# WICHE

**Western Interstate Commission for Higher Education**

ALASKA

ARIZONA

CALIFORNIA

COLORADO

HAWAII

IDAHO

MONTANA

NEVADA

NEW MEXICO

NORTH DAKOTA

OREGON

SOUTH DAKOTA

UTAH

WASHINGTON

WYOMING

3035 Center Green Drive • Suite 200
Boulder, Colorado 80301-2204

T 303.541.0200   F 303.541. 0291

**www.wiche.edu**

# Western Policy Exchanges

**In December 2008, with funding from the Bill and Melinda Gates Foundation and in partnership with the Data Quality Campaign, the Western Interstate Commission for Higher Education (WICHE) hosted a meeting of the West's key leaders in building state-wide integrated longitudinal data systems. Such data systems are essential to developing a better understanding of the mobility of individuals through K-12, postsecondary education settings, and into the workforce.**

**Many states are engaged in or considering ways to build or enhance their education databases, but fewer are including workforce information in that effort. While there are many reasons for this, one of the biggest impediments is an incomplete or inaccurate understanding of legal privacy restrictions. Thus, a central purpose of the meeting was to address how states and state agencies can share data while remaining fully compliant with the Family Educational Rights and Privacy Act (FERPA).**

## Fostering Collaborative State-Level Education and Workforce Database Development

Concerns about accountability in education have grown markedly in recent years. Unfortunately, many of the most important accountability questions cannot be adequately addressed by current data systems that rely principally on cross-sectional data. These "snapshots" can provide only a limited assessment of existing state performance in education because they do not fully capture variation in performance among students or groups of students, nor can they account for student mobility. There is a growing need to construct longitudinal data systems that track individuals through their K-12 and postsecondary education experiences, as well as into the workforce, in order to better evaluate how well state investments in education are fulfilling state workforce and economic development needs and to provide guidance for policy. To do this, states must foster collaborations among the responsible sectors and agencies within their state and among neighboring states. Given the mobility of the population, the most effective systems would also incorporate data and information across state boundaries.

State and federal policymakers do not have an adequate understanding of the degree to which individuals move into and out of education systems and, from there, into the workforce. Most studies to date have relied on aggregate analyses that imperfectly link the supply of educational programs to occupational demands in the economy. What's more, researchers are frequently forced to make major assumptions about the flow of human

capital, with an unknown impact on their results. This is not a criticism of such efforts, so much as it is a reflection of the very real limitations of the available data sources.

Accountability demands, along with an environment in which higher education has become the nation's top economic development priority, have made it vital to link all levels of education to workforce data. The capacity to build a better understanding of human capital flows, is growing in importance due to two factors: the global economy and a rapidly diversifying population.

To help its member states address issues surrounding their efforts to construct longitudinal data systems or enhance the capabilities of existing ones, WICHE, in partnership with the Data Quality Campaign (DQC, for more information see their website at www.dataqualitycampaign.org) and with support from the Bill and Melinda Gates Foundation, convened a meeting in Boulder, Colorado in December 2008. Invited to the meeting were officials responsible for the K-12, postsecondary, and workforce data systems from each of WICHE's member states (representatives from 14 states attended). The meeting was intended to spark conversations helping to dispel myths about the legal prohibitions on data sharing, consider ways to surmount political barriers, provide technical assistance, and catalyze further development of longitudinal data systems.

The gathering got underway with a dinner presentation by David Longanecker, WICHE's president, who emphasized the growing need for analyses that rely on longitudinal data, the use of which could lead to more effective policy development. He also cautioned that, too often, data systems are built with great attention to technical details but insufficient attention to the policy needs of the state. Data systems must be built with a clear sense of how they might be used and what questions they could answer. Absent that, data can be misused, often with nonsensical results: to illustrate this point, he exhibited the photo of the welcome sign posted by a small town in the foothills of the Rocky Mountains (Figure 1).

