

# California Cradle-to-Career Workgroup Meeting Summary

March 25, 2021

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 March 25, 2021 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; Leeza Rifredi, Bureau for Private Postsecondary Education; Barney Gomez, California Community College Chancellor's Office; Mary Nicely, Sarah Neville-Morgan, and Cindy Kazanis, California Department of Education; Tim Murphy, California Department of Technology; Natasha Nicolai, California Department of Social Services; Jennifer Schwartz for Elaine Skordakis, 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; Amy Faulkner, Employment Development Department; Joy Bonaguro, GovOps; Jeanne Wolfe, Labor and Workforce Development Agency; Sara Pietrowski, State Board of Education; Chris Furgiele, University of California Office of the President

## Updates

### Community Engagement

Kathy Booth of WestEd provided an overview of the work of the Community Engagement Subcommittee to document the types of people who will be likely users of the data system, which are codified in 23 personas. Each persona has been given a name, picture, and a demographic profile and includes the types of challenges they face, what they are looking for, and what they need to be successful. The personas are grouped into four archetypes—analyzers, planners, practitioners, individuals.

Kathy Booth of WestEd then described the audience matrix developed by the Community Engagement Subcommittee. Organized by persona type, the document lists goals that would be supported by information from the data system, which tools they are most likely to use, core messages about those tools, and how best to communicate with them. This information will be used to craft a communications plan. Finally, she noted that the personas and the audience matrix are being vetted more broadly with communities that embody the archetypes.

### Proof of Concept

Kathy Booth of WestEd noted that work on the proof of concept is on hold because the Legislature has asked to have until April 15 to review the December 2020 Legislative Report and the April 2021 Legislative Update before determining whether to release the remaining planning funds. However, the partner entities have begun work on fleshing out the proposed teacher retention dashboard including documenting underlying data structures, refining the list of key questions, and creating conceptual drawings on how outcomes could be visualized.

Cindy Kazanis of CDE shared that the delay by the Legislature has affected her agency's plans to improve CALPADS, which will impact both the analytical and operational tools. CDE is eager to provide any information that can help to move the proposal forward.

### Legislative Process

Kathy Booth of WestEd reminded the group that her team is compiling proposed edits to the trailer bill, as a courtesy, to share with the Department of Finance. Any remaining edits should be sent to LeAnn Fong-Batkin by March 31.

Then she noted that representatives from the Department of Finance, the Legislative Analyst's Office (LAO), and WestEd had testified at hearings at both the Senate and the Assembly in March 2021. An additional hearing is forthcoming for AB 99, a bill introduced by Assembly Member Irwin that references recommendations from the December 2020 Legislative Report.

Ed Sullivan of CSU noted that the idea of changing the governance structure was raised at both hearings. Kathy Booth of WestEd encouraged partner entities to reach out directly to the Legislature to share their thoughts on suggested edits, including increasing the number of seats for public entities.

Chris Furgiuele of UC flagged that AB 99 is missing language that exempts the Cradle-to-Career Data System from the Public Records Act.

Cindy Kazanis of CDE reflected on other suggestions raised by the LAO, such as requiring that the Cradle-to-Career Data System follow the Project Approval Lifecycle (PAL) process. Barney Gomez of CCCCCO inquired if the intention was to circumvent state processes. Kathy Booth of WestEd noted that the planning process had been implemented in accordance with CDT's requirements. For example, the Request for Information that informed the recommendation to use a master data management solution to link records was issued by CDT. CDT made changes to the process suggested by subcommittee members to ensure it conformed with state requirements.

Joy Bonaguro of GovOps noted that she has been in conversations with staff that oversee the PAL process. They have indicated that the work done through this planning process exceeds PAL standards. While some forms may need to be filled out to document work done to date, CDT is closely monitoring the work of this planning effort. She also noted that the proposed Cradle-to-Career Data System is a data infrastructure project, as oppose to a software development project, which may be subject to less restrictive planning requirements. She has had some preliminary conversations with CDT on a process for streamlining the PAL process for similar infrastructure projects, based on the positive experiences modeled in cases such as the Homeless Data Integration System.

