ISSUE BRIEF

Many Missing Pieces
The Difficult Task of Linking Early Childhood Data and School-Based Data Systems

Laura Bornfreund and Maggie Severns, Early Education Initiative
October 2010

Think about your school district and state: Are 5-year-olds entering kindergarten with everything they need to thrive? Do leaders know which early childhood programs work, which don’t and for whom? Is anyone tracking the impact of full-day kindergarten in the community and what happens when it’s not available? Which preparation programs or post-secondary credentials are associated with effective teaching in early education programs, including the elementary grades?

Questions like these usually go unanswered because early childhood data are not collected, not coordinated, or not accessible to the stakeholders who need them. Without longitudinal data – such as information on individual children’s progress from year to year – researchers struggle to conduct useful studies. Even aggregated information is hard to come by, hampering the ability of state officials, as well as school district officials, to make sound decisions about investments in early childhood education policies and programs.

Teachers and principals are in the dark too. Kindergarten teachers often begin the school year with little knowledge about the students entering their classrooms. They don’t know what skills their students have developed or whether they attended a preschool or childcare center the year before. Teachers in the first-through-third grades also lack information: They may have access to third grade standardized test scores but have no data from the prior grades to analyze student growth. Principals have no way to determine whether achievement is related to one good teacher; a series of teachers; the child’s attendance in a Head Start or pre-kindergarten program; or some combination of these factors.

Even parents struggle to find solid and easily accessible information about early childhood programs.

Putting relevant and longitudinal information in the hands of stakeholders – researchers, policymakers, school district officials, principals, teachers, and parents – would help to ensure that programs and districts are equipped to address the needs of children and families. States are starting to recognize the importance of collecting data across the full span of a child’s educational experiences, beginning as early as possible, and keeping track of that child’s progress.
Over time. Over the past five years, the federal government has dedicated roughly $515 million to help states build and expand longitudinal data systems to do just that.\(^1\) The latest round of federal grant funding – augmented by $250 million from the American Recovery and Reinvestment Act (ARRA) – explicitly called for states to make linkages between data on early childhood programs and the traditional K-12 system.

In many cases, states do not have a channel for collecting data on individual children in Head Start to integrate into their longitudinal data systems.

Even with these investments, an analysis by the New America Foundation has found that states have only just begun to make the right linkages. Gathering the right types of data and creating seamless systems will be a difficult task. Most states are still a long way from collecting early childhood information for purposes beyond compliance, much less connecting it in a cohesive way to existing K-12 longitudinal data systems. Some states are including – or have plans to include – data from state-funded pre-kindergarten programs in their longitudinal data systems, but this still leaves out a number of publicly funded early childhood programs and information about the children who attend them.

For example, data from Head Start, the federal government’s pre-kindergarten program for children in poverty, are reported collectively to the U.S. Department of Health and Human Services. In many cases, states do not have a channel for collecting data on individual children in Head Start, let alone integrating it into their longitudinal data systems. Even among state-funded programs, like those that provide subsidies to childcare centers, data are collected by human services agencies, not education agencies, making it difficult to link and share data. And these challenges don’t even touch the issue of how to integrate information from the private sector – the non-profit and for-profit organizations that run preschools, childcare centers, and afterschool programs as well as family-run childcare programs that serve the majority of children in early childhood settings in many states.\(^1\)

In short, gleaning an accurate picture of how well states serve young children is like trying to complete a jigsaw puzzle when you’re missing half the pieces.

In this report, we shed light on those missing pieces. We have analyzed proposals from recent grant winners to provide summaries of how they plan to collect data from early childhood programs and to integrate it with K-12 statewide longitudinal data systems. And we provide recommendations for making early childhood data – from birth through 3rd grade – more coordinated and far more useful to educators, parents, researchers, and policymakers.

**State K-12 Longitudinal Data Systems**

Since the early nineties, many states have developed systems that track data on K-12 student enrollment, demographics, achievement, and participation in educational programs– primarily for compliance with state and federal regulations. However, the purposes for tracking data have shifted as stakeholders have expressed a need for more and better data to inform classroom instruction, improve student learning, and guide policy decisions.\(^2\)

State departments of education are still capturing the necessary compliance data, but they are now using their statewide longitudinal data systems to gather and link data

\(^1\) These systems are often called P-16, P-20 or P-20/W systems in which the P typically stands for “preschool,” which itself may be defined as strictly pre-kindergarten (meaning programs for 3- and 4-year-olds) or preschool, meaning any early learning experiences that children have prior to entering school. The 16 or 20 typically derives from the number of years a child might be tracked in the system. The W stands for workforce.
Advocating for Effective Use of Data: The Role of the Data Quality Campaign

Since 2005, the non-profit organization Data Quality Campaign (DQC) has advocated for the effective use of quality data to improve student outcomes and promoted the development and implementation of statewide longitudinal data systems. DQC has identified 10 essential elements of high-quality longitudinal data systems:

- A unique student identifier that connects student data across key databases and years
- Student-level enrollment, demographic, and program participation information
- The ability to match individual students’ test records longitudinally to measure growth
- Information on untested students and the reasons they were not tested
- A teacher identifier system with the ability to match teachers to students
- Student-level transcript data, with information on course completion and grades earned
- Student-level college readiness test scores
- Student-level graduation and dropout data
- The ability to match student records between P-12 [preschool-through-12th-grade] and postsecondary systems
- A state data audit system assessing quality, validity and reliability.

The DQC regularly surveys states to determine the progress they are making towards incorporating the essential elements. So far, surveys show that 12 states have all 10 essential elements in place, and 34 have at least eight. The two least common elements are a statewide teacher identifier and student-level course information or transcript data.