## A National Overview: Current Efforts

The "educational pipeline" is a compelling metaphor for conceptualizing the development of human capital, and policy discussions of how to increase flow through it are now common in every state across the nation. Such discussions are fueled in part by more vocal participation on the part of business and civic

leaders, who have increasingly come to recognize the importance of the "supply chain" of educational capital into their companies and their states.

To strengthen that supply chain, and achieve the widely held policy goal of increasing the number of citizens flowing through the educational pipeline and attaining the postsecondary credentials needed for individual prosperity and national economic competitiveness, many states are working to develop integrated student unit record (SUR) databases. Only SUR data can provide a comprehensive understanding of how students move through (and, in too many cases, "leak" out of) the educational pipeline and into the workforce is needed. While postsecondary and K-12 SUR databases have been in existence for over 20 years, only recently have states begun to align their component data elements and share data across agencies. The meeting began by providing an overview of the development work states are currently doing and the progress they have made.

### K-12 Student Databases

Aimee Guidera, executive director of the DQC, noted that although each state's P-12 education system is unique, there is a set of 10 essential elements that are critical to any longitudinal data system:

1.  A unique statewide student identifier that connects student data across key databases and across years.
2.  Student-level enrollment, demographic, and program participation information.

## Figure 1. Data Versus Information

3. The ability to match individual students' test records from year to year to measure academic growth.
4. Information on untested students and the reasons they were not tested.
5. A teacher identifier system with the ability to match teachers to students.
6. Student-level transcript information, including information on courses completed and grades earned.
7. Student-level college readiness test scores.
8. Student-level graduation and dropout data.
9. The ability to match student records between the P-12 and higher education systems.
10. A state data audit system assessing data quality, validity, and reliability.

Since its inception in 2005, the DQC has documented considerable progress in how many of these elements states have incorporated into their data systems. Over the next three years, the DQC is determined to help develop longitudinal data systems that can follow individuals through the P-20 pipeline while ensuring broad yet appropriate access to these data.

### *Postsecondary Databases*

A 2007 paper by Peter Ewell and Marianne Boeke found that 42 of the 50 states have operational SUR databases covering their public postsecondary institutions, which together account for about 81 percent of the nation's total headcount enrollment. These databases share a number of important features: use of multiple databases, institutional coverage, historical data, data detail and periodicity, record identification, and the ability to link to other databases. (For more detailed information or to read the entire report, please visit http://www.nchems.org/c2sp/sur/)

Many states are enhancing their SUR data systems by adding new data elements or additional data capture points, and a significant number of states are finding ways to link their data with external databases. The only substantial area in which little progress has yet been made is in sharing data across state lines.

### The Regional Perspective: Where Are We Now?

In the WICHE region, there is considerable variation across SURs with respect to the data elements collected, reports generated, and collaboration efforts among states and stakeholder agencies. Idaho has no SUR database, while Oregon, Washington,

and Wyoming each have two such databases, and California has three. The majority of WICHE states collect transcript-level data. However, only Nevada and New Mexico are currently linking data from their postsecondary database with information from high schools and employment. The community college systems of Oregon and Washington are linking with other databases, but the university systems are not. Figure 2 provides a visual overview of the SUR capacity of the WICHE states. The five core data elements are: sex, race/ethnicity, date of birth, program/major, and degree awarded.

### FERPA Update

Among the largest obstacles to the development of longitudinal data systems are perceived legal prohibitions, especially those based on faulty interpretations of the Family Educational Rights and Privacy Act (FERPA). A lack of common understanding of what is permissible under FERPA, a fear of noncompliance, and widely varying opinions of state attorneys general regarding FERPA have together created a significant barrier to state-level sharing of student data for important and legitimate educational purposes. In addition, FERPA – and the way that it has been administered by the United States Department of Education – has had a significant and chilling effect on the development and implementation of robust longitudinal data systems.

Steve Winnick, senior counsel for EducationCounsel – a law, policy, strategy, and advocacy organization that focuses on education – presented a review of the new United States Department of Education–amended FERPA regulations, dated December 9, 2008, and focused on how these regulations affect state longitudinal data systems (SLDSs). The regulations (or language in their preamble):

▶ Authorize states to re-disclose education records to recipients and for purposes covered in FERPA-authorized disclosures.

