Defining Modern, User-Centered State Longitudinal Data System Design

Actionable Intelligence for Social Policy

Data Quality Campaign

Education Commission for the States

WestEd's Data Integration Support Center

May 2025

Defining Modern, User-Centered State Longitudinal Data System Design

As states work to build stronger pathways to economic prosperity, they are confronted with a challenge: they lack access to information on the long-term outcomes of education programs and the factors that drive those outcomes. Policy makers, educators, and community members alike are seeking data that will help them identify effective strategies and sound investments, taking into account how postsecondary education, workforce training, public benefits, and legal system involvement affect the relationship between school and work.

While many states have linked data from various sources, few people have access to this information. This is unlikely to change as long as data modernization efforts are framed as technology projects, which are often perceived as expensive and ineffective. Unfortunately, there are few roadmaps for how states can move from closed data systems with limited access to systems that provide the information needed to help learners succeed on their education journey into the workforce.

To help states understand how to get the information they need to inform decisions, representatives from Actionable Intelligence for Social Policy (AISP), Education Commission for the States (ECS), Data Quality Campaign (DQC), and WestEd have developed a framework that spells out the opportunities provided by linked data sets and the actions states can take to attain those goals. For example, rather than focusing on traditional choice points like whether to build a federated or centralized system, states should ensure that purpose drives design. States should contemplate the services that allow people to access information and how they will engage those constituencies in the design of data systems. These priorities can then help determine the appropriate technical, legal, and governance frameworks to meet those needs.

This document summarizes how state education data systems are evolving and offers specific guidance that can inform the design of linked information and desired services.

The Evolution of State Longitudinal Data Systems

More than 20 years ago, the federal government developed a grant program intended to help educators better understand the long-term outcomes of students. Early work on statewide longitudinal data systems (SLDS) focused on following individual students through their K-12 years.

Education data systems gradually expanded to include early learning experiences, postsecondary education, and employment outcomes (sometimes referred to as P20W data sets). For example, Minnesota's Statewide Longitudinal Data System shows whether high school graduates are enrolled in college or employed, and how much money they make when they complete their education. Its related Early Childhood Longitudinal Data System shows the impact of early care on academic achievement in kindergarten and 3rd grade.

Increasingly, states are seeking to further expand their education data systems to evaluate how experiences like receiving public benefits, participating in workforce systems, and legal system involvement shift education and employment outcomes. These types of integrated data systems, for example, such as Washington's Education Research & Data Center, include financial aid, foster care, homelessness, apprenticeship, juvenile justice, and incarceration records.

Education agencies are beginning to work with other linked data sets, such as information on social services and workforce development. For example, Rhode Island's Longitudinal Data System used data from the state departments of health and education in reports describing the impact of lead exposure on K-12 students' education outcomes.

At the same time that the scopes of linked data sets are expanding, the intended audiences, ways in which information may be used, governance structures, and technical infrastructure are evolving. Rather than creating a single database that is located on a dedicated server and strictly controlled by data providers, states are leveraging new technical capabilities that allow data to be combined flexibly and securely. For example, the Texas Higher Education Coordinating Board recently upgraded its internal systems and released dashboards showing student progress

toward state goals. In addition, Texas' Education Research Centers provide researchers with secure access to online spaces where they can conduct approved research studies.

States are also creating more opportunities for community input on allowable data system uses. For example, in addition to supporting accountability reporting or providing access to researchers to evaluate specific programs, states are leveraging linked data to identify children who would benefit from food aid or streamlining college and financial aid application processes. California's Cradle to Career Data System partners with CaliforniaColleges.edu to provide information to individual students on college and career options and whether they have taken the right courses to be eligible for the state's public 4-year colleges. The platform also provides services like sending transcripts for college applications and supporting financial aid applications.

Due to the expansion of data system scopes, the entities that manage data systems are shifting from the agencies that oversee K-12 education to other entities, such as universities, state information technology departments, or new governmental agencies. For example, Maryland's Longitudinal Data System Center is hosted by an independent agency and Nebraska's Statewide Workforce & Educational Reporting System is a joint administrative entity created by the data providers.

The role of these managing entities is also starting to change, from gatekeepers that limit access to data to service providers that help ensure information can be safely provided to address a broad range of uses. For example, Kentucky's Center for Statistics offers services like connecting employers to career and technical education programs, providing quarterly reports on workforce trends and commuting patterns, and conducting research on the predictors of education and workforce program outcomes.