However, building systems is not enough. Specifically, in 2009, DQC identified 10 state actions for states to undertake to ensure statewide longitudinal data systems are used for continuous improvement in public education:

- Link state K-12 data systems with early education, post-secondary and workforce systems, as well as other agencies such as social services
- Create stable and sustained support for longitudinal data systems
- Develop governance structures, assigning specific roles and responsibilities
- Build state data repositories that integrate student, staff, financial and facility data
- Implement systems to provide timely access to information
- Create individual student data reports that provide information that educators, parents and students can use to improve student performance
- Create reports that include longitudinal statistics on school systems and groups of students to guide school-, district-, and state-level improvement efforts
- Develop a research agenda and collaborate with universities, researchers and intermediary groups to explore the data for useful information
- Implement policies and promote practices, including professional development and credentialing, to ensure educators know how to access, analyze and use data
- Promote strategies to raise awareness of available data and ensure that all key stakeholders, including state policymakers, know how to access, analyze and use the information.

The DQC also surveys states about whether they have implemented the 10 actions. Right now, only Texas says it has acted on at least eight, with only Arkansas close behind, reporting seven. While many states are making progress on meeting essential elements as defined by DQC, few are taking the necessary steps to ensure data are used for continuous improvement.
on individual students – not just groups of students – across multiple years, multiple agencies, and the entire education spectrum.

While longitudinal data are useful to a wide range of stakeholders – from educators and parents to researchers and policymakers – not everyone needs data for the same reason.

Consequently, stakeholders do not need, nor should they have, the same level of access. Teachers need longitudinal data for the students in their classroom from their previous years of schooling to help them target their instruction and identify students who need additional help. Researchers and policymakers need aggregated data, or data that identifies individual students by numbers instead of names to determine the effectiveness of programs and practices and the performance of schools. State policymakers use data systems to evaluate the capacity of teacher preparation programs – traditional and alternative – and of professional development to produce effective educators. For district and state leaders, data systems provide essential information about individual student progress, teacher effectiveness, and school or program quality, by calling attention to both positive and negative trends over time.

In 2005, as states attempted to gather and process the information for these systems, several policy groups launched an effort called the Data Quality Campaign to improve the collection, availability, and use of high quality longitudinal education data. (See sidebar, page 3.)

A Short History of Early Childhood Data Collection

Efforts to collect and integrate early childhood data into longitudinal data systems have been more fragmented than efforts to coordinate K-12 data. In 2005, the National Governors Association (NGA) Taskforce on School Readiness recommended that states:

- Implement unified data collection requirements, training opportunities, and professional standards across pre-kindergarten, childcare, and Head Start programs;
- Establish common measurements and consistent data reporting mechanisms to enable information sharing and analysis across state agencies and programs and between the state and local levels; and
- Invest sufficient resources to support consistent data collection efforts.

In its conclusion, the report stated that “governors can focus on building ‘ready states’ by supporting a coordinated and comprehensive infrastructure for early childhood, integrating data systems and supporting evaluation efforts to inform decisions, and holding decision makers and stakeholders accountable for measurable results.”

States have started working on the NGA recommendations; but for the majority, early childhood data collection remains largely uncoordinated. This is true even for states like Florida that have been recognized as leaders in building K-12 longitudinal data systems.

Still, during the past few years, some states have made progress. For example, Illinois now assigns children a unique student identifiers, essentially ID numbers, when they enroll in a publicly funded early childhood program. It also gathers information on their family structures and parents’ income levels. For policymakers, these data help to answer questions about accessibility as well as about which programs work best for which types of students. Additionally, school administrators and teachers can access academic data on each child in their classroom, which enables teachers to tailor instruction for each child and principals to determine the types of professional development and support teachers may require. Maryland administers a school readiness assessment to all kindergarteners in public schools and filters the results by type of prior care. Maryland also attaches a unique student identifier, so this information is included in the state’s education longitudinal data system. Missouri has begun
collecting data on the workforce in early childhood programs, including the percentage with bachelor degrees (or higher) and the average wages of staff.  

Pennsylvania is a leader in this arena, as recognized by a national group called the Early Childhood Data Collaborative. (See sidebar, page 5.) The state's Early Learning Network collects data on children and teachers in programs administered by the Office of Child Development and Early Learning. Information collected includes children’s health information and enrollment details, as well as teacher qualifications and classroom quality rating scores. Future goals for the state include building the capacity for the Early Learning Network to encompass information on other services such as Temporary Assistance for Needy Families (TANF), Medicaid, child welfare, and juvenile justice programs. Because of the sensitive nature of much of this data, state policymakers will need to ensure that families’ privacy is protected. (See sidebar, page 13.)

Pennsylvania leaders are working to connect the Early Learning Network with its K-12 education data warehouse. While it does not yet have the ability to incorporate federal Head Start data, Pennsylvania intends to do so.

**Federal Support for State Longitudinal Data Systems**

Coinciding with the movement for more and better data, federal lawmakers established the Statewide Longitudinal Data Systems (SLDS) grant program (part of the Educational Technical Assistance Act of 2002) to help states design, develop, and implement longitudinal data systems. There have been four rounds of grant awards, in which a combined 41 states and the District of Columbia have received $515 million for multi-year initiatives to build out their SLDS and to improve their use of data to drive decisions and to facilitate longitudinal research. (See map, page 6.) The Department of Education intends to use the lessons learned by and experiences of grant recipients to aid

---

**The Early Childhood Data Collaborative**

The Early Childhood Data Collaborative (ECDC) supports state policymakers’ development and use of coordinated state early care and education data systems. The ECDC has identified 10 fundamentals of coordinated state data systems:

- A unique statewide child identifier
- Child-level demographic and program participation information
- Child-level data on child development
- The ability to link child-level data with K-12 and other key data systems
- A unique program site identifier with the ability to link with children and the ECE (early childhood education) workforce
- Program site data on structure, quality and the work environment
- Unique ECE workforce identifier with the ability to link with program sites and children
- Individual ECE workforce demographics, including education and professional development information
- A state governance body to manage data collection and use
- Transparent privacy protection and security practices and policies.