▶ Include provisions regarding recordation of re-disclosures that facilitate a state's ability to make re-disclosures.

▶ Permit sharing of data between P-12 and postsecondary data systems, but include confusing preamble language that may be read to limit such sharing.

▶ Provide what appears to be a narrow window for state disclosure of education records to research organizations for studies to improve instruction (or to develop or validate assessments or administer student aid programs).

## Figure 2.  Capacity of the WICHE States' Student Unit Databases in 2006

| | Link with High Schools | Link with Employment | 5 Core Elements | Transcript-level Data |
|---|---|---|---|---|
| **Alaska** | | | ✓ | ✓ |
| **Arizona** | ✓ | | ✓ | ✓ |
| **California** | | | | |
|   CC system | ✓ | ✓ | ✓ | ✓ |
|   UC system | | | ✓ | |
|   CSU system | ✓ | | ✓ | |
| **Colorado** | | | ✓ | |
| **Hawaii** | ✓ | | ✓ | ✓ |
| **Montana** | ✓ | | ✓ | |
| **Nevada** | ✓ | ✓ | ✓ | ✓ |
| **New Mexico** | ✓ | ✓ | ✓ | |
| **North Dakota** | | ✓ | ✓ | ✓ |
| **Oregon** | | | | |
|   CC system | ✓ | ✓ | ✓ | |
|   Univ. system | | ✓ | ✓ | ✓ |
| **South Dakota** | ✓ | | ✓ | ✓ |
| **Utah** | | | ✓ | ✓ |
| **Washington** | | | | |
|   CC system | ✓ | ✓ | ✓ | ✓ |
|   HECB** | | | ✓ | |
| **Wyoming** | | | | |
|   University | ✓ | | ✓ | ✓ |
|   CC system | | | ✓ | ✓ |

\* Idaho has no known SUR database.

\*\* Higher Education Coordinating Board

*Source:*  Peter Ewell and Marianne Boeke, *Critical Connections: Linking States' Unit Record Systems to Track Student Progress* (Indianapolis, IN: Lumina Foundation for Education, 2007).

> ▶ Appear to rule out disclosures of individual education records to a student's former school for evaluation or accountability purposes.
> ▶ Include flexible guidance on de-identified data.

Winnick indicated that states that wish to develop or sustain robust SLDSs should review state laws and policies, in light of the new FERPA regulations, to strengthen the legal foundations for appropriate sharing of data for evaluation and research purposes.

States also need to review their own privacy laws to ensure they do not create additional constraints beyond FERPA. States should consider enacting laws or regulations, consistent with state policy, to do the following:

> ▶ Authorize SLDSs to enter agreements for research studies to improve instruction for or on behalf of postsecondary institutions, schools, and local educational agencies.
> ▶ In the case of states with separate P-12 and postsecondary data systems, authorize each system to receive education records from the other system for purposes of evaluating, auditing, or ensuring compliance with the requirements of state and federal education programs.
> ▶ Authorize postsecondary institutions and data systems to disclose education records to a student's former school district or secondary school for the purpose of evaluating the education programs of that school district or school. (Winnick indicated that, despite negative language in the regulation preamble on such disclosures – language that is not legally binding – there is a very good legal argument that such disclosures are consistent with FERPA, and enactment of such authority in state law or regulations would strengthen that argument.)

Other steps that states should consider include: developing procedures for recording re-disclosures and transmitting the records upon request to postsecondary institutions, schools, and local educational agencies; developing appropriate standards and a process to determine whether data are de-identified and to code data for research purposes, as needed; and reviewing and, as appropriate, strengthening administrative and electronic safeguards to protect education records against improper disclosures.

