The meeting began with a presentation by Don Hossler and Doug Shapiro of the National Student Clearinghouse, which was intended to address questions participating state agencies have voiced about the role of the NSC in the exchange. Hossler and Shapiro began with an overview of where the NSC came from, its evolution from a loan clearinghouse to a membership organization, and what its potential in the project might be. NSC Research Center is a separate 501-C-3 from the NSC and has its own board of directors, which makes them an independent actor and should build credibility, but which requires trust and openness to be successful. Hossler and Shapiro acknowledged that being a membership organization creates a dilemma in that it puts constraints on how they are able to use the data for fear that institutional members will pull out. Members enable NSC to hold data on their behalf so NSC does not own the data.

Hossler and Shapiro described the advantages of using the NSC for this project as including:

- its large IT capacity with appropriate security protections and ability to receive longitudinal student records.
- associated legal expertise and familiarity with FERPA.
- lots of experience with matching data from both secondary and postsecondary sources.
- broad coverage of enrollment data (currently about 96% in Title IV, degree-granting institutions) for about ten years and a bit less for degree data.

Hossler and Shapiro described ways of making matches that they use without relying on the social security number—variables include name, birth date, HS diploma date, etc. Matching routines are proprietary to NSC. State agencies requested a brief document describing their algorithm sufficient to provide comfort with the algorithm but without giving away the store. NSC has focused largely on the match between secondary and postsecondary. They have checked the quality of the match compared to SSN match and got 95% of the original with multiple variables. They are finding between 10 and 18% differences between retention (within the same institution) and persistence (tracking students anywhere) rates for traditional-age students only. NSC is also seeking ways to collect additional data elements, including flag for degree-seeking (collected from institutions), major (by CIP code), class level (freshman, sophomore, grad), gender, race/ethnicity. Currently these are optional. As of Fall 2010, roughly a quarter of institutions provided these data, though coverage varies by element. Institutions are incentivized to provide these data because they get discounted services. Also they will be able to get that information back for the students who transferred away from their institution (i.e., major).

State agency representatives asked many questions as well, many of which emphasized the limitations of the data—e.g. no matriculation or first time data. But they are beginning to collect supplementary data that will eventually be useful. The research agenda is developed by
membership consultation and what the advisory committee wants to see done. USDOE reviews NSC’s security procedures, and they also have their own internal checks on this.

The Q&A session also touched on confusion about what the NSC’s role will be—data source or purely a matchmaker. For the matchmaker role, though, participating agencies need to be reassured about the integrity of the matching algorithm, what variables are included and how it actually works.

The bulk of the meeting following the NSC’s presentation addressed the MOA, next steps and other issues.

State agencies raised the question of how identifiable will the data be once it comes back from the matchmaking process and how much granularity will be there for disaggregation. State agencies were assured that the overall objective of the project is to test a functional data exchange that is a resource for participating states to use to address a wide range of research questions aimed at improving instruction and system-wide performance and meeting workforce needs.

The group subsequently walked through the draft MOA. While there was general consensus on the MOA, several activities are needed. Particular issues raised included a need to clarify that the MOA governs the grant project, but does not describe the permanent arrangement for exchanging data under the eventual governance structure. Oregon noted that the counts in the post-match data will probably not correspond to the original counts of the cohorts as chosen by the agency because of match fails, which may result in a credibility issue.

Working on nailing down a “data element dictionary” is clearly needed to align data element definitions and move the project forward. We agreed that it would be best to have everybody conform to the Gates-sponsored common data standards when these are available. We identified a subcommittee to work on the master list of the variables that everybody has as the foundation for building a DED. We should list all the elements that we would potentially need, even if the likelihood is low that all of them will eventually be used.

The MOA also needs a narrow specification of what the allowable uses of data are, as well as clarity around the “destroy the data when you are done” clause to ensure this also includes enhanced data that is now contained in each state’s data files.

Prescott next asked states to share their plans for the small funding amounts the project will be providing them.

- Hawaii will create data crosswalks and better definitions...also determine what data elements they don’t have that they need.
- Idaho has not decided yet, but will probably spend on labor costs.
- Oregon will use it for legal fees for review of MOAs etc.
- Washington will support a researcher to undertake ROI studies with the data they have.
Moving Ahead

What needs to happen to move this forward that we can do something about?

- SSNs for high school students
- Identity matching (independent of FERPA)
- Order of events and multiple attendances. What happens when and in what order?
- What data labor (UI) can actually share with whom (e.g., HI with WICHE issue)
- Consistency and data mapping
- Legal issues if FERPA does not get fixed
- What years for data will we need?
- What specific data will parties get back from the exchange? The common desire is for unit record data not just an aggregate report.
- Avoid the data being captured by accountability concerns
- Unintended consequences of data use outside the original scope of the project
- Use of SLDS data—limits on this?
- When does this cease to be a “project” (temporary, working within a constrained space)
- How to make the MOA consistent with what the feds require in exchange for SLDS funding in Oregon and Washington (i.e., number of years of data retained).

What needs to happen to make this a reality?

- Bring the legal folks into the conversation at this point
- Form a data element working group
- Finalize what the role of NSC should be

What will each state do in the near term?

- Revise the MOA based on this meeting’s input
- Take the revised MOA to the state’s legal people
- Plan meetings that coordinate SLDS planning and data exchange planning

Final Take Aways & Principles for Moving Forward

L’Orange asked if it was fair to say that the substance of the conversation has shifted from “IF” (we should undertake to exchange data among states) to “HOW” we will do so. The draft MOA is in a condition sufficient for it to be vetted through relevant attorneys general, but more specificity is needed for what data elements will actually be shared in the initial phase.

Activities undertaken by states in the project should be broadly generalizable. That is, we should (for the most part) not “jury rig” solutions to particular operational problems that are specific to particular states and would not work in other states.

The primary objective of the project is to demonstrate how a powerful, multi-use, shared data resource can be built. Certainly one objective is to answer some particular research questions
about student flow, but the project should avoid "solutions" that narrowly answer the research question without demonstrating how a more capable and general data resource can be built.

No decision should be made in the project that precludes a more inclusive eventual scope for the database.