



## SLDS Spotlight

### Mississippi's Approach to Building a P-20W Data Model

Nationwide, states are expanding K12 (kindergarten to grade 12) education data collection to include pre-kindergarten through higher education and workforce data, known as P-20W data. The state of Mississippi developed a P-20W data model to guide the development of a statewide longitudinal data system (SLDS). In an SLDS Grant Program webinar, representatives from Mississippi State University's National Strategic Planning and Analysis Research Center (nSPARC) outlined a process for building a data model that integrates P-20W data for the purpose of tracking education outcomes.

#### Step 1: Know Why You Want Longitudinal Data

As education policy decisions become increasingly data driven, there is a greater demand for analyzing not only K12 education data, but also looking at longitudinal data over a greater time span, from early childhood to postsecondary education and the workforce, in order to answer policy questions about education quality and outcomes. Before undertaking a P-20W data project, it is necessary to define the conceptual framework to determine what data are needed to answer policy questions and inform decisions made by stakeholders across the P-20W spectrum. Mississippi's conceptual framework is shown in Figure 1 (on next page). The framework, which depicts students' flow through the education system and into the workforce, guides the development of a detailed data model that becomes the blueprint for an SLDS project.

Mississippi spent several years visiting with agency partners to develop the conceptual framework and to determine how the framework could help improve the quality of educational outcomes beyond the agency level. The state also developed pilot projects in which participating agencies were asked to share sample data, allowing the state to demonstrate how the framework would guide the process. The result of the pilot project was used to show the benefits of the conceptual framework to the other partners, ultimately bringing all of the education and workforce partners on board.

#### Step 2: Find Out Where Data Are Located

All of a state's K12 data may be stored in one place, but that is unlikely to be the case for early childhood, postsecondary, and workforce data. Begin by identifying the agencies that already collect these data. Mississippi is currently gathering early childhood education data from agencies such as the Mississippi Department of Education, Head Start, and the Mississippi Department of Human Services; postsecondary education data from Institutions of Higher Learning and the Mississippi Community College Board; and workforce data from agencies such as the Mississippi Department of Employment Security and the Mississippi Development Authority.

#### Step 3: Determine What Data Are Available

After identifying agencies that have data relevant for the purposes of building an SLDS, the next step is to determine more specifically what types of data each agency collects. In order to do this, Mississippi first created an Agency Program List for every agency from which they were obtaining data. The Agency Program List is a spreadsheet that includes information on all of the programs offered by that particular agency. Information collected includes program contacts, types of services offered, and

This product of the Institute of Education Sciences (IES) was developed with the help of knowledgeable staff from state education agencies and partner organizations. The content of this brief was derived from a Statewide Longitudinal Data Systems (SLDS) Grant Program monthly topical webinar that took place on June 17, 2011. The views expressed do not necessarily represent those of the IES SLDS Grant Program. We thank the following people for their valuable contributions:

#### Webinar Presenters:

Domenico "Mimmo" Parisi  
Shon Myatt

*Mississippi State University National Strategic Planning and Analysis Research Center (nSPARC)*

#### Moderator:

Rosemary Collins  
*SLDS Grant Program*

*For more information on the IES SLDS Grant Program or for support with system development, please visit <http://nces.ed.gov/programs/slds>.*