One question that came up repeatedly during the meeting was how FERPA might affect the linking of educational data (K-12 and postsecondary) with workforce data. Winnick explained that states that wish to merge educational and workforce data should take steps to ensure consistency with FERPA. Under the FERPA statute, education records may not be disclosed (without written consent by a parent or by a student who is 18 or above or enrolled in postsecondary education) to a workforce agency for the purpose of evaluating or strengthening non-education workforce programs. An amendment to the FERPA statute is needed to permit these disclosures. However, to address these purposes, the workforce agency may

provide records on the individuals it serves to the SLDS to permit the SLDS to link the workforce data with education records. The SLDS can then report aggregate or de-identified data derived from students' education records back to the workforce agency.

Other practical steps to effect data matching between education and workforce agencies were discussed, including having a SLDS employee or contractor supervise the match or detailing a SLDS employee to the workforce agency for this purpose. In addition, Winnick discussed two areas where state law provisions might expand flexibility for at least a limited sharing of workforce and education data under current law:

▶ Designate the workforce agency as an authorized representative of the state education agency for the purpose of receiving education records to evaluate/audit education programs, or provide those records to a contractor of the education agency or workforce agency for that purpose.

▶ For purposes of evaluating education programs, define "education" broadly to include job training programs.

Winnick added that federal law sanctions and supports state longitudinal data systems which are intended to facilitate more effective use of data for improving education and meeting the academic needs of students. Through such systems, states, educators, and researchers should be able to achieve these purposes without violating FERPA. For this to happen, FERPA needs to be interpreted or, as necessary, amended to harmonize these state and federal policies. It should not be unnecessarily interpreted in a rigid manner that in effect thwarts the legitimate use of student data by education agencies to improve education.

Peter Ewell of the National Center for Higher Education Management Systems (NCHEMS) commented that FERPA is primarily concerned with an individual record being abused; but states rarely look at individuals. They are instead using data sets in the aggregate to inform policy decisions. Given this, states and policymakers need to explain more concretely how they use data to alleviate the fear of misuse.

In sum the goal of FERPA is to balance privacy with legitimate research. In many ways FERPA is not the real problem obstructing the building and use of state databases. Rather, the problem relates to resistance to change, system security, relationships, and resources, as well as political factors.

## Building and Using a Robust System

Too often, people just want to get on with building a data system without thinking concretely about the particular research and policy questions that such a system might answer. Up-front decisions also need to be made about whether the database will be housed in one system or several. According to Ewell and Hans L'Orange from the State Higher Education Executive Officers (SHEEO), there are a number of fundamental questions that a state must answer before embarking on this journey. As Ewell noted, "States must begin with determining the needs and interests of current and potential stakeholders. This includes careful consideration of the kinds of questions that potential users need to answer."

