

Minutes of the SDLS Data Management Committee Meeting
First day of two-day onsite meeting in Coeur d'Alene, Idaho - July 25, 2013
submitted by Archie George

Meeting started about 1:05 p.m. (Pacific daylight time).

Attending: Andy Mehl, Georgeann Griffith, Ann Lewis, Scott Grothe, Vince Miller, Archie George, Tami Haft, Todd King, Doug Armstrong, Linda Clark, Vera McCrink.

1. There have been two data requests so far. The first was from the Lakeland School District. Andy led the group through the final "Data Sharing Agreement" which has now been signed (the DMC has seen drafts of this previously). Discussion focused on aspects of the agreement various committee members felt could be improved in subsequent similar agreements.
2. The second request was contained in an email, from Marilla Antenez, a librarian at Boise State University. She is seeking data on older, degree seeking students for a grant proposal. Andy and Scott transferred the information in the email into the Data Request Form. Many of the data elements have not been submitted to the SLDS by the institutions (e.g. tuition waivers). A fair amount of discussion regarding this request addressed the appropriateness of involving the DMC members in formulating the response, and whether to solicit information from the individual institutions (e.g. age based tuition discounts) that is not in the SLDS (or not yet submitted) .
3. We discussed the template for the Data Sharing Agreement, incorporating several concepts that were introduced during the discussion of the Lakeland SD agreement. Ultimately, the data request portal will contain questions, the answers to which be used, automatically, to complete most items in the Data Sharing Agreement form.
4. Andy led the group through the fourth item, the SBOE "Report Request, Data Access & Security Policy." There have been a number of changes, mostly minor grammatical in nature. Many editorial and a few substantive changes were suggested. Andy incorporated these into the master document.

Meeting adjourned about 4:35 p.m.

Minutes of the SDLS Data Management Committee Meeting
Second day of two-day onsite meeting in Coeur d'Alene, Idaho - July 26, 2013
submitted by Archie George

Meeting started about 8:05 a.m. (Pacific daylight time).

Attending: Andy Mehl, Georgeann Griffith, Ann Lewis, Scott Grothe, Vince Miller, Archie George, Tami Haft, Todd King, Doug Armstrong, Linda Clark, Vera McCrink.

Andy led the group through the Project Service Order form, a request for proposals to “design, implement, test, and support initial deployment” of the data request interface. Basically, the goal is to replicate the Virginia system office automated data request system. To date, the cost estimates have ranged from about \$1.2 million (from the Virginia system folks) to about \$350 thousand (from Blackfin consulting, who have done work for Andy in the past).

Andy then asked those in attendance to assist work with the group deliberating the content of a standard high school feedback report. Andy intends to reactivate this group in the near future. This led to a brief summary, by Scott Grothe, of the reorganization that has occurred over the past several months, precipitated by Jackie Thronguard’s departure.

The data collection schedule and an update, by Doug Armstrong, of progress on the quantity and quality of the SLDS data sets. Doug will be getting back to the institutions’ data providers with detailed feedback on their submissions. He reported that he has lots of data to work with at this time, and things are “looking good.”

Andy discussed the WICHE cooperative tracking project, with Washington, Oregon and Hawaii. There is a lot of excitement about this project, most of the current work involves the future of the project, additional participants, rather than specifics of the data analyses.

Doug discussed expanding the SLDS data collection to student financial aid information. He has a statewide, online meeting scheduled with participation of student financial aid administrators from many of the colleges and universities.

Andy led a discussion of the interface between the SLDS and the Department of Labor, concentrating on the various individual IDs (EDUID, SSN, Labor ID, etc.) and associated matching challenges. There was a short discussion of a desire, on the part of several DMC committee members, for materials regarding all the protections built into the system that prevents any one individual or entity to see all the various parts of the tracking system, in order to counter some of unfounded fears and rumors that are going around regarding the SLDS, its data elements and capabilities.

