

SLDN Landscape and Recommendations

Allison C. Bell, Ph.D.
Principal
Three Arrows Up Consulting

Prepared for:
RTI International
3040 E. Cornwallis Road, PO Box 12194
Research Triangle Park, NC 27709
www.rti.org

November, 2022



Contents

Student-Level Data Network Landscape and Recommendations	1
General Observations	1
Current Landscape of SLDS.....	2
Recommendations for a Strong SLDN	3
Legislate the data system and ensure reliable funding source	5
Be clear about the purpose and audience.....	5
Involve a variety of audiences and demonstrate the value of the effort	6
Create strong governance and privacy protections	7
Build upon a foundation of existing data systems and efforts.....	8
Use common reporting standards	10
Provide flexibility for changing needs and contexts	11
Promote use of the data by various audiences	12
Other Points Brought Up in Interviews.....	14
Concluding Thoughts	15

Appendix. Data Availability for Select States

Student-Level Data Network Landscape and Recommendations

Since 2005, the Institute of Education Sciences at the U.S. Department of Education has awarded competitive grants to help states design, create, maintain, and expand statewide longitudinal data systems (SLDS). SLDS contain individual-level data on students from early education through the workforce, largely in response to a ban on the federal government from creating such a data system. In February of 2022, the U.S. House of Representatives recently passed the College Transparency Act (CTA), which would reverse this ban. As a result, the development of a federal student-level data network (SLDN) has become a real possibility. As states have developed and maintained student-level data systems for decades, there are lessons that these efforts can offer for such a network.

This brief provides recommendations based on lessons from state experiences with SLDS design, development, and implementation. These include examples of what states have done well in the design and implementation of their data systems as well as challenges and setbacks they have faced. These recommendations are based on a combination of:

- Conversations with experts and state representatives who have had a direct role in designing, developing, implementing, and/or maintaining longitudinal data systems. This included more than 20 representatives from associations, states, and other organizations, such as the State Higher Education Executive Officers Association (SHEEO), the Education Commission of the States (ECS), the Council of Chief State School Officers, the National Center for Higher Education Management Systems, the Data Quality Campaign (DQC), Arkansas, California, Kentucky, Mississippi, New York, Pennsylvania, Tennessee, Virginia, and the United Kingdom.
- Document and website review, including state data system websites, opinion papers, and research on SLDS.
- Experience and expertise with state data systems, including development, implementation, policy, and use.

General Observations

In addition to the specific recommendations for designing and implementing an SLDN, the following observations from the conversations with representatives provide context:

- Across the conversations for this project, there was not a consistent and clear understanding of the CTA, what is included in an SLDS (e.g., some conflated SLDS and state postsecondary student unit record systems [PSURS], others used SLDS to speak mainly of K-12 data systems—and, in fact, many state SLDS sites focus heavily on K-12, and others maintained that an SLDS consisted of multiple state agencies), or what the purpose of individual-level longitudinal data systems are. This points to a need for clear, concise, and specific communications for different audience types.

- Blockchain technology was brought up multiple times as emerging and important.
- To track migration of their students and to fill the lack of a national student-level data system, states have begun to connect their data systems across states with interstate data sharing agreements. Such examples include [Western Interstate Commission for Higher Education's \(WICHE's\) Multistate Longitudinal Data Exchange](#)¹ (MLDE), and the different members of the [Coleridge Initiative](#),² which promotes interstate data sharing via the Administrative Data Research Facility (ADRF).
- Several respondents talked about the lateness of the development of a federal system in comparison to the progress that states have made over the last 10 years especially.
- There are many different perspectives on best practices and weaknesses. For example, one state was identified by multiple people as an example of a leader in data system development, implementation, and use that has devolved into a cautionary tale (without always specifically describing why), whereas others point to the state's system as an example of a data system that should be replicated.
- Regardless of the presence and strength of an SLDS, such systems are still considered by many as necessary because state SLDS might not include all types of institutions. Many states' data systems do not include proprietary institutions or other nonpublic institutions. Even when nonpublic institutions are included in a data system, they may not report all data elements. A federal data collection system would include all institutions and allow for more complete comparisons and research.
- The successful implementation of an SLDN will require a careful balance of competing interests.