Amy Fong of CSIS clarified that the PAL process outlines a means for designing a structure, conducting a business analysis, and doing a procurement analysis. She suggested that the ways that the planning process aligns with the PAL process could be spelled out in the April 2021 Legislative Report.

Tom Vu of AICCU noted a concern raised by one Assembly Member about the cost and feasibility of the data system and asked whether WestEd had recommendations for how to address this type of statement. Kathy Booth of WestEd clarified that her organization's role is to summarize the recommendations of the workgroup in three reports to the Legislature, and is not participating in the legislative process. She encouraged partner entities to reach out directly to the Legislature to address concerns raised in the hearings. Cindy Kazanis of CDE noted that the AB 99 hearing will be the first week

of April and encouraged the partner entities to present a united front in support of the proposed data system.

### P20W Data Points

Kathy Booth of WestEd described the process used to refine the initial list of data points that was approved by the workgroup in 2020. She noted that this had been a very collaborative process that leveraged the expertise of data experts at the partner entities, as well as researchers and advocates who will use the data system.

The refinement of the P20W data set served several purposes. The task of documenting underlying data sets and determining how to reconcile them in a linked data set is a frequent stumbling block in developing longitudinal data systems, so the fact that this work is largely completed will speed the development of the analytical data set. It will also help to reduce costs for data system implementation because data can be utilized in its native format, rather than requiring the state agencies (and all the educational and social service institutions that provide data to them) to adopt uniform data definitions. Finally, it addresses the requirement in the Cradle-to-Career Data System Act that the workgroup ensure quality for data points included in the state system.

Kathy Booth of WestEd noted that as a result of reviewing the source data and its quality, the subcommittee made a number of important recommendations. In some cases, they suggest that data only be provided by one source that has higher quality information—such as foster youth status coming only from CDSS. In other instances, they flagged priorities for data improvement, such as information on a-g course completion and grade point averages (GPA). This is an example of where the operational and analytic use cases overlap—when CDE updates CALPADS to integrate with CaliforniaColleges.edu, it will help to improve the underlying data quality, such that some data elements currently removed from the P20W data set could be included in the future.

Finally, she clarified that WestEd is now consolidating all of the documentation and the partner entities are engaged in vetting the updated list of data points. This review will result in a finalized recommendation regarding which entities should provide each data point.

Kathy Booth of WestEd walked through the content of the reference materials, which describe how each data point has changed and lists implications for the dashboards and query builder. Some of the visualizations will be adjusted slightly to address data points that were removed. However, most of the proposed changes expand the P20W data set. More than 35 new data points have been added, such that the P20W data set will consist of 200 data points. Ten were removed because they are not currently collected by the partner entities or are of insufficient data quality. Four data points were flagged as being available by request only due to their complexity.

Finally, Kathy Booth of WestEd noted that a number of the postsecondary partners are concerned that they cannot share financial aid information that originates from the FAFSA due to guidance released by the U.S. Department of Education's Privacy Technical Assistance Center in 2017. WestEd is seeking clarification from the federal government on allowable use.

Patrick Perry of CSAC asked if data can be resubmitted, particularly if edits are made to data that appears problematic. Kathy Booth of WestEd noted that data can be resubmitted and the dashboards rebuilt. The system disclaimer makes clear that data may vary over time. More specific policies about the cadence and process for corrections will be made once the technology tools are selected.

Patrick Perry of CSAC noted that the trailer bill directs CSAC to get GPA data from CDE rather than directly from local educational agencies (LEAs). Given concerns raised by CDE about the quality of grade data, this requirement may need to be revised.

Cindy Kazanis of CDE noted that currently her agency does not collect transcript-level information. However, the funds requested for the proof of concept would allow CDE to begin doing the necessary to provide validated GPA information in the future. Glenn Miller of CDE concurred, noting that CDE will also need to work with the partner entities to develop a standardized methodology for calculating GPA.

