April 16, 2015

The California State University Report: Unique Statewide Pupil Identifiers Progress Report

As required by California Education Code Section 10803, enacted by Senate Bill 1298 in 2009, the California State University’s Unique Statewide Pupil Identifiers Progress Report has been posted to www.calstate.edu/budget/fybudget/legislative-reports/.

Education Code Section 10803: (a) The Chancellor’s Office of the California Community Colleges, the University of California, and the California State University shall each establish a process by which colleges and universities within those systems issue, maintain, and report
information Unique Statewide using the unique statewide pupil identifiers specified in paragraph (3) of subdivision (e) of Section 60900. (b) Annually, on or before April 1, the Chancellor's Office of the California Community Colleges, the University of California, and the California State University shall provide a progress report to the Governor and the appropriate policy and fiscal committees of the Legislature. The report shall include a detailed timeline for the implementation, maintenance, and use of the unique statewide pupil identifiers pursuant to subdivision (a).

The California State University remains committed to developing a cost-effective and useful statewide data system, and to this end, it continues to work with the California Community Colleges, the University of California and others in refining its methodology to capture, maintain, and use the unique statewide pupil identifiers developed by the California Department of Education. We have been meeting regularly for two years to discuss and develop cost-effective methodologies for collecting, maintaining, and using the information.

Should you have any questions about this report please contact Ed Sullivan Assistant Vice Chancellor, Academic Research and Resources,

Sincerely,

Steve Relyea
Executive Vice Chancellor and
Chief Financial Officer

SR:dr

Full report posted to www.calstate.edu/budget/fybudget/legislative-reports/

c: Mac Taylor, Legislative Analyst
Michael Cohen, Director, Department of Finance
Gregory Schmidt, Secretary of the Senate
Diane Boyer-Vine, Legislative Counsel
E. Dotson Wilson, Chief Clerk of the Assembly
Timothy P. White, CSU Chancellor
Garrett Ashley, Vice Chancellor, University Relations and Advancement
Edward Sullivan, Assistant Vice Chancellor, Academic Research and Resources
Karen Y. Zamarripa, Assistant Vice Chancellor, Advocacy and State Relations
Ryan Storm, Assistant Vice Chancellor for Budget
California State University
Report on Unique Statewide Pupil Identifier

The following report is forwarded in compliance with Section 10803 of the Education Code, which states:

"10803. (a) The Chancellor's Office of the California Community Colleges, the University of California, and the California State University shall each establish a process by which colleges and universities within those systems issue, maintain, and report information using the unique statewide pupil identifiers specified in paragraph (3) of subdivision (e) of Section 60900. (b) Annually, on or before April 1, the Chancellor's Office of the California Community Colleges, the University of California, and the California State University shall provide a progress report to the Governor and the appropriate policy and fiscal committees of the Legislature. The report shall include a detailed timeline for the implementation, maintenance, and use of the unique statewide pupil identifiers pursuant to subdivision (a)."

This document is the annual report responding to this legislative request.

Background
The move to adopt a unique Statewide Student Identifier (SSID) paves the way for longitudinally-linked student-level data. The ability for the California State University (CSU) to more accurately track students over time will prove an invaluable tool for assessing student progress in relation to academic preparation, participation in special programs, as well as campus outreach efforts. In addition, the SSID will facilitate a more comprehensive assessment of graduation and retention rates as students move across segments, as well as employment trends (for graduates and non-graduates alike). In the absence of such an identifier (or a reliable SSN) the CSU has relied on the use of "fuzzy matches", which use a combination of information about a student to uniquely identify them.

History of Endeavors to Date
The California State University (CSU) system has 23 universities. Applications for admission to the CSU campus are handled online through CSUMentor (http://www.csumentor.edu/), although hardcopy and other forms of applications are available upon request. In October 2006, the CSU began to collect the Statewide Student Identifiers (SSIDs, also known as California School Information Services (CSIS) identifiers), the statewide public K-12 unique student identifiers. The SSID was added as an item included in the CSU common application form for undergraduate admissions. The CSU, thus, collects the self-reported SSID from the student, and this data are routinely loaded from our admission portal to the CSU Common Management System (CMS). Common practice is to collect and store the self-reported SSID at the campus level.
Unfortunately, there is no online process to validate and confirm self-reported SSIDs. CSU admission officers were told to call high school registration offices for a case-by-case verification. The CSU system processes nearly a half million applications for first-time freshman admission each fall—largely between November and February. Any SSID verification system used in an admission process would require the functionality to handle large batch matches under extreme time and budgetary constraints.

The CSU, the California Department of Education (COE), and the State Board of Education have collaborated on a unique partnership. Through the Early Assessment Program (EAP)—all end-of-year eleventh grade students are provided with the opportunity to get an early reading of their readiness for college mathematics and English. Short fifteen-item augmentations to a set of 11th grade California Standards Tests (CSTs), along with an eleventh grade essay.