Current Landscape of SLDS

Over the past several decades, almost all states have developed student-level longitudinal data systems at some level:

- Forty-nine states (all but Delaware, where one is under development) have a PSURS of some kind. These PSURS may be maintained by multiple agencies and may not include all students ([SHEEO State Postsecondary Data](#)).
- In 34 states, there are linkages between K-12, postsecondary, and workforce data ([SHEEO State Postsecondary Data](#)).

¹ WICHE's MLDE is a data sharing agreement that includes six states: Hawai'i, Idaho, Minnesota, North Dakota, Oregon, and Washington. Participating states can request information on individuals through the exchange. The data are shared in an encrypted and deidentified format.

² The Coleridge Initiative is a project that started at New York University and is an effort to share education and workforce data across states. From its website: The Administrative Data Research Facility (ADRF) is a secure cloud-based computing platform designed to promote collaboration, facilitate documentation, and provide information about data use to the agencies that own the data. The ADRF was established by the Census Bureau with funding from the Office of Management and Budget to inform the decision-making of the Commission on Evidence-Based Policy.

- Forty states have an individual-level data system that contains data from at least two of the following core agencies: early learning, K-12 education, postsecondary education, and workforce ([ECS](#)).
 - California, which previously had several different data systems, is developing an SLDS that will include each of these agencies.
- In more than half of the states with a PSURS, there is an agreement for researchers external to the host agency to access those data ([SHEEO State Postsecondary Data](#)).
- Without a national data system, many states have engaged in interstate data sharing efforts to better understand what happens to students who move across state boundaries. Such efforts include the MLDE and Coleridge Initiative (as described above). The MLDE consists of six states (Hawai'i, Idaho, Minnesota, North Dakota, Oregon, and Washington). More than 250 state agencies, institutions, and organizations in 42 states collaborate in some way with the Coleridge Initiative. Both MLDE and the Coleridge Initiative include education and workforce data.

Recommendations for a Strong SLDN

When asked which states have the strongest data systems, respondents often identified Kentucky, North Carolina, Texas, and Virginia as exemplars.³ These noted exemplars have data systems that were legislatively mandated; include nonpublic institutions; link data across sectors; have clear purpose; involve multiple audiences in governance and design decisions; have clear and robust governance and privacy protection policies and procedures; and are used regularly by policymakers, institutions, researchers, employers, jobseekers, and students. Without exception, these data systems are supported by state leaders. A strong SLDN would have many of these same qualities.

Note: California, which was identified as a state that had been weak in terms of not having a cohesive SLDS, has recently committed time, resources, and energy to designing an SLDS that has the potential to emerge as a strong system (although this will take some time). The work of designing this data system has been based on lessons from other states' SLDS development. (See the exhibit "California: Learning from Other States" for a summary of California's recent work.)

³ These states may not all have a central data warehouse, but rather have linkages between sectors. North Carolina, for example, is a state that is working on improving its SLDS as a centralized system that provides statewide data, while the University of North Carolina (UNC) and community college PSURS systems are individually strong and are linked to K-12 and workforce data. Other states that are often identified as those with strong SLDS include Florida, Georgia, Maryland, Minnesota, and Washington. One of the reasons these states are "strong" rather than "exemplars" is reputation. Also, these states tend to be strong at one level or in one area (e.g., Maryland's SLDS is stronger in its K-12 than its postsecondary data elements and use), whereas the exemplars demonstrate all-around strength.

California: Learning from Other States

California had long been identified as a state with a weak data infrastructure. Although the state has had student-level data systems at the postsecondary level, they existed at each system separately with little opportunity for linking them. And, while each system linked with those of other agencies, it was impossible to answer some basic questions about California's postsecondary students, especially when students transferred between institutions.

