

# LEVERAGING THE POWER OF STATE LONGITUDINAL DATA SYSTEMS

## Building Capacity To Turn Data into Useful Information

May 2011

### Key Messages

- Transforming the robust data housed in current state longitudinal data systems (SLDSs) into useful information requires very different skill sets from those that have been employed to build the data systems over the past few years.
- To effectively use the data in the SLDSs, state education agencies face three key human capacity challenges:
  - Identifying the types of analyses required,
  - Matching skill sets to the required analyses, and
  - Acquiring the appropriate skills.

Faced with the need to create a competitive workforce and improve the quality of our education system, states are pursuing policy agendas to better prepare students for post-secondary education and careers. To inform these agendas, statewide longitudinal data systems (SLDSs) have become an invaluable asset and serve as a tool to help states answer pressing policy questions and improve practice. While states have made impressive progress in building SLDSs over the past 10 years, most activities have focused on the systems' information technology (IT) aspects, such as hardware, software, data warehouses and online portals. Many states have developed reports that can be shared with districts and teachers, but those reports are often designed to answer questions that could be answered without longitudinal data — for example, the number of graduates in a given year or the percentage of 5th graders who passed the mathematics test.

As states continue to build and refine these new SLDSs, the education landscape continues to change. Initially, states were focused on meeting federal reporting requirements associated with the No Child Left Behind Act. The goals for these systems, built to collect longitudinal data that allow analyses over time and across program areas, have since shifted from compliance reporting to informing educational policy and practice as a means of improving student achievement. Now SLDSs can provide much richer information than educators and policymakers are used to receiving, giving states the opportunity to analyze student progress across

school years and predict future performance; evaluate connections among outcomes and classroom experiences; and use better data to help inform interventions, classroom and school practices, and district and state policies.

In the long term, these robust data systems can enable state education agencies to create tools such as early warning systems and data dashboards, which equip educators with information to address issues including dropouts, teacher quality and college readiness. All stakeholders will benefit in a variety of ways:

- **Students and their parents** will be able to assess whether the student is on track to graduate college and career ready.
- **Teachers** will be able to determine which students in their classroom need additional support and whether their students will leave their classroom prepared.
- **School administrators** will be able to inform the school improvement planning process and review early warning indicators to determine how to intervene with individual students in their school.
- **District administrators and school boards** will be able to review trends over time (by school, grade, subject and combinations of each) to identify which programs work with which schools and/or groups of students.
- **State policymakers** will be able to review impact analyses of potential new policies before passing legislation.

However, fully harnessing the power of these robust SLDSs to answer key policy questions and provide information critical to various stakeholders depends upon state education agencies' being able to:

- Collect the data in a way that provides valid answers to a state's critical policy questions, even as these questions change over time;
- Define accurate measures of progress;

- Conduct appropriate analyses of the data;
- Present and disseminate findings to stakeholders; and
- Use the information to drive policy and practice decisions.

While IT staff and resources will always be required to maintain and update the SLDSs, the power of these systems will not be evident until education analysts and researchers also engage in the full scope of system design, maintenance and use over time.

## Findings from the Data Quality Campaign's 2010 *Data for Action*

Each year, the Data Quality Campaign (DQC) conducts an analysis of all 50 states, the District of Columbia and Puerto Rico to assess state progress on implementing the 10 Essential Elements of a Longitudinal Data System and 10 State Actions for Effective Data Use. In 2010, states reported the following information:

- **ACTION 5: IMPLEMENT ROLE-BASED ACCESS** — eight states report that they are implementing data systems to provide all stakeholders with timely access to the information they need while protecting student privacy.
- **ACTION 6: CREATE REPORTS USING STUDENT-LEVEL LONGITUDINAL DATA** — 23 states report that they are creating progress reports with individual student longitudinal data that provide information educators, parents and students can use to improve student performance.

- **ACTION 7: CREATE REPORTS USING AGGREGATE-LEVEL LONGITUDINAL DATA** — 27 states report that they are creating reports that include longitudinal statistics on school systems and groups of students to guide school-, district- and state-level improvement efforts.
- **ACTION 8: DEVELOP A P-20/WORKFORCE RESEARCH AGENDA** — 28 states report that they are developing a purposeful research agenda and collaborating with universities, researchers and intermediary groups to explore the data for useful information.