Todd led a discussion of a request, from Steve Hatchet (with the Department of Education), to add the mother’s first name to the EDUID process to resolve identity questions. Alternatives, such as city of birth, were discussed. Steve felt the mother’s first name would be commonly captured on the various student data systems and is not a common security question, and so would not raise concerns about identity theft. This conversation broadened to a discussion of the various problems and potential

solutions with “duplicate” EDUIDs. The DMC discussed the concept of notifying the schools and colleges that duplicates have been identified with instructions for incorporating the correct EDUIDs into their systems for subsequent data submissions. Todd indicated he has reached out to the institutions with this feedback on individual students, asking them to correct their data systems. To date, most if not all of this feedback has been provided to K-12 schools, not postsecondary. Attendee’s opinion, coming out of this discussion, was that EDUIDs should not be printed on transcripts until these issues have been resolved.

Scott asked Todd two big questions, 1) what is the process to resolve the duplicated once they have been identified and 2) why are there so many duplicates being created? If we only concentrate on fixing the duplicates, we might be missing the opportunity to improve the process so as to minimize the creation of duplicate EDUIDs. Todd discussed the “boot camp” training process, trying to insure that folks who are identifying matches or asking for new ones to be assigned are aware of the consequences and the help available for issues that come up. Scott wondered further whether there is evidence that the creation of duplicates is diminishing. Various data checks were discussed, e.g. names commonly associated with only one gender showing up with the other, but there is no easy way to identify all the duplicate EDUIDs.

Ultimately, the DMC committee does not object to adding the mother’s first name to the data elements collected from the institutions, for purpose of display on screens used for matching or assigning new EDUIDs. Doug will ask the student financial aid conferees, during the meeting next week, whether they collect the mother’s first name, for which students, and what the percentages of registered students likely have this information on the system. (Initial feedback, from the NIC student financial aid office, indicated the mother’s first name would not be collected except for a small number of students.)

Scott asked if it would be possible to develop a report or reports that would reveal the extent of data quality issues. For example, how many new EDUIDs are created over a specific time period (e.g. monthly), how many times a user is presented with a choice of two or more potential matches and then elects to create a new one. Can we count how many possible matches are displayed per potential match records submitted, what choices are made subsequently, etc. Issues with defining numerators and denominators for ratios that might correlate with data quality were debated. Even so, we know there are duplicate EDUIDs in the system, but we don’t know how many or what percentages of these exist, or whether there are proportionately more or fewer being created each year. Todd confirmed that there is no way to identify the number of duplicates without an institution reporting that they believe there are duplicate EDUIDs when they are attempting to assign or create them for a list of students. (Tami pointed out that now faculty and staff are also assigned EDUIDs which introduces some unique challenges.)

Tami suggested we need a process for notifying the state of name changes, for individuals who get married, are adopted, etc. Same goes for identified gender changes – or errors in data elements, especially those used for EDUID assignment, i.e. gender and birthdate. Todd indicated that to date, these changes are manual, Steve Hatchet receives notification from the school district, college or

university that an error was made or a change in information is needed, and the change is made manually.

Scott asked Todd who would be appropriate to meet with regarding defining some data quality indicators and reports that would be appropriate for monitoring these issues, especially the number of duplicate EDUIDs existing in the system.

Tami and Ann suggested they extend the EDUID assignment to non-credit students, which appears to be necessary in order to participate in the community college Virtual System of Accountability (VSA). NIC will take the initiative on this. This would be valuable for tracking workforce training participants into the labor force, which is necessary for compliance reporting.

Finally, Andy led a discussion of the agreement between the Idaho State Board of Education and the National Student Clearinghouse intended to determine how many Idaho residents, who have not attended an Idaho postsecondary institution, have completed Certificates. Certificates are not tracked in census records, thus the need for this additional data collection in order to fill a gap in tracking Idaho's progress toward the 60% goal of Idaho residents between 25 and 34 will have earned a certificate or degree by 2020. OSBE staff will be asking for feedback from the postsecondary institutions regarding opinions regarding the likely validity of these reports and inferences based on this approach.

A brief discussion of the data feedback reports Doug is providing ensued. The committee approved Doug's procedure for providing feedback to the institutions regarding data uploaded to OSBE for possible loading into the SLDS.

Future meetings were discussed, especially the next onsite meeting, likely in February. Likely dates would be near the end of February. SBOE meeting is February 26th and 27th, so the 24th and 25th might be a possibility.

Meeting adjourned about 11:50 a.m.