In 2019, after years of discussion and advocacy, California passed legislation to create a state centralized data system that includes K-12, postsecondary, workforce, and other agency data (see information on the planning process [here](#) and the implementation website for the data system [here](#)). Still in its implementation stages, this new data system has had the opportunity to implement the best practices learned from the implementation of other systems across the country. Some of the data system's features that experts are pointing to as best in class include the following:

- Codification of the data system in legislation.
- A budget (\$15 million was allocated in the 2021–2022 budget in addition to the \$10 million that had been appropriated for the system's planning stages, and the governor is asking for \$33 million for 2023-24) that, based on the best available information, is sufficient for the first year of implementation.
- Workgroup design that included stakeholders from all potential participating sectors and legislators.
- A series of planning and design meetings that focused on governance, privacy, technical aspects, and more. These have involved more than 200 people in conversations central to the system's design.
- The design process has been based on use cases and has also involved expert advice, which the state has considered.
- The plans include public dashboards, researcher access, and a library of prepared research.
- Two different advisory boards that will continue to provide advice to the governing board, with stakeholder outreach being a central component of the ongoing conversations.
- A phased approach for integrating different agencies' data over several years.

The data system is in the early stages of implementation, but all early indicators are that, if the blueprint is followed, California has the opportunity to be recategorized as a leader in SLDS.

Based on the success of these states and on the lessons learned from other states' challenges, several suggested recommendations for developing and implementing a strong SLDN are as follows:⁴

- legislate the data system and ensure a reliable funding source
- be clear about the system's purpose and audience

⁴ Some of these items might be addressed in the CTA, and specific recommendations and next steps for best addressing the requirements are included to underscore the importance of learning from states' experiences. Where possible, specific state examples are provided as support for each actionable step. In some cases, these are not the result of a specific state's success or misstep, but rather a recommendation born of the aggregate of respondents' experiences

- involve a variety of audiences in design and implementation conversations to help demonstrate value-added benefits to the current data environment
- create strong governance and privacy protections
- build upon a foundation of existing data systems and efforts
- use common reporting standards
- provide flexibility for changing needs and contexts
- promote use of data by various audiences

Legislate the data system and ensure reliable funding source

The most successful data systems are mandated by legislation and have a reliable funding source. Kentucky is an example of a state that has done this well. In 2012, the Kentucky Center for Statistics, which maintains the state's SLDS, was established through legislation. The legislation also stipulates that the agency may be funded through state appropriations, federal grants, user fees, and other grants or contributions from public agencies and other sources.⁵ Washington also passed legislation that established the agency that manages its SLDS and includes specific language that provides for state funding.

Establishing data systems through legislation can help secure the participation of all necessary data providers, help make explicit the governance structure and types of protections that must be in place, stipulate how the data can be used, and (as evidenced in Kentucky and Washington) specify dedicated funding. Although state budgets are often in flux and budget cuts can affect all state agencies, naming state sources as a funding source for data systems is a major step toward ensuring reliable funding sources for data system development, maintenance, and expansion.

Through the CTA, if passed, there will be legislation that establishes and describes the SLDN. Based on the language, which mandates that the data system be overseen by the National Center for Education Statistics (NCES), it is reasonable to conclude the system would be included in the agency's budget.

Be clear about the purpose and audience

Almost every person consulted on this subject began and ended their comments with the statement that clarifying the purpose of an SLDN would be extremely important. This can be accomplished through the following:

⁵ Section 151B.132 (Office of the Kentucky Center for Statistics).
<https://kystats.ky.gov/Content/KRS151B.132.pdf?v=20220801031302>

Developing use cases for the data that will be collected. When the DQC first started, the recommendations and advocacy platform were built with use in mind. DQC started with 10 essential elements of state data systems and then later developed 10 state actions. The clearly articulated elements and actions provided states with specific steps and a guide for designing their data systems.

Many states have use cases that are integral to their design process. New Jersey has specific “user cases” (fictional user profiles) to guide the state as it builds out its job seeker tools. These user cases help to facilitate discussions on user interface, what data is necessary to include, and communication efforts. As California has been developing its system, it has relied extensively on use cases.

Making explicit the specific questions that the data system will be used to answer. An SLDN should be able to answer specific, relevant questions, which should be clearly communicated. For example, the [MLDE based its design on specific questions that could be answered through connecting state postsecondary data systems.](#)⁶

Making explicit the primary audiences for the data. The development of an SLDN should include a clear statement that explicitly identifies the primary audience for the data collected. Many of the individuals consulted for this project made clear that data should not be collected for the sake of advocacy groups’ interests and stated that policy decisions should not be made based on “what would be interesting.” At the same time, many of the respondents indicated that use of the data, including by researchers of all types, will help to make the data of higher quality and more relevant, and will provide insights that may not otherwise surface.