Does your state have the human capacity and skills required to develop these longitudinal reports and to create and implement a thoughtful research agenda? Visit [www.DataQualityCampaign.org/stateanalysis/actions](http://www.DataQualityCampaign.org/stateanalysis/actions) to see all of the findings from DQC's *Data for Action*.

## Turning Data into Useful Information

To effectively use the data in the SLDSs, state education agencies face three key human capacity challenges:

- Identifying the variety of ways available data can be used and the types of analyses required to answer critical policy questions;
- Ensuring access to staff and/or consultants who have the appropriate skill sets and tools to conduct the required analyses; and
- Determining the extent to which they can or should rely on internal staff versus external partnerships to conduct the analyses.

### Identifying the Types of Analyses Required

Typically, when it comes to talking about analyzing data in education, people use "research" and "evaluation" as shorthand, but both of these terms conjure up certain images and emotions. "Research" often brings up images of long-term studies carried out by university researchers in which students are assigned to control and experimental groups. "Evaluation" typically brings up thoughts of accountability and negative ratings. Both of these images are true to a degree, but they do not convey the full spectrum of how SLDSs can be used to inform policy and practice. Other terms and phrases (such as "data mining," "business intelligence" and "analytics") are becoming more common. Regardless of terminology, all end users benefit

*Being clear about the purpose of the analyses will enable state education agencies to ensure that they have the staff with the most appropriate skill sets to guide the work.*

- **Analytics** — deals with modeling or analyzing data specifically to highlight useful information, address specific questions, suggest conclusions and support decisionmaking.
- **Business intelligence** — concentrates on analysis of aggregate data (classroom, school or district summaries) and business information, such as statistics about school graduates, dropouts and assessment performance.
- **Data mining** — focuses on predictive modeling, such as early warning indicator systems used to identify students at risk of dropping out.
- **Evaluation** — looks at the impact of a specific program or intervention in relation to specific outcomes, such as a teacher’s value-added to student achievement or the effectiveness of a reading or dropout intervention program.
- **Research** — references either academic or applied research that is conducted in a rigorous and systematic way, attempts to solve a problem with either new data or use of existing data for a new purpose, and is based on observable experience or empirical evidence.

Data in the SLDS can be used for analytics, business intelligence and data mining, without requiring specific or stringent research methodology. However, states also need to have the skills and capacity to carry out evaluation and research activities. All five activities will be necessary at different times for different purposes. Being clear about the purpose of the analyses will enable state education agencies to ensure that they have the staff with the most appropriate skill sets to guide the work.

### **Matching Skill Sets to the Required Analyses**

SLDS activities, including analytics, are often housed in IT departments because historically, by default, everything data-related goes to IT; however, state education agencies cannot assume that IT programming staff have the skills to perform the

in a variety of ways from having information from the more in-depth analyses that are possible with the current SDLSS. However, defining a common language will improve communication and expectations. What does each of these terms mean?

necessary analytics and data mining. Many chief information officers who are responsible for building and maintaining the SLDSs have for years made it clear that they and their staff are not “data” people. They can develop and program collection tools and reporting applications, but they generally do not have the statistical and research backgrounds required to ensure that the right data are analyzed and that the resulting analyses are valid.

To mitigate this problem, program area staff (e.g., special education, English as a Second Language) needs to be included in the design, collection and analysis phases of the SLDS to ensure proper data definition. Quite often both program area and IT staff then require additional involvement of research and evaluation staff to ensure appropriate analyses and interpretation are applied to the development of analytical, business intelligence and data mining activities. The education sector has historically underestimated the value that both data and analysts/researchers bring to the field.

To ensure that appropriate and valid analyses are conducted and interpretations of findings are applied, state education agencies need to make certain that personnel trained in statistical analyses are engaged to design, conduct and oversee the work. For example, **Louisiana** contracted with a researcher at Louisiana State University to design and oversee its evaluation of teacher preparation programs and has since hired this person as executive director for strategic research and analysis. **Virginia** has also hired an executive director for research and strategic planning, who provides support and advice to the superintendent of public instruction, the executive leadership of the department and Virginia Department of Education staff. The Virginia executive director also coordinates and oversees scientifically based research and analyses conducted by the department and in cooperation with partner agencies and organizations; provides internal consultation services to state Department of Education staff on issues pertaining to research, analysis and program evaluation; and supports the department’s strategic and operational process-improvement initiatives.