The ECDC’s key policy questions for state and district policymakers to consider are:

- Are children, birth to age 5, on track to succeed when they enter school and beyond?
- Which children have access to high-quality early care and education programs?
- Is the quality of programs improving?
- What are the characteristics of effective programs?
- How prepared is the early care and education workforce to provide effective education and care for all children?
- What policies and investments lead to a skilled and stable early care and education workforce?
other state and local education agencies with their respective longitudinal data systems.

The fourth round of SLDS awards was funded by the American Recovery and Reinvestment Act (ARRA) in February 2009, which provided $250 million for this purpose. The Department of Education required applicants to focus on the specific goal of expanding data systems to cover the span from early childhood through the start of a young adult’s career. The impetus for this advancement was the passage of the America COMPETES Act of 2007, through which Congress specified the essential elements for statewide longitudinal data systems, including the requirement that systems have “the ability to share data from preschool through postsecondary education data systems.” (See sidebar, page 7.)

The 2009 ARRA SLDS grants were a significant expansion of the funding available to states for improving data systems. The $250 million awarded in May 2010 (due to ARRA) was nearly two-thirds more than the $150 million awarded the year before. Other recent initiatives have also shown how much the federal government is pushing for the improved use of data. At the U.S. Department of Education, longitudinal data systems development is one of the four education reform priorities, the “four assurances,” central to the State Fiscal Stabilization Fund, Race to the Top, and the Investing in Innovations program. At the U.S. Department of Health and Human Services, coordination of data is a priority for state-level early childhood advisory councils that are receiving new federal funding this fall.

**How Many Federal Grants Have States Received?**

Interactive map available at earlyed.newamerica.net

---

**SOURCE:** U.S. Department of Education; New America Foundation analysis
Further amplifying the importance of linking data across the educational continuum and sharing data across agencies is President Barack Obama’s focus on education.

In his March 2009 speech to the Hispanic Chamber of Commerce, he said:

“[E]very state has a data system like the one in Florida that keeps track of a student’s education from childhood through college. And far too few districts are emulating the example of Houston and Long Beach, and using data to track how much progress a student is making and where that student is struggling. That’s a resource that can help us improve student achievement, and tell us which students had which teachers so we can assess what’s working and what’s not. That’s why we’re making a major investment in this area that we will cultivate a new culture of accountability in America’s schools.”

President Obama also included $65 million for SLDS grants in his budget request to Congress for fiscal year 2011, and the Senate’s Appropriations Subcommittee has voted to include $58 million in its proposed budget. At present, Congress has not yet finalized how it will appropriate federal money for the 2011 fiscal year.

Why Include Early Childhood Data? The Case for Integration

As states develop and refine their statewide longitudinal data systems, there are several reasons for policymakers to consider including early childhood data:

Better Teachers, Programs, and Schools
Currently, it is common for preschool teachers to see themselves in one system and for K-3 teachers to see themselves as part of another. This hampers teacher communication across grade levels about curricula and student progress. Kindergarten teachers who want information about a child’s performance in preschool, his attendance records, or what other state-funded programs served him in the past, face several barriers to accessing individual student information. Much of this information is not even collected at the student level. Instead, the data that

Required Elements of a Statewide Education Data System

America COMPETES Act of 2007

1. An unique statewide student identifier that does not permit a student to be individually identified by users of the system;
2. School enrollment history, demographic characteristics, and program participation record of every student;
3. Information on when a student enrolls, transfers, drops out, or graduates from a school;
4. Students’ scores on tests required by the Elementary and Secondary Education Act;
5. Information on students who are not tested, by grade and subject;
6. Students’ scores on tests measuring whether they’re ready for college;
7. A way to identify teachers and to match teachers to their students;
8. Information from students’ transcripts, specifically courses taken and grades earned;
9. Data on students’ success in college, including whether they enrolled in remedial courses;
10. Data on whether K-12 students are prepared to succeed in college;
11. A system of auditing data for quality, validity, and reliability; and
12. The ability to share data from preschool through postsecondary education data systems.

do exist reflect aggregate information such as the number of girls or the percentage of Hispanic children enrolled in a given program. In states that do collect individual student information for pre-k and other early childhood education programs, current system capabilities – and in some cases misunderstandings of privacy regulations – make it nearly impossible for teachers and administrators to obtain information in a timely and user-friendly manner.

Access to early childhood data not only helps K-3 teachers improve and differentiate their instruction to meet all students' needs, it also helps teachers in pre-k and other programs for children age 5 and younger. Knowing what happens to students in later grades can provide valuable insight on how academic skills and behavioral patterns formed in the early years develop and change over time, reflecting the impact of early childhood programs. And yet providers of Head Start programs, for example, rarely hear anything about how children progress after they enter kindergarten.

Improved Early Warning Systems
Tracking the needs and achievements of children from a young age can help researchers understand what places them at risk, enable specialists to improve early warning systems, and allow schools to begin interventions at the earliest possible signs of distress. For example, the Data Quality Campaign has laid out 10 “essential elements” of strong statewide longitudinal data systems (see sidebar page 3), one of which is the collection of student-level graduation and dropout data, to determine who drops out of high school, how, and why. With good longitudinal data collection, policymakers and educators can learn more about what types of students start school behind; struggle in the upper elementary grades, middle school, and high school; and identify warning signs of poor high school performance that are visible in the earlier years and grades.

Data for Researchers and Policymakers
For years, debates have raged over the effectiveness of pre-k, Head Start, and full-day kindergarten, and whether children’s gains dissipate over time. If more student-level longitudinal data were available to researchers and policymakers, they could answer many of the central questions surrounding the effectiveness of and need for aligned programs that start before kindergarten and extend through the early grades of elementary school. Longitudinal education information could be a key to deciphering when and how early childhood programs work and where better alignment is needed. For privacy and ethical reasons, of course data would need to be scrubbed of any information that could be used to identify individual students. Unlike teachers and principals, researchers and policymakers do not need personally identifiable information, nor should they have access to data that would be easily traced back to a specific student or family member.