In the case of an SLDN, the primary audiences for these data are

- federal staffers who need this information for purposes of Title IV operations
- policymakers at the federal, state, and institutional levels
- consumers (students, parents, and families)
- researchers
- postsecondary institutions (for benchmarking and analysis)

Involving each of these audiences and balancing their needs should be a central focus during SLDN development.⁷

Involve a variety of audiences and demonstrate the value of the effort

Audience involvement is another common theme in strong state data systems and can help build trust, ensure that the data reported are robust and accurate, engender support that can

⁶ Although not itself a state data system, the MLDE provides helpful lessons for the implementation of a data initiative. The linked report details some of the lessons and considerations from the MLDE experience.

⁷ This work is ongoing and it is difficult. The important takeaway for SLDN development is that the needs versus the requests of different audiences should be carefully navigated. The Department of Education has experience with involving different audiences in the design and maintenance of data systems. This may be a good opportunity to also review those practices and processes.

help facilitate sustainability of the system, and promote the ability for the system to evolve. One of the best ways to elicit audience buy-in and continued involvement is through the ability of system leaders to show the value-added benefits of a system. Given that there are already federal data collection efforts underway, involving various interested parties and articulating the value of the federal SLDN will be important. Specific steps include:

Involve a broad base of audiences in initial and ongoing design discussions.

California has learned from other states' implementation experiences and has involved a broad base of stakeholders from the outset in design discussions. The ongoing plan for the data system design continues to include broad-based stakeholder outreach and inclusion. In addition, California took the extra step of putting legislators on the governing board. This was done with the understanding that, ultimately, the system is only as good as the laws and funding impacting it, which are in the hands of the state legislators.

As an SLDN is designed and implemented, it will be important to include as broad a group of stakeholders as possible (across politics, geography, sectors, etc.). If decision-making is shared and broad, organizations can see themselves in it and will be more supportive of the effort. These efforts are underway with the expert panels hosted by the Institute for Higher Education Policy and RTI International. To continue to build the community and trust around an SLDN, this early work should be included in future discussions and the outreach should continue.

Offer explicit examples of what the value-added benefits of the SLDN will be, above and beyond current postsecondary data collection efforts.

It is important, especially as there is already a federal data collection that includes all institutions (the Integrated Postsecondary Education Data System [IPEDS]), that the value of the SLDN is clearly articulated. Kentucky, Maryland, and Washington are all examples of states that have clearly articulated the value of their SLDS.

Be transparent. According to DQC, and as seen in various types of data collections, transparency with stakeholders is key to building trust. When groups of stakeholders are in the dark and not part of the decision-making process, data systems and mandates are perceived as “something that is being done to them” rather than something they are part of. NCES needs to bring people along as part of the process, which will include

- communicating throughout all stages of the development process
- involving a broad base of stakeholders in decision-making
- being clear about how and why decisions were made and communicating about that

Create strong governance and privacy protections

A strong data system is clear about its governance and privacy protections, which go hand in hand. As written, the CTA legislation gives cursory acknowledgement to these issues. In developing an SLDN, the specifics of how the data will be governed and protected will need to

be clarified and codified. This is particularly true as audiences may not have full trust in the federal government to act in the best interest of *all* postsecondary institutions and their students. An SLDN will need to protect students while ensuring the data can be used.

Be clear about who “owns” the data that are provided as part of an SLDN and how those data can be used. Currently, there are questions about how the data that are reported will be used, who will have access to the data, and how the data will be shared with providers. A clear statement about those issues will help to assuage any concerns about data governance. States such as [Colorado](#) and [Maryland](#) clearly outline these considerations. In both examples, data governance is overseen by a committee.

Ensure adequate protections for students. The CTA legislation, in the view of many respondents who were interviewed as part of this project, does not provide adequate protections for students. One recommendation was to mirror the types of protections the U.S. Census Bureau has on its data—basically mandating that the individual-level data cannot be released to anyone. Another recommendation was to make clear what the federal government could and—more importantly) could not use the data for.

It is possible to protect identifiable data and still allow for data use. In Kentucky, researchers can apply for access to state data. If requesters are granted permission to use individual-level data, they are provided encrypted, de-identified data. The Department of Education already has policies and processes in place that allow researchers to access individual-level data through restricted-use licenses. The Department of Education also has significant experience in providing researcher access to public use files through data tools (e.g., PowerStats) that allow sophisticated analysis.