*Given the complexity of analyzing longitudinal data, the full power of the SLDSs cannot be realized without engaging education analysts and researchers as part of a robust state education agency team to turn the data into useful information.*

# Are You Getting the Answers You Need?

Statewide longitudinal data systems (SLDSs) now more than ever allow for robust answers to critical policy questions within states, and state policymakers are continuing to ask hard questions. However, while de-identified student-level data collected at the state level provide the necessary information, state education agencies (SEAs) often do not have the staff capacity to ask the related data questions, perform the analyses and provide the rich answers policymakers desire. By increasing research capacity within the SEA and through external partnerships, states will be better able to fully analyze the data within their SLDSs. Below is an example comparing typical SEA capacity in answering a state policy question to that of staff with ideal research capacity.

## Sample Policy Question: Do my state’s policies ensure a measurably effective educator workforce, and are these efforts evaluated to ensure that every student has an effective teacher?

	Typical SEA Staff Capacity (e.g., programmer, data manager)	Ideal SEA Staff Capacity (e.g., analyst, researcher)
Types of <i>Data Questions</i> This Staff Is Able To Formulate To Address the Policy Question	<ul style="list-style-type: none"> <li>■ What teacher data do we have?</li> <li>■ What student outcome data can we connect to individual teachers?</li> <li>■ Which school years do you want to show?</li> </ul>	<ul style="list-style-type: none"> <li>■ What are the indicators going to be used for? Do we need annual statistics? Trend analyses? Long-term evaluation or impact analyses?</li> <li>■ How do we discern the impact of the teacher as opposed to the impact of school, community, family or previous experience?</li> <li>■ Do we account for all teachers or just those in tested subjects and grades?</li> <li>■ How do we account for students who were in class less than a full year?</li> <li>■ What if a teacher was on leave for a significant period of time?</li> </ul>
Types of <i>Analyses</i> This Staff Can Perform Using SLDS Data	Calculate counts, percentages and averages for school, teacher and student performance	Use advanced statistical procedures on multiple years of connected student, teacher and school data to explore the influence of each on subsequent performance and outcome indicators and investigate trends over time
Types of <i>Information</i> This Staff Can Provide To Answer This Policy Question	<ul style="list-style-type: none"> <li>■ Highest degree obtained</li> <li>■ History of professional development for teachers</li> <li>■ Certification history</li> <li>■ Years of experience</li> <li>■ Average Advanced Placement scores, student attendance, assessment scores and course grades</li> </ul>	<ul style="list-style-type: none"> <li>■ Teacher value-added analysis (e.g., the Tennessee Value-Added Assessment System)</li> <li>■ Student annual growth model (e.g., Colorado Growth Model)</li> <li>■ Regression analyses to determine the amount of student test score variability that can be attributed to teachers or schools</li> <li>■ Correlation between student course grades and scores on state/district assessments</li> <li>■ Analyses of student performance one to three years later</li> </ul>

In the past, when data use was predominantly focused on mandated compliance reporting, limited access to analysts might have sufficed, but given the complexity of analyzing longitudinal data, the full power of the SLDSs cannot be realized without engaging education analysts and researchers as part of a robust state education agency team to turn the data into useful information for policymakers and practitioners.

### Acquiring the Appropriate Skills

Because so many resources have been focused on the IT activities of data collection and storage, many state education agencies are unable to conduct new analyses or research on longitudinal data to assist local education agencies or policymakers. A combination of both internal and external expertise is likely necessary in most states.

State education agencies that do have internal research and analytical staff have developed a variety of models for organizing these staff inside the IT department or in stand-alone departments, including research offices, evaluation offices, and analysts in data and accountability offices. For example:

- **New Jersey** has an Office of Research and Evaluation that designs studies and collects and interprets data.
- **Tennessee** has an Assessment, Evaluation and Research Division whose objective is to analyze and report assessment results, not conduct research based on policy questions.
- **California, Maryland, Virginia** and **Washington** have researchers on staff. (Although Washington has researchers, it relies on institutions of higher education for the bulk of its research.)

State education agencies do need at least a minimum level of internal expertise and capacity around data-driven decisionmaking and research to ensure they are savvy consumers of data, correctly interpret the results and effectively translate those results into actionable practices; however, practicalities and budget constraints may prevent them from building the kind of internal capacity they want and need.

