

California Cradle-to-Career Workgroup Meeting Summary

October 29, 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 October 29, 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; Joy Bonaguro, GovOps; Doug Leone (for Jeanne Wolfe), Labor and Workforce Development Agency; Sara Pietrowski, State Board of Education; Chris Furgiele, University of California Office of the President

Public Comment

Colleen Moore of EdInsights expressed appreciation for the hard work and openness of workgroup members. The unprecedented challenges experienced during 2020 reinforces the need for a longitudinal data system. She urged the workgroup to establish priorities for the components identified for phase one to ensure that the effort will be feasible, given current budget constraints.

Community Engagement Update

LeAnn Fong-Batkin of WestEd provided an update on the community engagement effort, noting that more than 400 people had participated and over 100 survey responses had been received. She went over key themes raised in the webinars and feedback surveys and noted where those issues were being addressed in the planning process. Workgroup members concurred with feedback that well-designed public tools could help foster data use.

Data Request Process

After the facilitator described the process used to develop the proposal, two members of the homework team walked through its components. Evan White of the California Policy Lab and a member of the Policy & Analytics Advisory Group noted ways that the original data request proposal had been amended to address advisory group concerns and outlined the revised responsibilities for the managing entity. Randy Bonnell of CDE described how the revisions resolved workgroup concerns and noted implications for data providers.

Questions raised by workgroup members included:

Could requestors ask for data elements that are not already in the Cradle-to-Career Data System? No. Data providers will only be required to provide information listed in the P20W data set. However, they could upload other elements for research purposes. Available data would be displayed through an index on the project website, and only those elements could be requested.

Would data providers have a role in examining the data set to evaluate the validity of these matches? Yes. Data match validity would be addressed as part of the implementation of master data management. Data made available for requests would be based on match algorithms that have been approved by the data providers.

Will data providers have sufficient capacity to review all of the requests that need their approval? It is not clear how many requests will be submitted, and data providers will need to allocate resources to respond to requests. However, the proposal calls for a review after the first year to see how well the process is working.

Will requestors get access to the full data set, or just those elements specified in the proposal? Just those data points listed in the proposal.

Will requestors have to get approval if they change the data points they want access to? Yes. In order to comply with federal law, which requires that unitary data can only be provided for an approved purpose, changes to the desired data points will need to be documented. In cases where the changes are minimal and align with the approved data questions, this can be implemented through a streamlined process where the existing agreement is modified. However, if the additional data elements signal a new set of research questions, then a new request will need to be submitted. This is one reason why it is important for the data providers to work with the requestor to ensure the desired data points will answer the research questions.

Would it be possible to give approved researchers access to the entire data set? Not in cases where researchers are accessing unitary data, per federal law. However, the workgroup has expressed interest in creating a synthetic data set or in data sets that contain de-identified data. This topic can be addressed during spring 2021 planning.

How will requestors access the data? Third parties will access information in a secure data enclave, which they can access virtually. Approved users will be given a secure log in that only provides access to the data points they are approved to use. The online environment will also include the analytical tools necessary to conduct the research.

How will data be evaluated when it is removed from the data enclave to make sure that individual identities remain private? Once the analysis is complete, researchers would be able to compile summary data that they want to remove from the system. The proposal specifies a disclosure review process where experts would examine these tables before they can be removed from the system, to ensure individual identities would not be disclosed, such as by having only a few students represented. This is a practice that has already been implemented in data systems in California and in the Texas data system.

VOTE

The Workgroup approved the proposed data request process, but two members agreed with reservations.

Ed Sullivan of CSU expressed concern that having the managing entity prepare requests for review will politicize the agency, particularly if it is either more conservative or more liberal than data providers. The Research Advisory Board provided a rotating group of reviewers that would ensure a diversity of perspectives over time.

Chris Furgiuele of UC was concerned that the data providers would not have sufficient bandwidth to review data requests and thought this task would be better implemented by the managing entity. In addition, he noted that the criteria for acting in good faith and approving requests that support the public good are vague, which leaves interpretation up to the data provider.

Publicly Available Data

After the facilitator described the process used to develop the proposal, Glen Forman and Doug Leone of CLWDA outlined the rationale for adding apprenticeship data to the P20W data set and integrating this information into the dashboards and query builder tools.

Some participants noted that they thought apprenticeship was already slated for inclusion in the data set and were pleased to see this information added, particularly as apprenticeships offer a strong pathway to a living wage and the information comes from a mature data system. Mary Nicely of CDE noted that including apprenticeship data would help to answer the question of whether students who participate in K-12 career and technical education enter apprenticeships.

Then, Alyssa Nguyen of the RP Group (and a member of the Research Agenda Subcommittee) walked through the proposal.