Kathy Booth of WestEd noted that the Policy & Analytics and Practice & Operations Advisory Groups expressed strong concern about GPA not being included in the data set.

Chris Furgiuele of UC stressed that none of the data in the analytical system will be coming from an official transcript and this level of information cannot be provided for research purposes. Ed Sullivan of CSU concurred.

Bruce Yonehiro of CDE asked for clarification about how data points, such as the Perkins Career and Technical Education Concentrator, could be requested. Kathy Booth of WestEd explained that researchers could request summary data using the expedited process or anonymized individual-level data using the comprehensive process, as laid out in the data request process recommended by the workgroup last fall. In either case, the data provider would determine whether the data point can be released.

In responding to questions posed by Assembly Republican Consultant Carolyn Nealon, Kathy Booth of WestEd noted that employment and earnings information will only be available for students who enroll in postsecondary due to legal restrictions about collecting social security numbers, which are used when matching employment records. She also stated that information in the employment dashboard and the query builder would only be available for those who graduate from college.

Chris Furgiuele of UC asked who will lead process of refining data points in 2021-22. Kathy Booth of WestEd clarified that the Definitions Subcommittee recommended that the technical definitions be fine-tuned once it is clearer how specific data points will be represented in the dashboards. For example, adjustments might be made to which students are included in a calculation, to provide a clear through-line between related measures. The staff at GovOps will lead this process.

Chris Furgiuele of UC asked whether a timeline had been established for data submissions. Kathy Booth of WestEd reminded the group that the data needs to be loaded during fiscal year 2021-22, per the plan submitted in the December 2020 Legislative Report. However, before data can be loaded, GovOps must staff up, the proof of concept project needs to be implemented, and the data tools need to be procured. Given the delay in implementing the proof of concept, the timeline is unclear.

Jennifer Schwartz of CHHS asked how additional data sets will be integrated for early childhood, particularly those that will be managed by CDSS. Sarah Neville-Morgan of CDE noted that CDSS had agreed that for the first year, information should be provided only by CDE. Then, once CDSS has built out its data system, additional data points or new data infrastructure could be integrated. Natasha Nicholai of CDSS concurred and clarified that the term “early learning and care” should be understood to mean “preschool.” Child care data will not be integrated until 2023, when the migration to new data systems is complete.

Jennifer Schwartz of CHHS flagged that the trailer bill language needs to be clarified so that it will be clear that CDSS will not provide early learning and care data initially. Bruce Yonehiro of CDE echoed this concern, noting that he had understood the model to allow data providers discretion about which data points to provide, rather than having the list be mandated by law. Furthermore, as currently written, the trailer bill appears to require each data provider to submit all data points.

Cindy Kazanis of CDE asked whether the P20W data elements are specifically listed in the trailer bill. Kathy Booth of WestEd responded that it references the list of data points approved by the Workgroup in the December 2020 Legislative Report. Cindy noted that the specific data points should not be codified in the legislation to allow the list to evolve over time.

*Note: the related trailer bill language reads:*

*10871. (a) Data Providers, as defined in Section 10861, shall submit, at least annually, the data points contained in the P20W data set as defined in Section 10861. <sup>1</sup>*

*(1) Source data shall be consistent with data definitions and standards adopted by the Governing Board. Data Providers shall make every effort to ensure source data is of the highest quality prior to submitting to the Managing Entity for inclusion in the Cradle-to-Career Data System.*

*(2) Any data sharing agreements necessary to effectuate this provision shall occur using the "IDEA/BUCP" process as defined in Section 10861, or a successor process by which state entities share data.*

*(b) The Managing Entity shall work with the Data Providers to ensure data quality, in furtherance of paragraph (8) of subdivision (a) of Section 10867.*

Ed Sullivan of CSU noted that his agency's general counsel are reading the trailer bill in the same manner as described by CHHS and CDE. His agency cannot support participation in the data system if it will be forced to collect new data points or provide information that may not be shared per state and federal law, such as financial aid information. He was also concerned that new data points will be added to the list through the legislative process and not in consultation with the data partners.