Major policy questions at the state and local level that could be answered with better longitudinal data include:

- In what ways and to what extent do early childhood programs – including home visitation, Head Start, childcare programs, state-funded pre-k, and full-day kindergarten – benefit children in their later years?
- Which students benefit the most from which type of early childhood program?
- What types of early childhood programs work, and which don’t and for whom? Is there a difference between pre-k programs offered by the public school district and those provided by community-based organizations? Do students who attend full-day kindergarten do better than those who do not?
- How important is it for pre-k teachers, childcare providers, and home visitation workers to have professional certification and/or post-secondary degrees? What types are associated with the best outcomes?
- How important is it for K-3 teachers and school administrators to have formal education and/or ongoing professional development related to early childhood development?
- What is the impact of early childhood programs, from birth to age 8, on children over time? What is the
incremental and cumulative contribution of ensuring that children have high quality experiences in grades one, two, and three? In what grades, and under what conditions, does this happen?

Questions on Quality

Of course, the utility of longitudinal information will hinge on the quality of data collected. Through our analysis, we were unable to get a clear picture of the type of data currently collected by states and whether those data are collected consistently across schools and districts. Consider attendance data, for example, and how schools and districts track absences. In 2008, the National Center for Children in Poverty released a report, written by researchers Hedy Chang and Mariajosé Romero, on the negative impact of early chronic absenteeism. They reported that children who miss numerous days of school, especially in kindergarten, are more likely to have lower levels of academic achievement in elementary school. They also found that many schools do not have standardized collection systems in place to determine who is absent from kindergarten and how many days they have missed. Student mobility between schools, let alone between districts or states, further exacerbates this problem. They recommend that school districts begin tracking children’s attendance when they enter pre-k programs “to identify if attendance is problematic prior to elementary school and to track whether participation in pre-kindergarten is helping to reduce chronic early absence in kindergarten.”

The very fact that inconsistencies and omissions are occurring in the collection of attendance data – a seemingly basic piece of information about students – begs the question of how to ensure the quality and consistency of more complicated data. State policymakers will need to employ data quality experts to ensure that the data chosen to be collected are reliable, valid, and able to answer the questions stakeholders identified.

Among the Missing Pieces: Pre-Kindergarten Data Collection

At first glance, the latest proposals from state policymakers to improve their systems suggest that states are already making major strides to connect data from early childhood programs to their K-12 longitudinal data system. According to the U.S. Department of Education website, 14 state grantees are able to link individual student pre-kindergarten data to their K-12 data systems, and another 10 states are currently working to do the same. According to a survey of state K-12 data systems managers conducted by the Data Quality Campaign, 41 states report having some capacity to link early childhood program data to their K-12 data systems.

But these numbers mask the extent of problems with current efforts at coordinating data and building a system that can provide a full picture of children’s progress. State and federal policies may conflict over how much data can be shared and with whom. Data sharing between agencies within a state is often hampered by different policies at the agency level. The inclusion of data on children who attend Head Start can be particularly challenging. It is encouraging to see the recent focus on eliminating these barriers but policymakers should approach the work with a clear-eyed sense of how many parts will need to come together to create a coherent whole.

First, multiple agencies oversee early childhood programs, which means that the information that a longitudinal data system is supposed to capture is spread across myriad government agencies administering numerous state, local, and federal programs each with its own set of rules. Officials in a state’s education department, which typically

---

**ii** Chang and Romero define early chronic absenteeism as students missing almost a month of school or more in a given school year.
manages the longitudinal data system, therefore face a daunting task. Although preschool programs funded by the federal Individuals with Disabilities Education Act (IDEA) are likely to be administered through state education departments, other programs are not. Departments of health and human services or labor, for example, often manage childcare subsidy programs. State officials who confer with local officials of Head Start and Early Head Start – programs administered by local entities – are also likely to be part of human services departments. Each of these agencies has its own mechanism for collecting information, which may not currently interact easily with another system. Consequently, administrators for early childhood programs may be collecting individual student or program data that are not included in the longitudinal data system designed by the education department. This is why it is essential for state agencies and sectors to develop common data standards and definitions for data systems.

There are no examples, to our knowledge, of any states that have fully incorporated data from the diverse array of early childhood programs into their K-12 longitudinal data systems.

In part, this disconnect is due to funding streams: Funding for early childhood programs can come from state and federal governments as well as local districts. And at the federal level, education funding for children under 5 comes from both the U.S. Department of Education – for example, programs funded by Title I of the Elementary and Secondary Education Act, a major program for disadvantaged students – and the U.S. Department of Health and Human Services. Moreover, while federally funded Head Start programs are required to report to the U.S. Department of Health and Human Services, they are not required to report information to the state in which the program resides.

There are some examples of states – Pennsylvania, as mentioned above, and Minnesota – that are working to integrate data from the various early childhood programs, but there are no examples, to our knowledge, of any states that have fully incorporated data from the diverse array of early childhood programs into their K-12 longitudinal data systems.

For instance, as mentioned above, the education community has long recognized Florida as a pioneer in developing a longitudinal system for education data. But, when it comes to making comprehensive linkages with early childhood education, the state still has a long way to go.

Florida’s “P-20” statewide longitudinal data system merges data from 26 state agencies, and since the 1995-96 school year has collected and analyzed student demographics, enrollment, courses, test scores, financial aid, and awards, as well as data on curricula, educational institutions, staff demographics, certifications, and professional development. This enables policymakers to identify:

- Schools that show high levels of student growth;
- High school factors that lead to success in college or the workplace; and
- Teacher preparation programs that produce effective teachers.