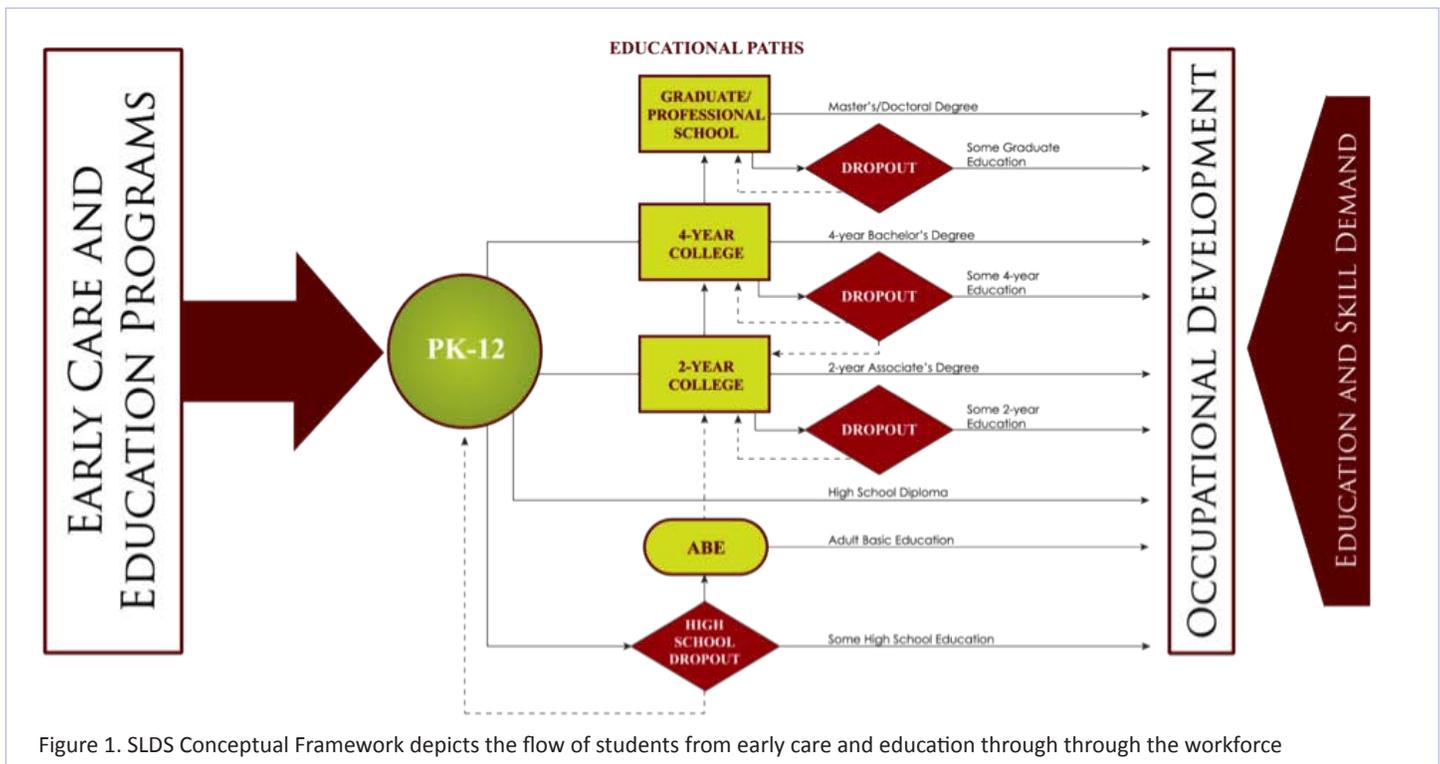


Figure 1. SLDS Conceptual Framework depicts the flow of students from early care and education through through the workforce

eligibility requirements for program participation. An example for the Mississippi Department of Human Services (MDHS) can be found at [http://nces.ed.gov/programs/slds/pdf/Program\\_MDHS\\_Example.pdf](http://nces.ed.gov/programs/slds/pdf/Program_MDHS_Example.pdf).

#### Step 4: Identify and Document the Data

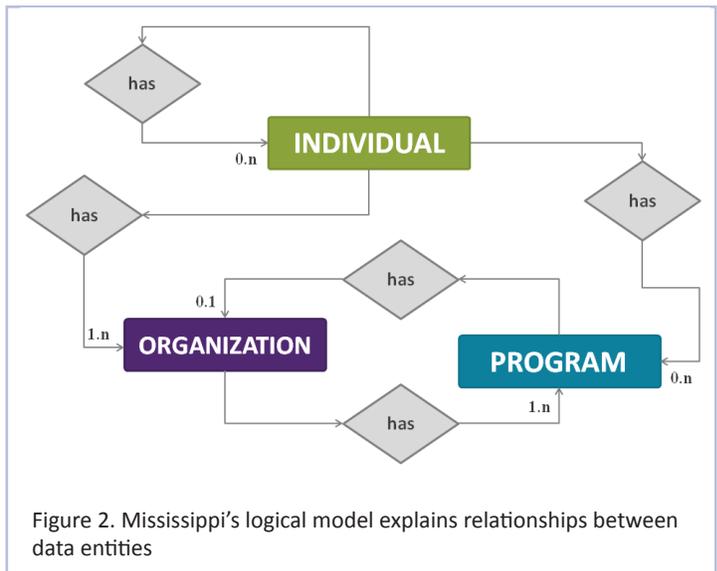
Next, Mississippi created a data dictionary (referred to as the “codebook” in Mississippi) for every program within an agency. The data dictionary provides basic information describing the data source. As an example, see the data dictionary developed for one of the MDHS programs called Mississippi Child Care Resource and Referral Network Professional Development Tracking System at [http://nces.ed.gov/Programs/SLDS/pdf/R&R\\_PDTs\\_Codebook.pdf](http://nces.ed.gov/Programs/SLDS/pdf/R&R_PDTs_Codebook.pdf). For each data element the program collects, the data dictionary includes its name, the type of data (e.g., character, numeric, etc.), the space allotted for each piece of data (field length), associated values (e.g., age group categories), and any relevant comments.

