Cradle-to-Career Definitions Subcommittee Meeting Summary

August 20, 2020

This document provides a summary of the key points that emerged from the August 20, 2020 meeting of the Definitions Subcommittee. The suggestions from this group will be used to craft specific definitions for elements in the P20W data set. 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/definitions-subcommittee (click on "Meeting Materials").

The goals of this meeting were to:
- Identify ways to improve the documentation process
- Establish display options and caveats for the public display of the following elements:
  - Race/Ethnicity
  - Sex/Gender
  - Sexual Orientation/LGBTQ
  - Social Services items (CalFresh, CalWORKS, Medi-Cal)
  - IEP
  - Financial aid items (family contribution, financial aid dependency)
- Note issues found when examining the proposed format for person-matching elements

The following Definitions Subcommittee representatives attended the meeting:

Randy Tarnowski, Association of Independent California Colleges and Universities; Joanna Murray, Bureau for Private Postsecondary Education; Todd Hoig, California Community College Chancellor's Office; Jerry Winkler & Channa Hewawickrama, California Department of Education; Patrick Delaney & Akhtar Khan, California Department of Social Services; Janet Buehler, California Department of Technology; Jennifer Schwartz, California Health and Human Services Agency; Patrick Getz & Daniel Rounds, California Labor and Workforce Development Agency; Rima Mendez & Amy Fong, California School Information Services; Monica Malhotra, California State University; Adrian Felix, California Student Aid Commission; Erin Skubal, California Teacher Credentialing Commission; Margo Gonzales, Employment Development Department; Sara Pietrowski, State Board of Education; Chris Handy, University of California Office of the President; Todd Britton, University of LaVerne; and Valerie Mendelsohn, West Coast University

Improving the Documentation Process

Subcommittee members reflected on their experience filling out the documentation templates for specific data points. The group noted that the process was easy and not overly burdensome.

However, there was some confusion about the concept of a data point being self-reported versus a provider-reported value. The facilitators clarified that information that comes solely from the individuals being served (such as on an application form or on a portal where individuals can make updates or corrections on their own information), the item should be listed as a self-report. But if professionals at the institution are involved in generating the value (such as amending what value is assigned or contributing their own information), then the item should be listed as provider-reported. When filling

out the form, partner entities should make sure they are aware of how the education and social service institutions that provide data to them collect the relevant data point.

The group agreed on two adjustments to the data that had been requested for this meeting:

- Create one form for identifiers associated with education institutions or service providers (such as a code assigned to a specific local educational agency or college) and a separate form for identifiers that providers use for the people they serve (such as a code assigned to a specific student or client).
- Work directly with the Department of Health Care Services rather than the Department of Social Services on Medi-Cal status.

The group also requested adjustments to the display options documentation to more clearly distinguish between data that would appear as a caveat in public-facing tools and information that is useful to understand about the availability of data from particular partner entities.

Race/Ethnicity

The subcommittee examined a proposed approach for displaying race and ethnicity data in the public-facing tools and discussed a series of questions based on differences in how the partner entities record this value.

Categories

Of central concern was the proposal to have a single list of values that treats Hispanic/Latino as a race rather than listing it separately as an ethnic group. While federal agencies have both education and social service providers first determine whether an individual is of Hispanic/Latino ethnicity and then collect race separately, in education contexts Hispanic/Latino is frequently shown in a single list with racial categories like Asian and African American. The list of proposed was the list associated with the Integrated Postsecondary Education Data System (IPEDS), which includes:

- American Indian or Alaska Native
- Asian
- Black or African American
- Hispanic or Latino
- Native Hawaiian or Other Pacific Islander
- White
- Two or More Races
- Not Reported

This values list was selected because the information was readily accessible from most partner entities due to federal conventions of grouping information in this fashion.

One participant expressed concern that it would be difficult for partner entities that record race and ethnicity data separately to fit into the IPEDS categories. A member of the public noted that the federal Office of Management and Budget assigns anyone who identifies as Hispanic/Latino to that category, even if they picked a different racial category.

The group discussed the possibility of having the public-facing tools break out Hispanic/Latino from the racial categories, but some were concerned that having too many options in the dashboard would be

confusing for data novices to parse. One participant suggested that a simple list be used for phase one, with the goal of determining ways to expand options over time. However, another noted that it would be critical for researchers to have access to more nuanced categories.

Some subcommittee members noted the importance of providing additional subcategories for broad groupings such as Asian. However, the facilitation team noted that the additional categories collected by the various partner entities are not consistent and not all partner entities collect more granular race and ethnicity data.

Changes to Definitions

One participant expressed concern about the change in IPEDS race/ethnicity definitions in 2008, which caused shifts in how students were assigned, and suggested that race/ethnicity data for postsecondary institutions should be available starting in 2010, when the new categories were fully in place. The facilitation team noted that most partner entities had race/ethnicity data starting in 2008. Channa Hewawickrama of CDE noted that there is race/ethnicity data on early care and learning starting in 2003. Given the long time horizons needed to evaluate early interventions, it would be valuable to make use of all available data.