Furthermore, the system should be developed under legislative authority that authorizes the links between the various database systems but does not specify exactly how these links should be forged. The system also requires secure, unique identifiers and aligned data element definitions and code structures. Finally, a robust system will have accessible, comprehensive, and up-to-date documentation. Ewell added that although a data system needs basic core elements, it should be flexible enough to add new elements at a later date.
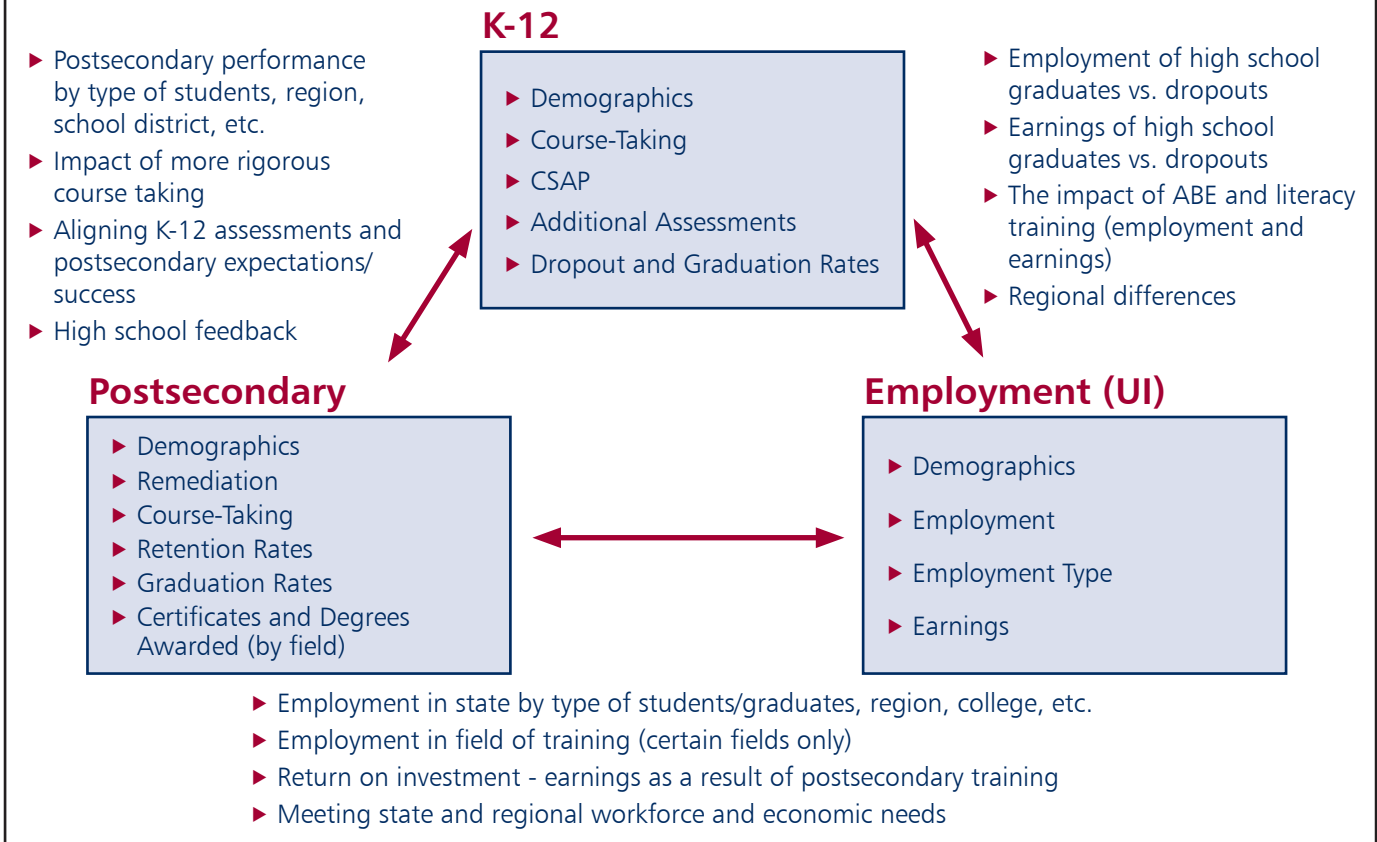
### *Process of Developing a System*

The ability to match student records with data on K-12 educational activities and with data on employment is critical. This does not necessitate a single data system containing secondary, postsecondary, and workforce records, but it does require that the systems be able to integrate with each other. Technical interoperability agreements between aligned but separate systems are a critical component. Likewise, the "ownership" of the data needs to be established in order to facilitate integration.

### *Elements of a Robust System*

Ewell and L'Orange presented a model of the elements required to create an integrated education and workforce data system (Figure 3). States must consider which elements will yield the best information and the most useful data and which analyses will help the state create or modify policy to improve education and workforce outcomes. They also need to identify and eliminate policies that are barriers to innovation in the education and workforce sectors.

In addition, Ewell reminded the audience, "We can't forget about mobility – 50 great databases are wonderful, but we need to cross state lines." The

## Figure 3. Connected. The Additional Information Available for Policy Decisions

**K-12**

- Postsecondary performance by type of students, region, school district, etc.
- Impact of more rigorous course taking
- Aligning K-12 assessments and postsecondary expectations/success
- High school feedback

K-12 box:
- Demographics
- Course-Taking
- CSAP
- Additional Assessments
- Dropout and Graduation Rates

- Employment of high school graduates vs. dropouts
- Earnings of high school graduates vs. dropouts
- The impact of ABE and literacy training (employment and earnings)
- Regional differences

