

CA DATA SYSTEMS MEETING
FEBRUARY 13, 2012
SUMMARY NOTES

A. Reasons developing a shared data system could be important as well as useful:

LAYER ONE:

Current Legal Requirements for Data that Crosses Multiple Segments:

- SB 1298 Compliance (Higher ed must capture K-12 IDs)
- Perkins legislation/Community College Accountability (budget act)
- Federal gainful employment report
- CSU disenrolled students report (budget requirement)
- CDE college-going report*
- CSAC gainful employment
- Annual Accounting of transfer #s/Rates (budget requirement)**
- UC College-Going rates for outreach programs
- Block Bill AB743 – CDE to community college information for placement
- School fiscal stabilization funds (SFSF) P-20 Data system(info on remedial courses required)
- UC transfer and articulation requirements
- CA subject matter programs, STEM
- Entry writing requirements
- Eligibility Study (UC/CSU)
- IPEDS student right to know
- CPEC held school/college codes for crosswalks

*Was collected by CPEC

**Transfer numbers (but not rates) were collected by CPEC

LAYER TWO:

CPEC Data – What data did CPEC collect that people accessed?

Primarily people used CPEC as a “one-stop shopping” area – all data could be accessed from individual segments, but CPEC had it all in one place.

- Headcount/demography
- Completions, by segment and by CIP code
- College-going rates, by district
- Transfer numbers (two-year to four-year)

Additional data that was collected by CPEC but that would require additional data for analysis:

- Fees/funding reports
- Future Enrollment Projections
- Capacity study (programs/facilities)

LAYER THREE/FOUR

What are the segments currently doing/what would they like to do that involve shared data? (Some in this list are currently being done; some are aspirational)

- Governor's office would like to see shared data be of use at the school/teacher/parent/student level to improve performance
- School feedback reports (where do students go/come from):
 - To K-12
 - To Community Colleges/CSU
- Improve assessment and placement
- Evaluate academic preparation
- Provide information to parents and students
- SARC – A to G requirements measures
- Postsecondary report cards
- Curricular alignment (CalPass, ASSIST, Credit Transfer, stat-finder)
- Wage Outcomes by campus and by program
- Labor force info
 - Unemployment Insurance
 - Long term wages
 - By industry
- Develop a consistent method to match K-12 wage data
 - What to do about non-high school graduates?
- Transfer – understanding student populations
- Outreach programs
- Currently do have local data-sharing agreements – this could make these easier

Aspirational goals:

- Participation rate analysis in general population
- Migration of graduates
- General equity reporting
- AB 540 students
- Financial Aid ?

B. What will it take to build a shared database, whether through centralized or federated model?

CORE DATA SOURCES:

- Data comes from the five signatories to the IA:
 - CDE, EDD, CCC, CSU, UC
- Additional Data sources:
 - National Student Clearinghouse
 - College Board
 - CSAC
 - Census
 - IPEDS
 - Department of Finance

Ultimately will need to develop common definitions of data elements for public presentation.

Core Data Available (Macro-Level):

Education Data

- Unit record student data (demographic, ID)
- Unit record awards (degrees, certificates, completion)
- Unit record financial aid
- Unit record student enrollment in institution
- Unit record course enrollment
- Unit record assessments
- Unit record application/admissions
- Special populations/programs
- Course characteristics
- Institution/school characteristics

Employment Data

- Quarterly wage records (by individual SSN, employer ID/Industry)
- Unemployment Records
- Federal/State workforce program participation

Outside Data

- NSC external enrollments

C. Decisions and Next Steps:

Governance Structure

Interagency Agreements will continue to be governing document – no JPA happening at this time.

Governing structure will, for now, include two representatives per segment, designated by that segment. There will be the potential to add Ex-Officio members as needed.

Advisory committee structure will include three committees:

- **Technical Advisory Committee**, to be comprised of those who can talk about the physical data structures, protocols of development, data definitions and data mapping. Need hardware, data and application people. This may include 2-3 people from each segment.
- **Users Advisory Committee**, which will include on-the ground folks who use the data on a day-to-day basis (the operational users).
- **Research Advisory Committee**, to include researchers who may use the shared data. This committee will be formed further down the road in the process.

Housing the Data

Group decided to table, for now, decision about where to house data. May ultimately go with a “federated” model where data is not housed in a single place. This model may allow group to better distribute the tasks and costs.

Regardless, there will be various levels of access – no one will have access to highest level of secure data (with SSNs). There will be various levels of lockdown, with public access to data marts/core tools/tables.

Next Steps

Next meeting will be with the Technical Advisory Committee (with members of the governance committee as well). Purpose of the meeting will be to:

1. Review/drill down on data domains - discuss relational table structures
2. Discuss whether federated storage model could support data queries, reporting and operational efficiency. What are potential benefits, costs and limitations of that approach moving forward?

Meeting scheduled for Friday, March 16th at WestEd Sacramento Office, 10:00am-3:00pm. Segment reps should identify the representatives to the technical advisory committee who should be invited to this meeting.