*State education agencies do need at least a minimum level of internal expertise and capacity ... [but] there are a variety of ways ... to acquire research and analytical capacity besides building a large internal staff.*

There are a variety of ways for state education agencies to acquire research and analytical capacity besides building a large internal staff. They should be open to partnering with external expertise to help design studies, define necessary data elements, conduct analyses or evaluations, and/or disseminate results or reports.

- **University-Based Researchers.** Universities with colleges of education or public policy programs typically have faculty who are both required to and interested in conducting research or program evaluations in the education sector. When the research and analysis needs of the state or local education agencies are aligned with the interest areas of university-based researchers, an excellent opportunity for collaboration exists, depending on the scope of the study and the data in the SLDS.

## Establishing a Common Education Vocabulary

Education stakeholders need accurate, timely, consistent and high-quality information about students and schools to plan effective learning experiences, improve schools and reduce costs. Given the high mobility of the student population — across districts and states and between K–12 and postsecondary — sharing high-quality data requires a common vocabulary for key data elements that exist in multiple data systems. The Data Quality Campaign is a member of the Common Education Data Standards consortium ([www.commonedatastandards.org](http://www.commonedatastandards.org)), which is working with states, districts, postsecondary institutions and the U.S. Department of Education to establish common education data standards because the absence of common definitions and formats across data systems increases the complexity of conducting essential analyses needed to improve education.

Examples of how vocabulary differences affect understanding of the data:<sup>1</sup>

**Retention** — In K–12, this term refers to a student being held back a grade level (and is viewed negatively); in postsecondary, it is a positive occurrence indicating that a student has returned to advance his or her education.

**Discipline** — In K–12, this term refers to information about suspensions, expulsions, behavioral modifications, etc.; in postsecondary, it refers to an academic field of study.

Example of differences across state longitudinal data systems:

**Exit codes** — States often mandate categories used to classify reasons students leave a school or district (e.g., pregnancy, incarceration, transfer, graduation, homeschooling and dropout). There is quite a discrepancy across states — and sometimes within states across years — in terms of the number, types and definitions of these coding categories, which prevents meaningful comparisons of exit data across organizations. In 2009, for example, the Texas Education Agency collected data on 14 exit codes, but in 2005 it documented more than 30 exit codes. In 2008, Indiana collected data on 30 exit codes, while Massachusetts collected data on 21 exit codes, eight of which were for dropouts.<sup>2</sup>

1 [www.commonedatastandards.org/faq1.html](http://www.commonedatastandards.org/faq1.html)

2 R. Curtin, Massachusetts Department of Elementary and Secondary Education, *A Long Road to a Longitudinal Data System*. Presentation to the National Academy of Education, 2008.  
W. Bruce, Indiana Department of Education, *Accounting for Every Student: Graduation Rates in Indiana*. Presentation to the National Academy of Education, 2008.  
Texas Education Agency, *Secondary School Completion and Dropouts in Texas Public Schools 2004–05, 2006*, <http://ritter.tea.state.tx.us/research/abs2.htm>.  
Texas Education Agency, *Secondary School Completion and Dropouts in Texas Public Schools 2008–09, 2010*, <http://ritter.tea.state.tx.us/research/abs2.htm>.

In addition to alignment and agreement on the research topic and scope, timelines present another issue. Aligning the state education agency's reporting needs, the academic calendar and the timing of SLDS collection periods may be difficult in some cases. However, university faculty bring substantial expertise about both content and research methodology that state education agencies often do not have, and their skills and knowledge can be invaluable assets to agencies potentially lacking analytic capacity.

A potentially inexpensive partnership model is to have graduate students analyze SLDS data under the supervision of experienced professors. Again, the analysis needs to align with the state's data and research requirements, but students may be motivated by the opportunity to gain access to data that can be used for a dissertation. This model necessitates supervision by a professor and consultation with state education agency staff about the questions to be addressed as well as the establishment of appropriate data sharing/use agreements.