Florida is able to capture information about pre-k services offered by public school districts, and it collects information from both public and private providers who receive public funds to operate Florida’s voluntary pre-kindergarten (VPK) programs. Additionally, using the data collected, the state “grades” VPK providers based on their students’ performance on the kindergarten readiness assessment. Florida is also studying VPK students’ longitudinal progress. Students in the first VPK class were third graders in the 2009-10 school year, and the Florida Department of Education intends to release a report on their success over time.
Where does Florida’s system still need work? Connecting data from other publicly funded programs beyond VPK. The state does not collect data, for example, on students enrolled in Head Start, Early Head Start, or subsidized childcare programs. In the 2008-09 school year, Head Start programs alone served 9 percent of 4-year-olds in the state; another 23 percent of the state’s 4-year-olds are in programs other than VPK, Head Start, or special education services.33 The state is missing valuable information on tens of thousands of children.

Additionally, pre-kindergarten teachers who work in VPK participating centers that are not public schools are not assigned unique identifiers. That means there is no way to link them to the students they have taught. Valuable information about teachers is lost. Finally, Florida’s longitudinal data system does not incorporate information on other birth-to-5 services in which children and their families may participate, such as home visitation.

If Florida, a recognized leader in K-12 longitudinal data systems, still has this many missing pieces related to early childhood, it is likely that most other states have even more work to do.

### States’ Top Challenges to Linking Early Childhood and K-12 Data

<table>
<thead>
<tr>
<th>Lack of resources</th>
<th>Lack of common student ID</th>
<th>Lack of coordination between ECE and K-12 authorities</th>
<th>Incompatible systems</th>
<th>FERPA and other privacy concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska</td>
<td>Alaska</td>
<td>Washington, DC</td>
<td>Alaska</td>
<td>California</td>
</tr>
<tr>
<td>Alabama</td>
<td>Alabama</td>
<td>Idaho</td>
<td>Alabama</td>
<td>Kansas</td>
</tr>
<tr>
<td>California</td>
<td>California</td>
<td>Louisiana</td>
<td>California</td>
<td>Maryland</td>
</tr>
<tr>
<td><strong>Colorado</strong></td>
<td><strong>Colorado</strong></td>
<td>Delaware</td>
<td><strong>Colorado</strong></td>
<td>Michigan</td>
</tr>
<tr>
<td>Idaho</td>
<td>Idaho</td>
<td>Montana</td>
<td>Idaho</td>
<td>Nebraska</td>
</tr>
<tr>
<td>Kentucky</td>
<td>Maine</td>
<td>Louisiana</td>
<td>Kentucky</td>
<td>Ohio</td>
</tr>
<tr>
<td>Louisiana</td>
<td>Mississippi</td>
<td>Montana</td>
<td>Mississippi</td>
<td>South Dakota</td>
</tr>
<tr>
<td>Maryland</td>
<td>North Dakota</td>
<td>New York</td>
<td>North Dakota</td>
<td>Wisconsin</td>
</tr>
<tr>
<td><strong>Maine</strong></td>
<td><strong>Maine</strong></td>
<td>Ohio</td>
<td>South Dakota</td>
<td></td>
</tr>
<tr>
<td>Michigan</td>
<td>New York</td>
<td>Rhode Island</td>
<td><strong>Virginia</strong></td>
<td></td>
</tr>
<tr>
<td>Missouri</td>
<td>Ohio</td>
<td>South Carolina</td>
<td>Washington</td>
<td></td>
</tr>
<tr>
<td>Montana</td>
<td>South Dakota</td>
<td>South Dakota</td>
<td>Wisconsin</td>
<td></td>
</tr>
<tr>
<td>Nebraska</td>
<td><strong>Virginia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Hampshire</td>
<td><strong>Wisconsin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nevada</td>
<td>South Dakota</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Oregon</strong></td>
<td><strong>Wisconsin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Dakota</td>
<td>Wyoming</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: 2009-10 Survey Results, Action 1, Part 1, Question 92, Data Quality Campaign. (Note: 2009 ARRA SLDS grant recipients are in bold.)
In 2009, the Data Quality Campaign conducted its annual survey of states on the status of their accomplishments towards the 10 essential elements of statewide longitudinal data systems. States identified several challenges with connecting data on publicly funded early education programs and K-12 education.34 See the table above for issues that states reported as particularly problematic.

Some states offered specific comments about the challenges they face: 35

- Kansas does not link information about children receiving subsidized childcare because the state’s Department of Education does not provide oversight for that type of childcare.
- Ohio is currently in the process of eliminating barriers to linking publicly subsidized programs that are administered by agencies other than the Ohio Department of Education. Under new legislation, these programs will become part of the new Center for Early Childhood Development under the authority of the ODE.
- Oklahoma, in its SLDS grant application, requested funding to expand its Unique Identifier System. The state wants to assign children a number as early as possible in birth-to-pre-k programs that would then be utilized through their entire educational career. Oklahoma’s application was not funded.

Nine states noted the challenge of incompatible systems. For the purposes of this survey, they are most likely referring to a lack of common data standards and definitions among state agencies. However, this challenge could also be true for school district data systems. Are they compatible with one another? When children move from one school district to another what happens to their data?

The same questions arise for students who move from state to state. Massachusetts and Connecticut are examples of states working to make their longitudinal data systems more compatible to transport students’ information if they transfer across state lines.36

Another challenge comes with the influx of federal funds directed toward longitudinal data systems. For example, the official charged with managing a given state’s federal SLDS grant might not have been part of conversations about the development of their state’s application for the Race to the Top competition. States may inadvertently start creating duplicate efforts around data collection and use.37 As the work to improve and expand longitudinal data systems progresses, it will be essential for state agency leaders to be aware of parallel data system efforts in order to share the work instead of duplicating it.