After all program data elements are documented, it is necessary to compare them across programs. To accomplish this, Mississippi created a Program Data Mapping spreadsheet, which compiles information from all of the data dictionaries. This mapping process provides an efficient method of determining the precise location of each data element; it also links entities and common data elements, which allows sorting of the data. After mapping the programs in one agency, programs for the other agencies can be added. Although mapping is time-intensive, the benefits of data linking are clear, as it permits cross-referencing information within a program as well as across programs. Essentially, mapping involves categorizing data elements and defining their relationships in a way that facilitates locating data across programs, helps to understand how various data elements are connected, and provides an easy way to sort data by desired categories. As an example of a data mapping solution, see an excerpt from a spreadsheet created for MDHS data at [http://nces.ed.gov/programs/slds/pdf/MDHS\\_PDM.PDF](http://nces.ed.gov/programs/slds/pdf/MDHS_PDM.PDF). To better understand the power of program data mapping, watch five minutes of the presentation that focuses specifically on this topic, starting at time 7:40 of the *recording*. (Note: The Common Education Data Standards offers an Alignment Tool to help education agencies map their data dictionaries to those of other organizations. This free resource greatly simplifies the process described here.)

#### Step 5: Create a Logical Model

Once all available data elements and their sources have been identified, the next step is to create a logical model explaining the relationships between data elements. As depicted in Figure 2, Mississippi’s model links data using three overarching categories (or entities): Individual, Organization, and Program. An individual could be a parent or a child, an organization could be a high school or place of employment, and a program could be K12 enrollment or workforce course/training.

Linking can occur in multiple directions, which are illustrated by arrows in the model. For example, an individual can be linked to another individual (e.g., a student and teacher), or an organization could be a part of a program, etc. All of the linking in Mississippi is done through personal identification numbers (IDs), which are generated using social security numbers (SSNs), which are generated using social security numbers (SSNs). When SSNs are unavailable, personal IDs can be created based on a combination of other personally-identifiable data, such as birthdate, race/ethnicity, gender, and age. (Note: CEDS Version 2 provides a logical data model and associated tools to assist states and other organizations through the process of creating their own data models and subsequent implementations based on those models.)



## Step 6: Build and Populate the Physical Model

The next step is to build a physical model based on the logical model. This includes creating a database where all of the data will be stored. After the infrastructure is set up, it can be populated with all of the data that has been organized. Be conscious of data quality—the SLDS team should have a plan in place to make sure the data are as accurate as possible. All data included in the Mississippi SLDS are first sent to a clearinghouse, where they are cleaned and all personally-identifiable information is removed. To facilitate the process of populating the model with data, Mississippi has coordinated and worked closely with their partnering agencies. Memoranda of understanding guide data sharing, and each agency sends data to the clearinghouse according to an established schedule.

## Step 7: Build a Business Intelligence System

In order for users to access an SLDS system and its data, a user interface needs to be developed. The SLDS team should determine what types of reports will be generated initially, but also design the system to be flexible enough to accommodate changes in the reporting requirements. Allowing access to the system requires regulation—each state must determine issues of data governance, including who can access what data, for what purposes, and how data can be used. A data governance program should include policies and regulations ensuring that all data usage and sharing is conducted in compliance with applicable federal, state, and local privacy laws.

## Step 8: Put the Data to Good Use

Once the P-20W data model is built, the information collected should be used to its full benefit. Provide the data back to the users in a way they can readily use, and tailor training to meet the specific needs of various users. Once users are properly trained to use the tools and understand the data, the data can be used to inform research and decisionmaking at all levels of the education system.

In Mississippi, P-20W data are primarily being used in postsecondary settings. For example, the data have been used to generate Perkins reports by linking career technical education records from Mississippi Community Colleges with records from the Mississippi Department of Employment Security. Similarly, Mississippi's institutions of higher learning have used the system to measure workforce outcomes by institution and field of education.

### Additional Resources

SLDS webinars are available at <http://nces.ed.gov/Programs/SLDS/webinars.asp>. The Common Education Data Standards (<http://ceds.ed.gov>) offers an Alignment Tool and a Logical Data Model—free resources that greatly simplify the processes of P-20W data mapping and modeling (as described in steps 4 and 5 of this document) and the development of subsequent physical implementations. The National Forum on Education Statistics publishes best practices guides that provide additional information on many of the topics included in this webinar. For more information on longitudinal data systems, education data management, and education data privacy laws, see the list of Forum publications at <http://nces.ed.gov/forum/publications.asp>.