Timeframes

The group contemplated the question of timeframes that would be used for the public facing tools. Although educational institutions may have information on a term basis and EDD may have data by fiscal quarter, partner entities will only be required to upload data once per year and the focus of the data system is on transitions across partner entities.

Conflicting Values

The group focused closely on the question of how to handle cases where different partner entities reported different values for individuals in the same time period, such as if a dual enrollment student were marked as African American at their high school in the same year that they were marked as Latino at the community college. The proposal suggested that the most recent value should be used, and if there were more than one value recorded in the same time period, that the student should be assigned to the Two or More Races category.

One participant suggested following the IPEDS guidelines for handling conflicting values, which uses the first value in an academic year rather than the last. Another noted that this was based on the concept of following a cohort of students over time, and suggested that different rules be established for cohort views (such as following Native Hawaiian or Other Pacific Islander students who started in early care and learning in a specific year), versus snapshot views (such as seeing the proportion of students who graduated from a postsecondary in a specific year who are Native Hawaiian or Other Pacific Islander). Jerry Winkler of CDE clarified that in K-12, unlike for IPEDS, the most recent value is used.

The group was uncomfortable with reassigning students with multiple values in the same time period to the Two or More Races category because this is a value that students can select, not a category for resolving conflicting data. One participant wondered if it is important to force a choice between categories, even though this could result in figures that add up to more than 100%. Others felt that it would be critical for the categories to be displayed as percentages that add up to 100%, particularly when working with non-researchers.

The group considered assigning students with multiple values reported in the same time period to a new category called Multiple Values Reported. One participant mocked up how data might be displayed using this approach, compared to the other approaches that had been discussed. Many subcommittee members were concerned that putting individuals into the Multiple Values Reported category essentially made them invisible, because students in various categories would all be lumped together.

Sex/Gender

Categories

The proposal included the following categories:

- Male
- Female
- Nonbinary
- Decline to State
- Not Reported

One member questioned the value of having separate values for Decline to State versus Not Reported. Another wondered if the proposed Nonbinary category would include all options other than Male and Female. The facilitation team confirmed that all gender options that are not Male or Female would be assigned to Nonbinary, given California regulations to create a nonbinary gender category for both education and social service organizations. However, partner entities that collect additional gender categories could make that information available for research requests.

Conflicting Values

While using the first value in a school year felt appropriate in the context of race/ethnicity, the group noted that gender identity would be better represented through the most recent value, given that this status may change over time and individuals may not feel safe reporting their nonbinary status in specific venues. This raised the question of whether some providers might be viewed as a more reliable source of information for some variables. Several subcommittee members thought that where possible, specific data providers should be the single source of truth, to reduce the complexity of parsing conflicting information.

Sexual Orientation

The facilitation team noted that the proposal recommended changing the title for this metric from Sexual Orientation to LGBT, which stands for Lesbian, Gay, Bisexual, and Transgender. This recommendation was made because transgender is a gender identity and not a sexual orientation.

Military Status

One participant noted some of the data points listed are calculations that include multiple data points rather than a singular element. For example, military status may be derived based on receiving financial aid that is specific to members of the military. The facilitation team noted that the templates will be updated so that partner entities can clarify whether an item is derived and if so, if they would prefer to provide the calculated result to the managing entity or have the managing entity calculate the variable on their behalf.

Additional Student Characteristics

Given the time spent discussing foundational decisions such as how to handle multiple values, many of the planned topics for the meeting were not covered. The facilitation team will use Survey Monkey to collect initial input on the proposed display options.

Recommendations

Conflicting Values

The group requested that the facilitation team gather information on how other states handle conflicting values for the same individual, to inform future discussions. However, the subcommittee also determined that it would be better to not resolve the question of how conflicting values will be handled at this stage in the planning process, as decisions may need to be made on a variable-by-variable basis, where being able to see the data in context would make the appropriate course of action clearer.

Race/Ethnicity Recommendations

- IPEDS race/ethnicity categories should be used for the public facing tools
- Researchers should be able to request other race and ethnicity categories that are collected by the partner entities
- The managing entity should work with the intended dashboard audience to determine ways to improve how race and ethnicity are displayed
- Use data from the earliest available year, but include a carefully worded caveat about the change in postsecondary race/ethnicity categories in 2008 for dashboard users, particularly related to examining trends
- When there are duplicate values for race/ethnicity in a school year, use the first value

Sex/Gender Recommendations

- When there are duplicate values for sex/gender in a school year, use the last value

Additional Recommendations

- For all data points, have the documentation clarify how information would be managed in a cohort view versus a yearly snapshot
- For public tools, display data at the level of school years, which run from July 1-June 30
- Allow researchers would be able to request data at the term or fiscal quarter level, if it is available from the partner entity in the Cradle-to-Career system

The group also agreed that it would be helpful to see revised versions of the suggested display options that reflected the conversation at the meeting. The facilitation team will send out updated documents and provide a Survey Monkey where the subcommittee members can provide feedback.