What’s Next: How States Plan to Use New Federal Dollars to Improve their Data Systems

What are states doing to overcome these and other barriers? And what policies are states putting into place to improve data collection, sharing, and use?

The New America Foundation analyzed the 20 state applications from the 2009 ARRA funded Statewide Longitudinal Data System grant program administered by the U.S. Department of Education’s Institute for Educational Sciences. The analysis found plans to improve interoperability (the ability of data systems housed in different agencies to “talk to each other” and exchange data) between agencies as well as between states; include data from federally funded programs including Head Start into longitudinal systems; make data more accessible and usable to teachers across the pre-k-12 spectrum, parents, researchers, and the public; and assign teachers in state-funded pre-k or other preschool programs with unique identifiers. These are important steps for states to be taking.
### A Look at How States Plan to Use 2009 ARRA SLDS Grants

<table>
<thead>
<tr>
<th>State (Percentage of 4-year-olds served in 2009 by state-funded pre-k)</th>
<th>Pre-k-12 tracking using a unique student identifier</th>
<th>Data sharing between programs and/or departments that serve young children</th>
<th>Kindergarten readiness assessments included in SLDS</th>
<th>Access portal for parents, nonprofits, and the public</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arkansas (44%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorado (20%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Florida (67%)</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illinois (20%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kansas (21%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maine (19%)</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Massachusetts (11%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michigan (19%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minnesota (20%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mississippi (0%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| New York (43%) | | | | *
| Ohio (8%) | | | | |
| Oregon (8%) | | | | |
| Pennsylvania (16%) | | | | |
| South Carolina (38%) | | | | |
| Texas (45%) | | | | |
| Utah (0%) | | | | |
| Virginia (14%) | | | | |
| Washington (17%) | | | | |
| Wisconsin (48%) | | | | |

* Tracks students using social security number

Note: States were left blank if a state did not include information about a capability on its 2009 ARRA SLDS grant application. This could indicate that a state has made no progress on this capability or that the state simply chose not to mention it in this particular grant application.

Complete

In progress. Progress may be part of 2009 ARRA grant.

This capability will be planned and/or started with 2009 ARRA grant.

Sources: 2009 ARRA SLDS grant applications; 2009 NIEER yearbook.
The chart on the previous page provides details on all 20 states. Below, we highlight nine proposals and the federal funding they have received to date. These proposals were selected because they show the variety and depth of improvements that states are hoping to undertake in the coming years using SLDS grant funding:

Arkansas
$9.8 million from 2009 ARRA funds; $18.1 million total
The Arkansas Department of Education has entered into a data-sharing agreement with the state’s Department of Human Services, which administers the state’s pre-k program, Arkansas Better Chance. The Department of Education anticipates integrating 9,000 additional children into its longitudinal data system through this agreement and plans to remove any duplicate records and assign a unique identifying number to each child. Arkansas Better Chance maintains detailed demographic information on enrollees, data from health and developmental screenings, and information on the educational attainment of teachers and paraprofessionals who work at the partnering agencies, all of which will be incorporated into the system. Additionally, Arkansas Better Chance will supply the Department of Education with information on individual students, from 2005 to present, to allow the agency to begin conducting longitudinal research.

Colorado
$17.4 million from 2009 ARRA funds; $21.6 million total
Three different Colorado state agencies collect and maintain information about children who participate in publicly funded early childhood education programs: the Department of Education, the Department of Health Care Policy and Financing, and the Department of Human Services. Colorado plans to integrate data on the children attending these programs into its education longitudinal data system so it can collect more accurate information about services, link information about the programs in which children and their families are enrolled, and identify programs that are associated with strong readiness outcomes. In addition to the SLDS grant, Colorado was relying on a Race to the Top grant to help fund the upgrades to its data system. The state did not make the cut in the first Race to the Top round, nor was it awarded in the second round. Without Race to the Top funding, it is unclear whether the SLDS grant by itself will be enough.

Maine
$7.3 million from 2009 ARRA funds; $10.5 million total
The Maine Department of Education plans to assign students participating in state-funded early education programs a unique identifier and integrate the Department of Health and Human Service’s early childhood program data into its system by next summer. Additionally, Maine will establish a pilot project to track participation in Early Head Start, Head Start, and Educare, a non-profit program that serves children from infancy through age 5. These three programs serve more than 3,000 children who would be added to Maine’s longitudinal data system. The Department of Education intends to use its system to help researchers evaluate the effectiveness of early childhood programs.

Massachusetts
$13 million from 2009 ARRA funds; $19 million total
The Massachusetts Department of Early Education and Care plans to unify its separate data systems for preschool, homecare, and after school programs. In addition, the state will add information about state-funded early education programs, including student outcomes data that is linked to state assigned student IDs. Massachusetts also plans to:

- Provide school districts and schools with data on programs in which students are enrolled beginning at birth;
- Develop a system for assigning ID numbers to prospective pre-k – 12 educators;
- Design and implement data audits, a data quality curriculum, and a certification process for early learning centers; and
- Create a regional data sharing system with Connecticut.
Minnesota
$12.4 million from 2009 ARRA funds; $15.7 million total
In Minnesota, information collected about preschool students is currently limited to those who receive special education services—only about three percent of the total preschool population. Under the proposed expansion, enrollment and demographic data would be added for children attending the state’s Early Childhood and Family Education programs, School Readiness programs, and Head Start. Information about the specific sites and the teachers at those sites will also be collected. Data on kindergarteners will be expanded to include children’s developmental assessment results at kindergarten entry. In its application, the Minnesota Department of Education stated it would use this information to evaluate the quality and understand the impact of early childhood services.

Additionally, the SLDS grant will fund the development of a “dashboard” – a password-protected website that early childhood educators can use to analyze data on children in their care. The dashboard will maintain information on health factors (immunization and special education), development (screening, special education, and preparedness for kindergarten), and parent education (involvement in early childhood family education or Head Start programs).