- **Research Consortia.** Research consortia are emerging both at the local and state levels to mine and analyze data. The Consortium on Chicago School Research has supported research in Chicago Public Schools for a number of years, and the Research Alliance for New York City Schools was recently created. The Education Research Centers in Texas conduct research for the Texas Education Agency on policy and performance with the objective of improving educational decisionmaking. Consortia such as these provide local and

state education agencies with research and analytic capacity that is aligned to the pressing questions that the agencies face. Unlike institutions of higher education, this is the sole mission of these organizations.

- **Independent Research Organizations.** Independent research organizations, whether for profit or nonprofit, can provide research assistance to state education agencies. Because the research is contracted, the organizations bring their content and research expertise to conduct the analyses needed by the agency. The challenge is the cost of such collaborations, which may be prohibitive. The state education agencies must either fund the research organization to conduct the studies or find external funding.
- **Regional Education Laboratories (RELs).** The RELs can be used to provide analytic capacity or research methodology expertise to state education agencies. The U.S. Department of Education's Institute of Education Sciences funds 10 laboratories to work with state and local education agencies to provide rigorous research solutions and technical assistance to educators. The mission of the RELs is to help education agencies meet their research and analysis needs. The RELs are uniquely positioned to provide the analytic capacity that the agencies lack. For example, REL Appalachia has a research analyst stationed in the Kentucky, Tennessee, Virginia and West Virginia Departments of Education to provide needed research skills and respond to departmental requests.

## Increasing Research Capacity in States

State education agency leadership must work with stakeholder groups to identify the most pertinent policy, practice and data questions; prioritize research and analytical activities; and guide communication about the use of the information housed in the SLDSs. With the necessary tools in place, including expert analysts and valuable partnerships among all stakeholders, state education agencies have many resources available to ensure that the plethora of information now captured in SLDSs can be used to drive successful policy and practice and improve student achievement. The following examples show some of the ways states are tapping these resources.

### South Carolina's Use of External Research Capacity at Universities

#### *Housing basic skills within the state education agency while using external partnerships for sophisticated analyses*

The South Carolina Department of Education (SCDE) has an Office of Data Management and Analysis whose mission is to provide timely and accurate data to the department to enable stakeholders to make informed decisions about policy and practice. Within the office, three groups focus on data: Research Services, Programming Services and District Technology Services. The Research Services group provides the required data reports and responds to ad hoc requests from stakeholders. These stakeholders may include legislators, graduate students,

## Start the Conversation: Develop a P-20/Workforce Research Agenda

To make full use of the longitudinal data they are collecting, states need access to individuals with high-level analytical skills and research training who can mine the data and answer the multitude of policy and evaluation questions. The Data Quality Campaign (DQC) State Action 8 begins to define these needs by encouraging states to start the conversation through the development of a P-20/workforce research agenda.

According to the 2010 DQC state analysis, **28 states** have taken the necessary steps to:

- Develop a P-20/workforce research agenda in conjunction with other organizations; and
- Establish a process by which outside researchers can propose their own studies for approval and/or obtain state data for external research.

external researchers and staff from other state organizations that are conducting research projects. The work centers more on gathering data than on doing actual analyses.

The SCDE staff does not include statisticians, researchers or others with sophisticated analytic capacity. SCDE reaches out to Clemson University and the University of South Carolina to conduct the actual research projects. Research Services takes requests for data and research, reviews them, and sets priorities according to internal versus external requests. The group's role is primarily to provide the data needed for research that will be conducted by others. According to SCDE, the group would benefit from more staff to fulfill requests but also from having researchers on staff. The department recognizes that developing internal research capacity is unlikely and has, therefore, developed a strategy to establish and maintain partnerships with other government agencies (e.g., South Carolina's Office of Research and Statistics) and institutions of higher education to address SCDE's questions.

## Formalizing Partnerships for Sustainability in Kansas

### *Establishing a consortium to enhance the minimal analytical capacity of the state education agency*

The Kansas State Department of Education (KSDE) has a three-person research and evaluation group in house that fulfills internal requests for descriptive research questions. More significant research requests are fulfilled by affiliated university collaborators because KSDE realizes that the department

cannot support and retain the kinds of qualified researchers needed to answer and analyze sophisticated research questions. Thus, KSDE incorporated into its 2007 SLDS grant proposal the creation of the Kansas Educational Data Users Consortium (KEDUC), which includes KSDE, its board of regents, various stakeholders, and universities and colleges that supply the research capacity.