Build upon a foundation of existing data systems and efforts

Some SLDS have been in existence for decades. Institutions of higher education have been reporting to the federal government going back as far as the 1930s. In developing an SLDN, it is imperative to recognize that it is building on a foundation of many decades of state- and federal-level data collection.

Recognize that there are valuable lessons to be learned. Throughout conversations about a potential SLDN, one of the recurring themes was that people wanted to be recognized for the data system work that has already occurred and be assured that even though something new is being built, states and institutions will not be marginalized. This can be accomplished through actively seeking out states to share their experiences and lessons. Even though this may seem like a simplistic step, the public acknowledgment of previous experience and lessons will be significant in easing state and institutional anxiety.

Ensure that states are involved in the conversations. Each of the states interviewed indicated that an SLDN would serve different needs than their SLDS and therefore could

not replace the state longitudinal data system. As such, states should be instrumental in discussions that can inform how an SLDN can be complementary to existing state efforts. One major characteristic of an SLDN that provides value over an SLDS is the compulsory inclusion of all institutions, which will allow for important questions comparing public and private institutions to be addressed. As a result of including all institutions, an SLDN would collect data on all students and could potentially provide information on students who move across state lines.

Leverage existing federal data systems to reduce duplication of effort and institutional burden. The CTA specifically names existing federal data systems with which data sharing agreements should be established. Where possible, these data sharing agreements should be used to reduce institutional reporting burden and help ensure data quality. The specific data that need to be collected from outside of institutions includes the following:

- ***Earnings data.*** The CTA calls for earnings data through data sharing agreements with several agencies. The Internal Revenue Service (IRS) has millions of records of data on income and earnings that could be used to provide program-level earnings information. Additionally, the Census' Longitudinal Employer-Household Dynamics program's Post-Secondary Employment Outcomes provides earnings information and has data sharing agreements with several states. The two data sources are similar, although the IRS data is more complete as it includes self-employed individuals and is based on tax information rather than employer-supplied data. In addition, the CTA language calls for data sharing agreements with the Social Security Administration for the labor market outcomes of former postsecondary students (data elements would include retirement, disability, and insurance information) and the Bureau of Labor Statistics for wages (which is distinct from earnings, which encompasses all types of money an individual receives [investments, alimony, other benefits, etc.]) of former postsecondary students and could include work stoppage data.

As data sharing agreements are drafted for an SLDN, it is possible that the inclusion of these data sources could help to relieve institutional burden for those that are collecting these data through alternative means (e.g., alumni surveys). Critics maintain that including these data could lead to a breach of privacy or student data. As a result, the specific need for the data sharing agreement with each of these entities, how the data will be used, and how privacy will be protected will need to be clearly stated and communicated.

- ***Data on service members and veterans.*** Both the Department of Defense and the Department of Veterans Affairs track information on active-duty military and veterans' education and education benefits that postsecondary institutions may not have direct access to. Benefits such as the GI Bill are issued directly to

students and not tracked by institutions as financial aid. Additionally, these benefits can be used at institutions that do not currently report to IPEDS.

- ***Students' federal financial aid data.*** The Office of Federal Student Aid (FSA) maintains the National Student Loan Data System, which contains information on all students who receive federal student aid. Even though institutions will likely have this information, it could be that a data sharing agreement with FSA could reduce the overall number of data elements an institution would have to report.

If possible, use the SLDN to reduce the reporting burden for other data collection efforts. At the postsecondary level, the federal government already collects data from each institution (in IPEDS) and from students through the sample surveys (the National Postsecondary Student Aid Survey, Beginning Postsecondary Students, and Baccalaureate and Beyond). With an SLDN, it is possible that the effort that goes into these data collections could be reduced. In the case of IPEDS, there should be no (or very little duplication) between IPEDS data elements and those reported to an SLDN. In addition, there are several data collections that are not administered by the federal government but which contain much of the same information. These include Complete College America, Achieving the Dream, and the Postsecondary Data Partnership. It is possible that the development of an SLDN could help to reduce the burden on states of reporting to these efforts.

Where possible, take advantage of existing Department of Education initiatives.