**Postsecondary**

- Demographics
- Remediation
- Course-Taking
- Retention Rates
- Graduation Rates
- Certificates and Degrees Awarded (by field)

**Employment (UI)**

- Demographics
- Employment
- Employment Type
- Earnings

- Employment in state by type of students/graduates, region, college, etc.
- Employment in field of training (certain fields only)
- Return on investment - earnings as a result of postsecondary training
- Meeting state and regional workforce and economic needs

reality is that students move from state to state, not only within the K-12 arena but for postsecondary purposes and for workforce opportunities as well. The impetus to build robust systems that transcend K-12, postsecondary, and workforce sector boundaries must come from the states themselves. But to do this, states and agencies must work together to ensure that all the players understand the need for such a database system and the multitude of state, regional, and national benefits that it can bring.

### Examples from Selected States

Florida is a state that has used education and workforce data together in effective ways. It has an integrated, longitudinal PK-20 education data system, which includes student-level data for public schools, community colleges, career and technical education, adult education, and the state university system. Also included are postschool employment and non-education system program data.

Washington is another state that has used education and workforce data to answer pertinent policy questions: one example is the 1996 Dislocated Worker Analysis, which asked the question "Should laid-off

workers just go back to work, or does it pay to train for new jobs?" The research found that workers who took classes or training in STEM (science, technology, engineering, or mathematics) had less of a decrease in wages then those who either had no training or training in non-STEM courses. Based on this research, the state has continued to fund training for dislocated workers. Additionally, as part of its Tipping Point Project, Washington found that one year of college credit plus a credential is the "tipping point" for students needing to find career pathways. New policies have increased funding to certain grants and scholarships in Washington to assist low-income students get to the tipping point and beyond.

Florida and Washington have several fundamental factors in common with regard to their databases that account for their success. Both states stress longitudinal data and "just-in-time" reporting for answering key questions related to state, higher education, and workforce development issues. That reporting informs legislative and other policy decisions. Washington's success with the Tipping Point Project and the Dislocated Worker Analysis answered strategic questions for the state using data resources.

Jeff Sellers, assistant deputy commissioner at the Florida Department of Education, noted that "research using these data can track students through work and education to determine where the leaks in the pipeline are, what can be a predictor of educational success, and what can be a predictor of work success." He added that "the linkage between employment opportunity, economic prosperity, and the educational system – particularly postsecondary education – is critical to the nation's future." This discussion concluded with five important points that will aid states in building, using, and maintaining a robust longitudinal data system:

- ▶ Make the data meaningful.
- ▶ Emphasize security.
- ▶ Use data appropriately.
- ▶ Remember that in some cases, federal reporting drives data collection.
- ▶ Keep it (data collection and data reporting) simple.

## Discussion Themes

Policymakers must have data to support decisions that will ensure students leave high school prepared for college and work. To meet this need, robust longitudinal data systems must be in place. Discussions throughout the meeting identified some common issues that must be addressed in the process of creating, maintaining, and using such systems.

### Record Identification

Several states indicated that record identification was an issue. In some states, the problem is with collecting social security numbers (SSN) in addition to creating a unique identifier. For example, K-12 systems in some states do not collect the SSNs at all. Other states have challenges around creating a key link without using the SSN.

### FERPA Compliance

One of the biggest issues discussed at the meeting is how states can comply with FERPA and still share data across various departments. Many states have been able to link higher education data with K-12 data in limited ways but have been unwilling to link workforce data, based on FERPA interpretations. These inconsistent approaches to FERPA compliance cause contradictory policies, frustration, and confusion.

### Governance Structure

In some states a major problem is the way higher education record keeping is organized. California, for example, has three postsecondary student unit-record databases, corresponding to its major sectors of higher education, and they tend to work within their own silos. This structure leads to ongoing challenges with data sharing and privacy within postsecondary education, even before an attempt is made to link in data from the K-12 system and workforce information. Thus, the existing structure of data governance will help or hinder the construction of robust data systems and their effective usage.