New York
$19.7 million from 2009 ARRA funds; $27.5 million total
New York plans to add individual-level data from the Health and Human Services department to its longitudinal data system. The state already collects data from educational institutions in the State Office of Children and Family Services. Additionally, New York plans to strip its longitudinal data of all personally identifiable information so that it can be used for federal reporting, research, and analysis for policy purposes. The state plans to design a “public access portal” to help the public measure program effectiveness.

Ohio
$5.1 million from 2009 ARRA funds; $13.7 million total
Ohio currently collects aggregate data from its state-funded early childhood programs: Early Childhood Education, Preschool Special Education, as well as data from the Early Learning Initiative, for which funding was cut in the state’s 2009-10 budget. Currently, however, individual-level pre-k data are not included in Ohio’s longitudinal data system.

Ohio has passed legislation (the Ohio Revised Code) that enables the state to use its unique student identifier for students in its state-funded pre-k programs. Ohio plans to extend its data system capabilities to track childcare licensing, Head Start, Even Start, and Early Childhood Mental Health Consultation, and Nutrition Programs as well.

Oregon
$10.5 million from 2009 ARRA funds; $18.9 million total
Currently, Oregon has three education data systems which are operated by different agencies: one system collects data on children before they enter school and up through 12th grade, and two other data systems track higher education data.

The state stores “minimal preschool data:” demographic, program model, and attendance information on children in Oregon Head Start, and the Early Intervention/Early Childhood Special Education programs. Oregon’s current pre-k database does not store data collected from formative assessments; previous plans to do so were thwarted when state pre-k programs began using two differing assessments, and limited resources had prevented the state from inserting both of these assessments into the early childhood database separately. Oregon now plans to incorporate these assessment data into the system that is operated by the state department of education. The state also plans to synchronize early childhood data collection across its wide spectrum of pre-k providers so that all providers use a common transcript.
### Protecting Privacy

Student privacy is a major concern when it comes to sharing data, particularly when information about a child could be sent to an array of recipients — teachers, school districts, researchers, and federal agencies, for example. Here is a brief overview of some basic laws and regulations that govern discussions of data systems development:

**FERPA**

Programs administered by the U.S. Department of Education must comply with regulations derived from the Family Educational Rights and Privacy Act (FERPA) of 1974. FERPA was enacted to protect the privacy of students’ information, such as transcripts and disciplinary records. Additionally, the law gives parents the right to see their children’s educational records and to consent to their disclosure. Historically, some ambiguities in FERPA – coupled with changing technology and data environments since the statute was written – have led some states and districts to hesitate when it comes to sharing data across agencies or with outside parties such as researchers. However, current understanding is that an agency can provide data to a state’s department of education without contradicting FERPA as long as the data is not personally identifiable. Data systems must therefore employ security protection to ensure that a student’s name and address cannot be determined by someone without authorized access. However, data systems can tag children with a unique and protected ID number so that researchers and education leaders have the ability to trace a student’s progress from year to year and assess what factors may have contributed to that student’s achievements.

It is important to keep in mind that FERPA was written before longitudinal data systems were conceptualized. FERPA was not instituted to prevent longitudinal data tracking, nor the benefits it may bring, but instead to protect student privacy—a practice that must remain intact. Later this year, the U.S. Department of Education is expected to release new guidance on what FERPA allows.

**HIPAA**

If data collected on a child in an early childhood program contains health information and is not already protected under any FERPA regulations, that data may be regulated under the Health Insurance Portability and Accountability Act (HIPAA) of 1996. HIPAA regulations protect the privacy and security of a patient’s health records. The extent to which HIPAA regulations apply to early childhood programs will depend on whether those programs collect data on children’s health and what type of data it is.

**Head Start**

While the Office of Head Start encourages grantees to transfer student records to local school districts, it does not have a privacy policy except to say that Head Start programs cannot share data on individual children without written parental consent. Some Head Start grantees request that parents sign privacy policies stating that student records may be shared with school officials. But the performance standards are silent on statewide longitudinal data systems. And unless a state licenses or provides additional funds to Head Start providers within its borders, states have no jurisdiction to require what a Head Start program does with its records.

At the federal level, Head Start providers are required to send the Office of Head Start information about their programs such as the number of students served and their demographics, for example, but those data never include information on individual students.
Moving Forward

It is becoming increasingly important for states to have the ability to: use longitudinal data to guide instruction; identify students who would benefit from intervention or other services; evaluate programs, schools, principals, and teachers; conduct research; and inform local and state policy decisions. But only recently has this notion of linking data from the birth-to-five realm with data from K-12 generated the same sense of urgency.

For there to be meaning behind states’ assertions that their longitudinal data systems “connect to early childhood,” they must take steps to gather and integrate information from many early childhood programs, including Head Start and even subsidized childcare.

Ensuring the inclusion of early childhood data in longitudinal data systems is an essential step toward redefining the primary years- moving away from envisioning education only as K-12 toward a Pre-k-12 framework. This will help to advance efforts to create a seamless and aligned early childhood system for children through the third grade, an approach known as “PreK-3rd.” Data are needed to assess the effectiveness of PreK-3rd efforts. And data sharing within and across grade levels would provide teachers with the critical information they need to target their instruction to meet the needs of their students as they move from grade to grade.

In its recent white paper, the Early Childhood Data Collaborative recognized that states have made some progress when it comes to implementing, connecting, and using longitudinal data systems for early childhood programs, but at the same time it emphasized that simply building the infrastructure is not enough. States and the federal government must take steps to make data accessible to all stakeholders and to use – not simply collect – data to drive policy decisions. Doing so, while ensuring privacy is maintained, will lead to continuous improvement in early childhood programs, teaching, and learning.

