered if it would be helpful to provide a range of options, including bare-bones and more comprehensive technology solutions. Joy Bonaguro of GovOps indicated that it would be difficult to do so at this point because there are so many unknowns about the costs until the technical architecture is more fully fleshed out. Therefore, the group determined that it would be preferable to provide one overall budget figure with descriptive detail about the cost categories and their rationale.

Sara Pietrowski of SBE suggested that the language should provide context about how budget categories relate to each other, particularly investments that are precursors to others—such as the need to upgrade the CALPADS infrastructure before the K-12 eligibility tools can be scaled.

The group considered more prescriptive language, such as guidance on whether funds would flow through GovOps or go directly to the agencies, or recommending whether funds should be continuously appropriated or come from the general fund. They also weighed trying to provide more precise figures.

However, they determined that this level of specification should be determined through the legislative and budget process.

## **VOTE**

The group voted on whether to take the following approach:

The narrative for the fiscal impact section of the report should include the following categories for 2021-22:

- Staffing and operational costs
- Governing board
- Supporting partner entities to provide the necessary data
- Building the analytical tools (list what these are)
- Upgrading CALPADS to allow for real-time data for transcripts (impacts success and durability, must happen first for CCGI to scale)
- Beginning scaling CCGI (list the services)

Estimated Total: \$15-20 million

The section should also emphasize the need for a commitment to ongoing general funding.

All but one of the workgroup members approved the fiscal impact recommendation. Ed Sullivan of CSU abstained because he did not want to vote for or against revenues that could impact the CSU budget.

## **Governance Proposal**

The workgroup considered a proposal from CLWDA to add one more seat to the governing board so that there are two positions for the agency: one for the Employment Development Department for employment, earnings, and labor market information; and a second to represent the workforce training data managed by the agency.

Patrick Getz, who has compiled workforce data for CLWDA's CAAL-Skills system, outlined the types of information that the agency could provide to better understand how workforce training is helping people out of poverty. For example, the agency could provide access to information on adult basic education, vocational rehabilitation, and programs serving justice-involved individuals, as well as many additional disaggregation categories on barriers to employment.

## **VOTE**

Most workgroup members approved the expansion of the governing board seats. Tom Vu of AICCU voted no because he felt departments should not have separate seats from their agency and suggested that the second position be an ex officio seat instead. Ed Sullivan of CSU abstained due to concerns that increasing the size of the governing board would dilute his agency's voice. Joy Bonaguro of GovOps abstained because her agency has been recommended as the managing entity.