There are two Department of Education initiatives that might have particular relevance for an SLDN:

- [Education Blockchain Initiative \(EBI\)](#). This initiative, launched in partnership with the American Council on Education in 2020, focuses on understanding how blockchain technology can facilitate the “secure, traceable, and verifiable exchange of educational data among institutions in the learning and employment ecosystem.” The goals of the EBI are aligned with those of a potential SLDN and perhaps there are areas of collaboration, especially where there are needs for secure data dissemination.
- [Digital Strategy Initiative](#). Although this initiative has a website and information available publicly, it is less clear what the actual work has been. NCES should investigate whether there are clear areas for collaboration with this initiative.

Use common reporting standards

Currently, data collections at the state and federal levels struggle with standardization so that comparable data are being reported. Using common reporting standards will help to reduce institutional burden and can improve data quality. Specifically, as an SLDN is implemented, it is important to:

Base data definitions on existing definitions. Institutions are used to the definitions in IPEDS and other federal data sources. Wherever possible, state representatives would like to keep consistency with those, or at least define data elements in a way that will align with key data elements. For example, in Kentucky, postsecondary data elements collected in the statewide data system are those that are also reported to IPEDS.

Additionally, the Common Education Data Standards (CEDS) is an ongoing project that has more than 550 postsecondary data elements. The Institute for Higher Education Policy recently released [a report on recommendations for integrating CEDS into the SLDN](#). North Carolina, which has a strong SLDS, is an example of a state that relied on CEDS as a foundation for its data dictionary and definitions ([as evidenced in this 2015 case study that includes Virginia](#)).

Clarify the process for reporting the data. Data providers want to know how the data will be reported—this is often a concern voiced in discussions about SLDS development as well as in discussions about how data currently get reported, so the issue is not unique to individual-level data collections and is one with which NCES is familiar. There are several options for how data are to be reported—through the state, through system offices, or through individual institutions. During the development of the reporting standards for an SLDN, the implications of each method of reporting should be carefully considered.

Be clear about the contacts for questions and guidance. In Virginia, leadership has taken a case management approach to administering the data system. As a result, it is clear to institutions who they should contact if there is a question and they have direct access to resolve issues with the data. In some states, IPEDS state coordinators have been fulfilling this role. However, different people may give different advice. A successful SLDN should offer clear guidance to people who might field questions in states and provide access to support personnel who can assist with clarifying and resolving discrepancies.

Provide flexibility for changing needs and contexts

State systems that have been able to evolve with the changing postsecondary landscape are those that have been the most useful in providing answers to important questions and helping to guide policy decisions. While SLDS are often thought of as static structures, states have been able to incorporate design elements into their systems that allow them to be flexible. Specific recommendations based on state experiences and practices include the following:

Build in fields that allow for institutions and states to customize their data report. In the Arkansas state data collection, there are empty fields that allow institutions to add custom data elements. In addition, the system office can add custom data elements. This ensures that the data reported can serve the needs of multiple audiences.

Allow for changes in individual demographic reporting. Not all demographic characteristics are static. Throughout the course of their educational careers and lives, students may identify with more than one gender, race, or ethnicity. A student is not legally required to choose demographic characteristics on intake forms or surveys that match their historical background. In Virginia, the data system is designed to handle this. Each student has a demographic record that is attached to their identity record, and that record is by institution attended. If a student answers multiple demographic questions—either by attending multiple institutions or by filling out the information multiple times—multiple demographic records are attached to the student’s identity record. The demographic characteristics that are appropriate for the time period for which the data are being used are the ones that are applied to the student.

Provide a way to capture data and match multiple records per student. In the current IPEDS, a student might be reported within the data of multiple institutions each year. Furthermore, because institutions might not have a complete picture of a student’s history, student status (e.g., first-time, non-first-time student) could be miscategorized. An SLDN should be able to clearly identify unique students (even across enrollment spells) and provide a method for matching students with multiple records (e.g., if a student enrolls in two institutions at the same time).

Promote use of the data by various audiences

Regular use of data reported through a data system serves multiple purposes. It makes the data system relevant, improves data quality, and ensures that the most time-appropriate information is available to decision makers (including students and their families). Regular use of the data provided through the SLDN can be promoted by:

Creating data dissemination tools that are appropriate for different audiences. The SLDN data users will have different needs and different levels of data literacy. NCES is accustomed to providing data for different audiences and that practice should be continued and enhanced with an SLDN.