KEDUC has been tasked with three foci: (a) professional development around data and research, (b) the development of a research agenda, and (c) communication around data and research findings. KEDUC vets research requests to ensure the questions are relevant and aligned to pressing educational issues. KEDUC's objective is to ensure that researchers are asking the right questions the right way and using the right data to answer them. Researchers are granted access to the SLDS, within the bounds of the Family Educational Rights and Privacy Act and protecting individual privacy, and work with KSDE staff to obtain answers to the research questions.

## Establishing an External Research Center To Build State Education Agency Capacity in Arkansas

### *Developing a model for building analyst skill sets, including partnering externally*

The Arkansas Department of Education (ADE) uses both an external and an internal model for research capacity around its SLDS. Internally, ADE has limited research capacity and, until 2009, needed to seek outside assistance, primarily from the University of Arkansas and the University of Central Arkansas. The department relied on people it knew through personal relationships and professional networks to conduct the required research.

The Arkansas Research Center (ARC) at the University of Central Arkansas was created in 2009 "to foster effective educational data use and to serve as a clearing house for state agency educational data needed to benefit Arkansas schools." ARC was created through legislation and with SLDS funding to help ADE answer its policy questions.

The center has only two full-time employees (one of whom is a developer of the SLDS) and two part-time employees to respond to research requests and questions statewide. ARC recognizes that individuals need a special composite of skills to effectively mine data from the SLDS. They need a combination of technology, research and data skills linked to knowledge of education and schools. ARC also recognizes that a variety of student-level data are needed to address its questions — not simply state summative assessments, but also more local and formative assessments. To that end, ARC is working with

## Lessons from a District: Jefferson County (KY) Public Schools' Substantial Commitment to Internal Research Capacity

### *Prioritizing the need to build research capacity within the district to answer sophisticated questions*

Each state, regardless of its demographics and size, potentially has a handful of local education agencies with substantial technical and human capacity around data-driven decisionmaking from which the state education agency can learn. Some local education agencies have had sophisticated data warehouses and other technology solutions to support their data for many years. Some are commercial products; others are customized home-grown solutions.

One such district, Jefferson County Public Schools (JCPS), is the 26th largest district in the country, serving Louisville, KY, and its surrounding area. JCPS is able to support a large information technology staff and a research department of more than a dozen highly trained researchers. These two departments have produced, through a series of grants, a powerful data warehouse, an assessment system that all educators use and a variety of other technologies to support data-driven activities.

The research staff works with administrators and teachers to address the district's research questions by providing valid, reliable and useful data to stakeholders. The staff can handle the simplest to the most sophisticated research questions and has the technological capacity to support the process.

Core lessons that can be extrapolated from JCPS' experience include:

- Committing to the importance of data;
- Developing a vision for data use that comes directly from senior leadership;
- Creating technology to support that vision and the associated educational objectives; and
- Recognizing that having the technology is necessary but not sufficient.

UCLA to integrate a model of formative assessments into its system. The goal is to integrate the local data into the state system in a transparent manner so that the SLDS can function as the data repository for all local education agencies as well as ADE. It then is up to ARC to mine those data, make sense of them and communicate the findings to stakeholders.

## Ensuring at Least Minimal Internal Capacity

These promising practices highlight states that are encouraging research and analysis through minimal state education agency capacity, relying mostly on external partnerships. While DQC applauds their efforts and emphasis on acquiring these skills, the Campaign encourages all states to build at least a minimal degree of internal agency capacity. This capacity will allow the state education agency to ensure that any external partnerships are conducting credible research and analysis that is reliable and valid.

## Acknowledgments

The Data Quality Campaign would like to thank **Ellen B. Mandinach, WestEd**, and **Nancy J. Smith, DataSmith Solutions**, for their expertise, guidance, research, draft writing and editing that, combined, form the heart of this paper.



To download DQC resources,  
visit [www.DataQualityCampaign.org](http://www.DataQualityCampaign.org),  
follow us on [Twitter](#)  
or visit us on [Facebook](#).

The **Data Quality Campaign (DQC)** is a national, collaborative effort to encourage and support state policymakers to improve the availability and use of high-quality education data to improve student achievement. The campaign provides tools and resources that will help states implement and use longitudinal data systems, while providing a national forum for reducing duplication of effort and promoting greater coordination and consensus among the organizations focused on improving data quality, access and use.