Questions raised by workgroup members included:

How does this proposal relate to the work underway in the Data Definitions Subcommittee? The subcommittee is reviewing each of the data points listed for the P20W data set. Each partner entity records whether and how they capture the information, the dates for which the item is available, the values associated with each item, and the quality of the information. After examining the similarities and differences between agency definitions, the subcommittee determines how the information should be displayed. For example, the partner entities found they identify economically disadvantaged status in a variety of ways, including whether students are eligible for free-and-reduced price meals, receive financial aid, or if a parent has a college degree. Therefore, they determined that these more precise variables should be used, rather than creating a single data point that is defined differently by each data provider.

Will data points be included if they are collected by only some data providers? This will be determined on a case-by-case basis. If the information is considered of sufficient quality, then the subcommittee would indicate it should be included. They can also note data points that may need further examination, such as if information may be missing for a large number of people.

Will data points be assigned a definitive provider? Yes. In some cases the subcommittee has recommended that a data point be provided by a single agency, even though it may be available from other partners. For example, the group determined that all foster youth data should come solely from CDSS. In other cases, data will be combined from multiple providers, such as financial aid information that would come both from postsecondary institutions and CSAC.

How will the public tools handle data points that have specific definitions for individual partners, such as the allowable definition of foster youth needed for the Local Control Funding Formula? Information will be contextualized in the tools, such as providing a set of options for various definitions of foster youth or noting when data is provided by only some entities. In cases where the data point is derived—meaning that it is constructed from several elements in a proscribed manner—data providers are asked to identify whether they would like to provide that derived element to the managing entity, or if they would like the managing entity to calculate the data point on their behalf (and if so, using what formula). The public tools will also be examined as part of the design process to make sure that they don't create a competing narrative when compared to agency dashboards or reports, which would undermine the Cradle-to-Career mission of being a reliable data source.

Will the data set include curriculum data, such as whether courses are eligible for transfer to UC and CSU (the ASSIST database) or associate degrees for transfer that are accepted at specific four-year colleges? While this was considered as part of the process, it was not prioritized.

VOTE

The Workgroup unanimously approved the proposed publicly available data, including the addition of apprenticeship information.

Managing Entity Responsibilities

Amy Fong of CSIS walked the workgroup through the proposed responsibilities for the managing entity, including the new responsibilities outlined in the data request process.

Doug Leone of CLWDA noted that it will be a challenge for the managing entity to handle the various data structures for each data provider individually and suggested that it be empowered to convene small teams to resolve issues shared across providers. This recommendation will be included in the legislative report.

Barney Gomez of CCCC suggested that a service level agreement be implemented with the managing entity—a topic that is slated for the Technology & Security Subcommittee.

VOTE

The Workgroup approved the proposed data request process, but three members agreed with reservations and Joy Bonaguro of GovOps abstained.

Ed Sullivan of CSU reiterated his previous concern that the managing entity will become politicized due to its responsibilities for the data request process.

Chris Furguele of UC noted that the scope of responsibilities entails significant expertise but does not indicate how many staff or how much funding will be required.

Tom Vu of AICCU echoed these concerns, particularly regarding the need to develop a budget.

Priority Research Questions

The workgroup discussed how to determine the scope of data that should be included in a future phase of the data system related to health, social services, and workforce training, particularly given that the California Cradle-to-Career Data System Act did not include any priority research questions on these topics.

Elaine Scordakis of CHHS felt that the priority questions should be identified by the system users.

Amy Fong of CSIS felt that it would be advisable to wait until the system is built and has been underway for a period of time before seeking to expand it. Empirical information on the types of questions that users prioritize and where they experience difficulties using data should inform this decision. Furthermore, there are a number of data projects underway now that may be able to provide high quality information by the end of phase one.

Chris Furgiuele of UC concurred that social service, health, and additional workforce data would be valuable for research and public facing tools, but that it was premature to try to create any specific requirements. This process should wait until the system has taken shape and research questions become clearer.

Doug Leone of CLWDA noted that integrating workforce training data will be complex given that information is held by a number of different departments and agencies. It may also require additional work to ensure that appropriate privacy protections are included for these populations.

Amy Faulkner of EDD clarified that the workforce training data sets include the following sources that are not among the partner entities: Employment Training Panel, Trade Adjustment Assistance, and Workforce Investment Act Title I.

Cindy Kazanis of CDE noted that the workforce training data that her agency collects for adult education is not as robust as the information in CALPADS. She agreed with Amy Fong about the value of examining needs after the data system is built and predicted that research questions related to workforce training will be significantly different as the project nears the end of phase one.

Tom Vu of AICCU expressed interest in getting addition information about social service needs of individual students.

Ed Sullivan of CSU agreed that social service and health data would be valuable in operational tools. However, it will be important to secure student consent before sharing their information, particularly when data are available on specific individuals.

The facilitator asked whether the group would like to convene a homework team during November to develop a list of recommended health, social service, and workforce training data points for phase two of the data system. The group recommended waiting, but some wanted to revisit the question in the spring.