States make their data available to audiences in different ways. In some states, such as Georgia and Kentucky, researchers can request de-identified individual-level data. The process for requesting data varies from state to state. In general, researchers need to provide a rationale for their request, describe how the data will be used, have their request reviewed, and agree to data protection measures. Many states create interactive dashboards and data tools that present aggregate-level data (e.g., Kentucky, Maryland, and Minnesota). Often these data are in the form of charts, tables, or maps. Finally, states produce static reports that use the data. One example of a state that does this is Washington through its Education Research and Data Center (ERDC).

Ensuring that data are as up to date as possible. A common request of data reported through IPEDS is that it be disseminated more quickly. If an SLDN is developed, one of the priority conversations should be around timely dissemination, knowing that data reporters and users would prefer that the data release not take longer than the current IPEDS data release schedule.

The exact timing of data release varies depending on the type of data (i.e., enrollment, completion, course taking behavior, and financial aid) and the state. Exemplar states release data tools and/or reports with updated data at least annually and may even update data quarterly as appropriate (usually this is enrollment data). In general, state data may be available earlier than IPEDS. For example, the Kentucky Council on Postsecondary Education has fall 2021 enrollment data available, whereas IPEDS only has fall 2020 data publicly available (via both “Compare Institutions” and “Look Up an Institution”). See the appendix for a sampling of states’ data availability compared to IPEDS.

When regular updates are not obvious, the reputation of the entire data effort may be questioned. For example, personnel changes at Washington’s ERDC have led to a lack of updates of at least the publicly available data and dashboards. One of the more recent dashboards was updated in [January 2021](#). More recent reports have been published, but the delay in updating the dashboards led to the perception that all the data was out of date.

Texas: Data Tools and Collaboration

Texas has a strong data infrastructure. The Texas Higher Education Coordinating Board collects data on students at all institutions and the data are linked across education and workforce. As a result, Texas can answer comprehensive questions about its students and provides data dashboards for multiple audiences and allows researcher access.

In addition, the University of Texas System is regularly lauded as a leader in data use. Its [SeekUT](#) is an exemplar student tool that connects data across education and workforce to give students comprehensive information on which to base their decisions. Texas was one of the [first states to collaborate](#) with the Census Bureau to give a more accurate picture of graduates’ post collegiate outcomes.

Other Points Brought Up in Interviews

Although not necessarily specific lessons learned from state data systems, the following advice arose from research conducted for this project:

- Prior to beginning any specific conversations about the design and implementation of an SLDN, NCES should create a roadmap for the structure and progress of those conversations (California is an example of how those might be structured).
- NCES should consider the use of pilots to implement the data system as a whole but also to try to answer some of the specific questions about how the SLDN should be administered.
- There are several philosophical issues that NCES needs to address. One is the tolerance for “dirty” data. Collecting quality, robust data that will allow questions to be answered comes with a cost. Interviewees acknowledged some of the data quality challenges with IPEDS (misidentification of first-time students, differences in census dates) and expressed a preference for “doing it right,” in terms of collecting data for an SLDN. This could require high time and resource costs. A common theme in conversations was that *decisions should not be made on what is easiest and most cost-effective to collect*. Additionally, state representatives and others noted that this is a good opportunity to create the ideal data system, rather than replicating issues with IPEDS at the student level.
- Several of the interviewees, including DQC, mentioned that an SLDN should serve K-12. Although the legislation prohibits the inclusion of K-12 students, it does specify feedback reports, which would benefit K-12 students. In implementing an SLDN, it will be important to clearly communicate that these feedback reports are being produced and what questions they will be able to answer.
- Especially those who will be instrumental in administering an SLDN should have an appreciation for the history of postsecondary data collection going back at least as far as the 1970s. Having a historical context will help to clarify what missteps federal data collections have encountered and help clarify how an SLDN can be most meaningful.