Jennifer Schwartz of CHHS noted that this issue could be taken up by members of the Legal Subcommittee. Amy Supinger of WestEd noted that the subcommittee had begun this conversation in their meeting earlier in the week and determined that the education partners would work together to draft potential alternative language.

Joy Bonaguro of GovOps suggested that language could be added to the trailer bill that ensures that data can only be provided "consistent with prevailing law." Ed Sullivan of CSU replied that he was concerned that state law could trump what is codified in the legal agreement between the data providers and the managing entity.

Bruce Yonehiro of CDE noted that under the Family Educational Rights and Privacy Act (FERPA), education agencies may not disclose records without student approval, unless usage falls under an exemption. This requires that the agency make the decision to provide the data to an authorized representative for limited purposes. Education agencies must be allowed to negotiate a participation agreement with GovOps that specifies the data elements to be provided by each entity, rather than

---

<sup>1</sup> Section 10861 covers definitions, with the P20W Data Set described as "the data set adopted by the Governing Board, including but not limited to the approximately 160 data points described in the Report to the Legislature required by pursuant to subdivision (a) of Section 10856 and as adjusted by the Governing Board. The Governing Board is authorized to add or remove data points based on changes in the usage of the element."

putting this requirement in statute. Putting requirements in state law or delegating this authority to the governing board will violate federal law.

Ed Sullivan of CSU echoed this analysis. He clarified that CSU is not concerned about the proposed list of data points. His agency just wants to ensure that they will only disclose those data points that CSU confirms are in compliance with the law. That is why they want to change the trailer bill language.

Thomas Vu of AICCU agreed with making a change to the trailer bill to address this issue.

Joy Bonaguro of GovOps suggested referencing language used in other states that have already implemented similar systems.

Amy Supinger of WestEd thanked the workgroup members for clarifying the nature of the concern, and posited that it may be more of an issue of wording than intent. Bruce Yonehiro of CDE suggested adding language such as “provided that data sharing and contribution is voluntary and flexible over time.” Ed Sullivan of CSU did not think this language was sufficient, but Tom Vu of AICCU thought if this language were included in the trailer bill, it would address the challenges that some of his member institutions would face, particularly if they do not currently collect specific data elements.

Jeanne Wolfe of CLWDA asked whether the concern was about the list of data elements, or the proposed requirements around data sharing.

Ed Sullivan of CSU noted that his agency does not want to be put in the position to disclose some of the data points that it thinks cannot be legally shared such as financial aid and some information collected in college applications.

Chris Furguele of UC indicated that his agency was not concerned about the current list but was uncomfortable with how the data set might expand. The proposed governance process does not provide a specific mechanism for expanding the data points, beyond noting that the managing entity will conduct a feasibility study and the request will be considered by the governing board. UC does not want to be put in a position where they will be forced to collect new data that they do not have the resources to gather.

## **VOTE**

Based on this discussion, Kathy Booth of WestEd edited the item up for a vote to read:

Provided that data sharing is based on a contractual agreement between data providers and the managing entity, affirm the recommended changes to the P20W data points and the public-facing tools.

Most of the workgroup members voted to approve this item. Chris Furguele of UC, Tom Vu of AICCU, Leeza Rifredi of BPPE, Jeanne Wolfe of CLWDA, Jennifer Schwartz of CHHS, and Tim Murphy of CDT voted “yes with reservations” and stated that the trailer bill must be amended to address the workgroup concerns.

Ed Sullivan of CSU voted no, per the issues raised in the discussion.

## **Opt Out Language**

The workgroup reviewed the revised language provided by the Legal Subcommittee. They made one additional edit for clarity.

## **VOTE**

The workgroup voted unanimously to adopt the updated language, as amended in the meeting.