Federal Recommendations

The Early Education Initiative at the New America Foundation offers the following recommendations for federal agencies that administer early childhood programs:

- Include data-system development among the priorities of the soon-to-be-formed “Interagency Policy Board” between the Department of Education and Department of Health and Human Services. This new board should explore how to improve system coordination between federally funded early childhood programs and states so that state longitudinal data systems include a fuller picture of children’s early childhood education experiences; and
- Make sharing early childhood data across agencies, in accordance with privacy laws and regulations, a priority and encourage state agencies to do the same.

State Recommendations

We recognize that every state offers different early childhood programs, has different data needs, and is in a different stage of data system development. With that in mind, the Early Education Initiative at the New America Foundation offers the following recommendations for states to consider as they take up the challenge of integrating early childhood data into their longitudinal data systems:

- Work to implement the Early Childhood Data Collaborative’s 10 fundamentals of coordinated state early care and education data systems;
- Provide timely information to pre-k-12 teachers and principals on individual students and their academic...
backgrounds so that instruction can be tailored to their needs;

• Expand agreements to share data responsibly with other state agencies so that data systems include information from a fuller range of education and social service programs;

• Develop guidelines and professional development programs on the responsible use of and security of data to ensure the privacy of student information;

• Collect student-level information (where permitted) on children enrolled in federally funded programs such as Head Start, Early Head Start, Even Start, Title I, IDEA, and those funded by Child Care and Development Block Grants;

• Incorporate data on kindergarten, first, and second grade assessments so that districts can track student progress in the early grades and identify effective early childhood programs;

• Collect information about early childhood educators’ credentials, the pre-service training they received and their participation in professional development programs, and use that data to help identify the knowledge, skills, and ongoing support teachers need to be effective;

• Include Quality Rating & Improvement Systems (QRIS) information about centers and programs providing early childhood education;

• Allow the public to access aggregated data on the long-term success of children who attended early childhood programs;

• Ensure that educators have access to data about their students’ early childhood education experiences and that early childhood educators have information about students’ achievements in later grades;

• Assist school districts in collecting more complete attendance data so that districts and the state can use the information to guide future funding and policy decisions to address chronic absenteeism more quickly; and

• Ensure that researchers have access to longitudinal, unidentifiable, student-level information to conduct research on the effectiveness of programs from birth through the third grade, including not only state-funded pre-k, Head Start, and subsidized childcare, but also home visitation, parent engagement, and other social services on children’s success in school.

For there to be meaning behind states’ assertions that their longitudinal data systems “connect to early childhood,” they must take steps to gather and integrate information from many early childhood programs, including Head Start, childcare, and home visitation programs. We recognize that doing this – and doing it well – will take an enormous amount of coordination, collaboration, and consultation with privacy and data systems experts. But this is the only way to meet demands for quality data that can help teachers in the early grades and early years answer questions about the types of experiences their students have had; drive principals to understand the impact of early childhood programs on a student’s success in later grades; provide parents with details about their child’s progress; and enable local school districts and state officials to make tough decisions about when and where to direct their investments in early childhood education programs.

This report has been made possible by generous grants from the Foundation for Child Development, the W. Clement and Jessie V. Stone Foundation, and the A.L. Mailman Family Foundation.
Notes

2 Data Quality Campaign, Creating a Longitudinal Data System: Using Data to Improve Student Achievement (Washington DC, 2006).
5 Ibid.
9 Ibid.
11 Ibid.
13 Ibid
14 According to the grant, these systems are intended to enhance the ability of States to efficiently and accurately manage, analyze, and use education data, including individual student records. The data systems developed with funds from these grants should help States, districts, schools, and teachers make data-driven decisions to improve student learning, as well as facilitate research to increase student achievement and close achievement gaps.
19 The other three “assurances” are standards and assessments, great teachers and leaders, and turning around the lowest-achieving schools.
21 Barack Obama, “Remarks by the President to the Hispanic Chamber of Commerce on a Complete and Competitive American Education,” (Washington DC, March 10, 2009.)
25 2009-10 Survey Results, Action 1, Part 1, Question 89, Data Quality Campaign.
28 “Linking Data across Agencies: States that are Making It Work.”
31 Interview with Jeff Sellers, Acting Deputy Commissioner of Florida (July 17, 2010).
32 Ibid.
33 The State of Preschool 2009, p. 47.
34 2009-10 Survey Results, Action 1, Part 1, Question 92, Data Quality Campaign.
35 Ibid.
37 Sarah D. Sparks, “Race to the Top Winners Face Data System Challenges,” Education Week, Vol. 30, Issue 3, Page 23, (September 15, 2010.)
38 ABC provides places for 3 to 5 year-olds in families with incomes below 200% of the federal poverty threshold. There are 318 partnering agencies, including 145 school districts and cooperatives and 173 private providers in either center-based or home-based sites.
46 Ibid, p. 20.
50 Ohio Department of Education, p. 19.
52 Ibid, p. 2.
54 Ibid, p. 15.
55 Ibid, p. 11.
63 Ibid.
64 Sara Mead, Education Reform Starts Early: Lessons from New Jersey’s PreK-3rd Reform Efforts, New America Foundation, December 2009, p. 34.
66 Quality Rating and Improvement Systems rank the quality of early learning programs and provide coaching to teachers and administrators to help improve them.
This report carries a Creative Commons license, which permits re-use of New America content when proper attribution is provided. This means you are free to copy, display and distribute New America’s work, or include our content in derivative works, under the following conditions:

Attribution. You must clearly attribute the work to the New America Foundation, and provide a link back to www.Newamerica.net.
Noncommercial. You may not use this work for commercial purposes without explicit prior permission from New America.
Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under a license identical to this one.

For the full legal code of this Creative Commons license, please visit www.creativecommons.org. If you have any questions about citing or reusing New America content, please contact us.