Through this collaboration, students and their parents, high schools, and the CSU (and soon the CCC) receive information about the college readiness of participating students. As part of a larger review of the EAP, the CSU collected applicant data from CSUMentor’s fall 2007 admission processing, which included typical student identifiers, as well as student self-reported SSIDs. These data were merged with data from the spring 2006 administration of Statewide Testing and Reporting (STAR) system which included the spring 2006 CSTs and EAPs. The spring 2006 EAP data (extracted from the STAR data) which included typical student identifiers, as well as some district supplied SSIDs. The results were not promising. There were 45,298 fall 2007 CSU first-time freshmen from California public schools. Only 59 percent of these students were from districts that supplied SSIDs on the student’s test record. This suggests that SSIDs were not routinely used in many districts by spring 2006. Where there were self-reported SSIDs from CSUMentor and district-supplied SSIDs for students, there were no matches; the file structure supplied to the CSU research consultant from CSUMentor for fall 2007 had a 9 character SSID field, while the STAR SSID field had 10 characters. An inspection of the SSIDs fields for individuals who had STAR and CSUMentor SSID fields indicated no observable matches.

In the first annual SB 1298 report, the CSU was heartened to hear that utilizing existing student data at the CCC Chancellor’s Office and data from the COE, the CCC Chancellor’s Office's recent 2010 pilot study matched an average of 65% of the students in the Chancellor’s Office database to K-12 student records. This match allowed the CCC Chancellor’s Office to extract the CSIS/SSID number from the K-12 data and store the CSIS/SSID number in the CCC Chancellor’s Office database. The CSU concurred in the first report with the CCC assessment that moving forward, the matching algorithm might be refined to yield a higher match than any local implementation model which relies on student reported data. The CSU recognizes that it would be cost-effective and non-disruptive to regular CSU institutional operations to work with the CCC in its refinement—hopefully leading ultimately to a cost-effective statewide method for meet the needs of the state, the postsecondary segments, and K-12.
All parties were disappointed when funds for SB 1298 were not available CY 2010-11. The proposals to the federal government through its statewide database improvement program and through the Race to the Top program also failed to garner support. The final efforts to gain support from foundations was unsuccessful.

Given this reality, the principals from the CDE, the CCC, the CSU, the UC, and the Economic Development Department (EDD) who manage their data and analyses, along with the California Postsecondary Education Commission (CPEC), decided in 2010-11 to discuss whether each of the entities might be able to build a statewide data sharing methodology and agreement, as was described in the SB 1298 reports and in the federal proposals. All but CPEC brought data and expertise to the collaboration, and the Commission evidenced its technical capacity to collect, merge, and report on data from multiple postsecondary educational segments. During 2010-11, CPEC piloted ways to incorporate CSIS/SIS numbers with postsecondary data. In addition, pilot studies have been undertaken between the EDD and the CSU it was furthermore heartening to see that Senator Simitian continued his interest in the value to the State that SB 1298 offered. SB 885 (Simitian) Public education accountability: longitudinal educational system -- proposes much of what was outlined in SB 1298 and that which the aforementioned entities have been trying to accomplish. The principals developed a joint interagency data sharing agreement, vetted by counterpart General Counsel, for each to match and to merge files, as is envisioned in SB 1298, towards the end of fiscal year 2010-11.

CPEC was to be defunded, beginning in 2011-12. The CCCCO, the CSUCO, and the UCOP continued to work together with the CDE to ensure that the data holdings collected by CPEC were not destroyed, as is required by law under FERPA. Intersegmental general counsel and segmental researchers crafted agreements and communications with CPEC counterparts to return segmental FERPA-protected data holdings to each segment which had provided the data, but to leave such within the CPEC architecture. Much of 2011-12 has been spent on the CPEC data holdings, although some progress was made in building stronger understandings and relationships with the COE. At its most recent statewide data sharing meeting in mid-March 2012, it became clear that no "fuzzy matching" approach can enable all public K-12 students to be tracked to the CCC, the CSU, and the UC on the basis of SSIOs. Subsequent meetings of the intersegmental consortium have not produced enhanced results.

Self-reported SSIOs simply are not validated and, hence, have not proven to be useful. In order for the SSIO to be used at CCC, CSU, and UC institutions, they must be received at the institution as validated identifiers with additional validated identifiers (e.g., first name, last name, date of birth, high school codes). This can be accomplished most simply by having the SSID included on the official electronic high school transcript or a facsimile electronic high school transcript available through COE.
through the data sharing agreement. Alternatively, the CCCCO, the CSUCO, and the UCOP could request its campuses to provide electronically self-reported SSIDs along with other contact data that might be programmed to match against official SSIDs and contact data stored at the state level. In addition, the COE and others may wish to consider collecting and encrypting social security numbers (SSNs) which are required by the federal government, banks, for financial aid, for employment, for verification of postsecondary enrollment, for verification of postsecondary degree receipt. The CSU institutions and its system office who manage applications, admissions, enrollments, financial aid, and degree awards have never lost "personal identifications" nor caused problems with the loss of individual confidentiality and privacy.

**Detailed Timeline for the Implementation, Maintenance, and Use of the Unique Statewide Pupil Identifiers**

CSU student data systems are ready to incorporate the SSID but until there is an agreed plan for validated SSIDs, such is not in implementation plans. The CSU is participating in efforts to facilitate the sharing of student data among the three public segments of higher education (University of California, California State Universities, and California Community Colleges) and K-12 institutions (California Department of Education - COE). The goal is to directly match segmental data with COE student data, providing a way to validate self-reported SSIDs. Access to CDE student SSID data will also provide CSU with a means for deriving an SSID for those who did not provide this information elsewhere. Attaining this goal continues to be dependent on getting SSID included on all student high school transcripts, ideally electronic submission of all such transcripts (eTran), and provision of a comprehensive dataset of valid SSID's matched with student name and high school from CDE (so data received can be validated).