### **Security Framework**

After Kathy Booth of WestEd provided a review of the various components of the recommendations from the Technology & Security Subcommittee, Ed Sullivan of CSU noted that the U.S. Department of Education's Office of Federal Student Aid just released new guidance related to cybersecurity. Baron Rodriguez of WestEd was unaware of this change, but noted that the proposed Cradle-to-Career Data System security framework is likely to meet the new standards because it is already designed to address financial, health, and California-specific requirements, which are stricter than FERPA. Jennifer Schwartz of CHHS concurred, stating that health information must be managed according to the higher standards just required for financial aid. Baron Rodriguez of WestEd also noted that the Technology & Security Subcommittee had flagged that the framework should be edited in the fall to reflect other updates to federal requirements.

Tim Murphy of CDT asked for clarification on how the security framework will inform the procurement process. Baron Rodriguez of WestEd noted that it would shape requirements for procurement and that all contractors would be required to comply with it.

Chris Furgieuele of UC noted that he had concerns about one portion of the permissions protocol that had unclear language about the ability of parents to view information in the data system—it can be interpreted as applying to either the operational tools, the analytical tools, or both.

Tim Murphy of CDT agreed that the document needed some additional clean up work. However, the overall direction of the document is sound.

Chris Furgieuele of UC noted that he also had concerns about the RACI analysis that is included in the permissions protocol. Tim Murphy of CDT agreed that a few items could be refined.

Given these concerns, the permission protocol was removed from the list of documents being approved by the workgroup. Because there were other areas of the permission protocol that the Technology & Security Subcommittee felt could not be finalized until after the procurement process, the document will be revisited in the next fiscal year, with these specific concerns flagged.

## **VOTE**

Most of the workgroup members voted to approve this item. Ed Sullivan of CSU voted “yes with reservations” and stated that the security framework should be updated to reflect current federal requirements.

### **Deidentification Protocol**

Kathy Booth of WestEd described the proposed protocol and enumerated how it was built upon existing policies implemented by CHHS, recommendations from the National Center for Education Statistics, and rulesets implemented in other tools like the federal College Scorecard. She shared that the advisory groups were concerned that the deidentification protocol would be undercut by the expansive definition of personally identifiable information.

Chris Furgieuele of UC felt that the wrong people had been asked to develop and recommend the policy and that the partners had not been given sufficient time to review it. He was concerned that some of

the terms used were not sufficiently defined, that specific examples in the policy were confusing, and that the examples were insufficient for the level of complexity of the information in the data system. He argued that this policy does not need to be finalized now and could be referred to the managing entity for development in the next fiscal year.

Kathy Booth of WestEd clarified that the Cradle-to-Career Data System Act requires the workgroup to specify how individual privacy will be protected, and that that deidentification protocol is a key component of this strategy. She noted that the people who developed the draft policy included the authors of CHHS's and CDE's deidentification policies, and mirrored a policy posted by the CCCCO. However, she stressed that it was critical that all parties have sufficient time to review and weigh in on the policy.

Chris Furgiuele of UC stressed that it will be critical for the managing entity to have the expertise to address complex policies such as the deidentification protocol, particularly because it is unlikely that the partner entities can develop a comprehensive protocol before the data are integrated.

Ed Sullivan of CSU agreed that it would be helpful to have more time and bring in additional experts to review the proposed protocol.

Amy Fong of CSIS asked where the boundaries between the workgroup and the managing entity will be. Will GovOps be able to make operational decisions to support the data system once it undertakes the development?

Kathy Booth of WestEd reminded the group that the Cradle-to-Career Data System Act tasked the partner entities with developing recommendations to the Governor's Office and the Legislature to inform the implementation of the data system. Having conversations in advance will help to smooth implementation later, but the managing entity is the one that will ultimately be responsible for enacting the data system. Ed Sullivan of CSU noted that by establishing general policies, the workgroup can help create guard rails to guide what gets implemented by GovOps.