### Communication and Trust

Representatives from New Mexico noted the importance of continuously educating people about their state data system and what it can do. State database leaders need to take the time to work with their own stakeholder communities about the data they collect and maintain and discuss how their data can be shared for the benefit of improved policy and practice. Several participants echoed the fact that communication at multiple levels is critical for success and that communication and trust must also be constantly cultivated.

### Resources

Designing, building, and maintaining a sufficiently robust longitudinal data system is not an inexpensive proposition. At the meeting conversations inevitably turned to the challenge that a lack of adequate funding presents. Given the current economic status of our country and proposed budget cuts, states may have to build and maintain data systems with even less money. This emphasizes the importance of working together (across states and within states) in order to keep momentum up. The people needed to maintain and analyze data are frequently the first to be cut when budgets are tight. This is a mistake because data-informed decisions are most needed during difficult periods. Fortunately, the federal stimulus package contains a substantial amount of funding for state longitudinal data systems. But states will need to be thoughtful about their design so that ongoing maintenance costs are affordable after the stimulus money is gone.

### Legislation

According to Winnick, federal law supports robust state data systems, but FERPA has

been interpreted and understood in a manner that impairs these systems. With the issuance of new FERPA rules, there is an opportunity for states to issue policies in the form of state laws or regulations that strengthen their data systems and the legal support for their proposed uses of data, consistent with FERPA. As states continue to move forward in expanding their data systems, they should consider whether changes in state law are needed to ensure that data administrators have the authority to make necessary disclosures on behalf of higher education, K-12, and local agencies that are needed to permit effective use of data for essential research and evaluation purposes.

### Start small

One key to successful implementation is to get the database up and running so that core data elements and basic reporting are in place quickly, leading to tangible and useful analyses and reports. If a state can produce even a few thoughtful reports for key policymakers, they will want more, as the experiences in Florida and Washington show. The Arizona delegation added that building on common goals (such as reporting tied to specific legislators' districts) is another way to begin the process of entrenching a data system in the state's political culture.

### Keep it simple

It makes little sense to expend political capital, resources, and effort it to build data systems that turn out to be of limited use in informing public policy discussions or in creating more effective accountability systems. States indicated that setting down a select set of core data elements that address the most important and basic questions, while making sure that the system design allows for expansion and flexibility, is a good way to start. States trying to construct "the perfect system" are likely to get bogged down in design or implementation weeds before the promise of more effective information can take root.

## Concluding Thoughts

Based on breakout sessions, ongoing discussions, and the general comments and questions from the presentation, there are a number of ways in which states working with state longitudinal data systems could use ongoing assistance. These include:

▶ Continuing to make connections among states and state players.
▶ Producing a document on "rules of thumb" to follow in designing data exchanges based on collective experience.
▶ Developing the architecture for a common "data hub" to integrate data from the three types of agencies that multiple states could use.
▶ Developing aligned definitions and data structures for a limited number of "common core" data elements.
▶ Working with the federal government to secure funding to help meet the resource needs of SLDS projects.

At the conclusion of the meeting, WICHE's Longanecker, DQC's Guidera, NCHEM's Ewell, and EducationCounsel's Winnick provided some last-minute reminders:

▶ Don't think about this as "database development" but as producing useful information you can put in front of people, and add on from there. As one participant aptly stated, "It is all about using data to tell the story."
▶ In many states there has been a shift in mentality. Producing a fully integrated longitudinal data system is no longer an accounting or IT decision. It needs to be a policy-driven venture.
▶ Each state is different and will thus approach the task of developing a data system differently. However, there may be a better way to approach this task than dealing with one state at a time. For example, a central hub administered by an independent agency that is not under the direction of K-12, postsecondary, or workforce could house the data system. Furthermore, such a hub could service multiple states.
▶ Ultimately, successful longitudinal database development and effective usage is all about people. This is true of both problems and success stories.

Momentum is building for the kind of analyses that can only be accomplished with longitudinal data. Such analyses are vital if states are to remain competitive in a global economy that depends on human capital. WICHE and its partners remain committed to helping states move forward on the path toward robust longitudinal data systems that link information from K-12 and postsecondary education with information about the workforce.