Concluding Thoughts

If implemented correctly, an SLDN could be of great value. States can provide important lessons for designing and implementing an SLDN, but ultimately the function and audience for an SLDN will be different from SLDS in nuanced but important ways. Just as with other federal education data collections, the design of an SLDN should serve its main purpose first (i.e., provide the Department of Education and students with relevant information to evaluate postsecondary education in this country as a whole and to better make informed decisions), protect students, and be responsive to changing contexts. Top priorities include the following:

- Clearly communicate about the SLDN throughout the development process. The progress and decisions made may not satisfy all audiences, but it will be important to include them in the process. It will also be necessary to clearly communicate how an SLDN provides value over the current data collection systems (IPEDS and the NCES sample surveys). Ultimately, buy-in and support for an SLDN will dictate its sustainability.
- Ensure that students are protected. Strong privacy and security policies and procedures are paramount. However, it will be equally important to ensure that the data are not being used in a way that can harm students who are included. For example, recent administrations' immigration policies have made some concerned that citizenship status could be used to identify students for potential legal action such as deportation.
- Be thoughtful about what data are reported and why. There are many different motivations for promoting an SLDN. It is not the federal government's job to provide all data for all needs.

Appendix. Data Availability for Select States

Data Availability for Select States and State Systems (as of August 2022)

State (state system)	Fall enrollment	Academic year enrollment	Degrees awarded	Graduation rates	Source	Notes
Arizona		2021–2022	2019–2020	2020–2021	https://www.azregents.edu/university-system-quick-facts	Four-year colleges only
California						
CCC	Fall 2021	2020–2021	2020–2021	Spring 2022 (success rate)	https://datamart.cccco.edu/datamart.aspx	
CSU	Fall 2021	2021–2022	2020–2021	2020–21	https://www.calstate.edu/data-center/institutional-research-analyses/Pages/reports-and-analytics.aspx	
UC	Fall 2021		2020–2021	2020–2021	https://www.universityofcalifornia.edu/about-us/information-center	
Colorado	Fall 2020		2020–2021			
Florida						
DOE	Fall 2021	2020–2021	2020–2021		https://www.fldoe.org/accountability/data-sys/CCTCMIS/reports.shtml	
BOG			2020–2021		https://www.flbog.edu/resources/data-analytics/dashboards/	Lack of data availability online
Georgia						
USG	Fall 2021	2021–2022	2021–2022		https://www.usg.edu/research/enrollment_reports	
TCSG	Fall 2021	2021–2022	2020–2021			
Illinois	Fall 2021	2021–2022 being collected	2021–2022 being collected	2021–2022 being collected	https://www.ibhe.org/datapoints.html (fall enrollment); https://www.ibhe.org/iheis.html	
Indiana	Fall 2020	2019–2020	2019–2020	2019–2020	https://www.in.gov/che/data-and-research/reports-and-analyses/	
Kentucky	Fall 2021	2021–2022	2020–2021	2019–2020 (fall 2013 cohort)	tinyurl.com/22v9rs5b	
Michigan		2020–2021	2020–2021	2020–2021	https://www.mischooldata.org/	
New York	Fall 2021			2019–2020	http://www.nysed.gov/information-reporting-services/higher-education-reports	

State (state system)	Fall enrollment	Academic year enrollment	Degrees awarded	Graduation rates	Source	Notes
North Carolina						
NCCC	Fall 2021	2021–2022	2021–2022	2021–2022	https://www.nccommunitycolleges.edu/analytics/dashboards	
UNC	Fall 2021	2021–2022	2020–2021		https://www.northcarolina.edu/impact/stats-data-reports/interactive-data-dashboards/	AY enrollment extrapolated based on term availability
Tennessee	Fall 2021	2020–2021	2020–2021	2020–21	https://www.tn.gov/content/tn/thec/research/fact-book.html	
Texas	Fall 2021			2020–2021	http://www.txhighereddata.org/	
Washington		2019–2020	2019–2020	2019–2020	https://erdc.wa.gov/data-dashboards/public-four-year-dashboard#about-dashboard	

Notes:

This table should be considered a rough guide for data availability for these states. This does not ensure states are able to report these data.

Not all states' reports contain data for all institutions.

AY = academic year; BOG = Board of Governors; CCC = California Community Colleges; CSU = California State University; DOE = Department of Education; NCCC = North Carolina Community Colleges; TCSG = Technical College System of Georgia; UC = University of California; UNC = University of North Carolina; USG = University System of Georgia.