Amy Fong of CSIS followed up to emphasize that it will be important to ensure the managing entity has the flexibility it needs so it can adapt to future environments as it brings the system online. She reminded the group that it is doing best it can with current information.

Patrick Perry of CSAC noted that there may be other options for deidentification. For example, the Integrated Postsecondary Education Data System (IPEDS) perturbs data to get around small cell suppression, or the managing entity could create a synthetic data set.

Chris Furgiuele of UC noted that it might be preferable to have different cell size restrictions based on specific data elements. For example, application and admissions data might implement a higher threshold than other types of information.

Ed Sullivan of CSU suggested creating another homework team to revisit the recommended protocol, with an understanding that the managing entity may make additional revisions at a later point. Chris Furgiuele of UC agreed and suggested that this be prepared for the May workgroup meeting.

Kathy Booth of WestEd asked the partner entities to alert LeAnn Fong-Batkin about the people they would like to join the homework team.

## April 2021 Legislative Update

Kathy Booth of WestEd described the proposed content for the April Legislative Update and asked the workgroup if it would be possible to provide a more detailed budget, as requested by the LAO. To contextualize the discussion, she shared insights gleaned from reading the McKinsey report that was released in 2008 regarding the estimated costs for a longitudinal data system.

Some of those requirements reveal how far technology has evolved in the ensuing decade, particularly regarding how to link and align data. The report also specified a number of line items that have already been addressed through the Cradle-to-Career Data System planning process, as well as items that are no longer considered part of the scope of this process (such as revamping CDE's dashboard or providing broadband connectivity to specific regions). However, there were also many similarities, such as focusing on upgrading underlying data quality, embedding staff within data providers, implementing a user-centered design process, and providing professional development.

Cindy Kazanis of CDE thought it would be helpful to compare the current proposal to technology requirements in the McKinsey report, which was released a lifetime ago (in IT terms). For example, CALPADS had not been fully implemented at that time. Similarly, costs have been reduced significantly. What previously cost thousands of dollars in hardware and software costs can now be implemented for a few dollars a month in the cloud.

Ed Sullivan of CSU noted that it is the scope of the data system, not the underlying technology that will drive up costs. Keeping the data system within the scope as currently described will help to keep it on budget. He contemplated the value of increasing the size of the ask, particularly given that there are now significant one-time dollars available—which was not the case when the first Legislative Report was submitted in December.

Patrick Perry of CSAC preferred leaving the budget figure at its current level. If the ask is too big, it will compete with other items that are high priorities. It would be better to get the data system off the ground and reevaluate the budget after a few years. Joy Bonaguro of GovOps and Amy Fong of CSIS agreed.

Tim Murphy of CDT concurred, noting that the largest costs for technology projects are always operational—for the people that implement the system and manage process like change and stakeholder management. The current plan is aligned with effective technology practices, such as emphasizing interagency agreements, picking the most meaningful data points, and determining how to deidentify information. Any attempt to guess on a larger dollar value without the proof of concept will just be a spitball estimate.

Sara Pietrowski of SBE reminded the group that the recommended Cradle-to-Career Data System fiscal impact analysis reflects how CDE's School Dashboard was developed. A first year of appropriations was used to get the dashboard developed, and then CDE asked for ongoing funding source to maintain the system.

Assembly Republican Consultant Carolyn Nealon noted that the initial P20W data set will be primarily education data. But when other data sources are added in later years, those agencies will need funding to support the embedded staff and planning activities. One possibility would be to budget for those positions that will be needed in the future.

Kathy Booth of WestEd noted that she will integrate the issues raised at today's meeting into the April Legislative Update and send it out by the end of the day. Partner entities will have until 9am on Tuesday, March 30 to give input, so that it can be submitted on April 1.

Finally, she noted that the April meeting will be cancelled. In May, the workgroup will review recommendations from the Community Engagement and the Legal Subcommittees. The June meeting has been moved to June 17 and will cover which entity will provide each data point and the final Legislative Report.