Given the many different ways linked data sets can be used, one of the first steps of a planning process should be narrowing down the functions that best align with overarching state goals.

A User-Centered Design Approach for Expanding Education Data Systems

To help reframe the concept of designing modern data systems that integrate data across multiple state agencies, AISP, ECS, DQC, and WestEd focused on the functions that the data system serves for a state. The organizations built upon a framework developed by AISP to describe three common functions that states prioritize when linking education data to other information:

- ***Public Reports & Dashboards***: provide transparent, consistent information that displays and clarifies outcomes.
- ***Research & Analytics***: make data available to authorized parties to develop nuanced analyses of the factors impacting outcomes.
- ***Supporting Individuals***: provide personalized services to individuals so they can access education and public services.

Many states have developed education data systems that allow for both Public Reports & Dashboards and Research & Analytics functions. However, their governance structures, technical infrastructure, and legal agreements may make one of the data system functions more feasible than another. Fewer states have built tools for Supporting Individuals, which require different technical and legal frameworks. Nevertheless, there are strong examples of all three functions across the country, which makes it possible to create guidance regarding how states could modify their current data system so it better meets the needs of data providers, policymakers, and communities.

Expanding on previously published work by AISP, the organizations examined 17 core elements of education data systems that link individual records across agencies and how implementation varies for each of the three functions. For example, the frequency with which data are provided for the Public Reports & Dashboards function is generally annual, in a manner that is aligned to accountability requirements. For the Research & Analytics function, annual data submissions might be supplemented with other data sources to fulfill specific research requests. For Supporting Individuals, data are

provided in near real-time. Other core elements include topics like governance, the managing entity, data types, legal agreements, technical architecture, security, privacy, policy requirements, community engagement, and sustainability.

Read more about AISP's work on the core purposes for data sharing, which informed this paper: https://aisp.upenn.edu/about-data-sharing/

A Roadmap for Design Decisions

The table below clarifies how data systems need to be designed in 17 elements to support the three primary functions. This information can be used to prioritize action steps as states modernize their education data systems.

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Demand		
Rationale		
Make outcomes of learners more transparent and fulfil planning and reporting requirements	Enable detailed analyses of long-term learner outcomes and the impact of participating in multiple public programs	Reduce the burden on individuals to access public services and understand their opportunities
Approach		
Data are linked, anonymized, and results are reported at the aggregate level (for example, results are shown for groups of 10 or more people)	Data are linked, de-identified, and shared for a specific purpose (for example, an authorized researcher can see outcomes for specific people but not their names)	Data are linked and shared to help individuals (for example parental income level about an individual student could be shared to qualify that student for public benefits)
Audiences		
• Agencies providing the data • Policymakers • Institutions (such as schools and colleges) • Media • Intermediaries that support agencies and institutions	• Researchers • Evaluators • Research and planning staff from institutions (such as schools and colleges)	• Service providers (such as counselors) • Individuals (such as learners, workers, and individuals receiving services) • Caregivers (such as parents or guardians)

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Systems		
Types of Data		
Data required for the managing entity to match records across data providers		Data required for the managing entity to verify the identity of individuals
Data necessary for accountability reporting	Data necessary to answer questions associated with a research agenda	Only the information necessary for a service delivery process (such as eligibility requirements for financial aid)
Data that document a more complete set of experiences for individuals over time		
Data that are aligned with the priorities of the state		
Data Frequency		
Data are provided to the managing entity at least annually, with frequency associated with when data providers receive information and timing for reporting requirements	Data are provided to the managing entity at least annually, with frequency associated with when data providers receive information Additional data sharing may occur to fulfill approved data access requests	Data provided to the managing entity in near real time, such as every night
Data Quality Standards and Documentation		
Data are evaluated by the managing entity for correctness, missingness, accuracy, and stability over time		

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Data definitions align with requirements for accountability reporting or widely accepted standards (such as a definition for foster youth used for standardized disaggregation across agencies)	The managing entity may combine information in ways that do not align with accountability reporting or widely accepted standards to address research questions (such as an expansive definition of an individual that was ever engaged in the foster care system)	Data definitions align with requirements for service delivery processes (such as whether an individual is considered a foster youth in a manner that qualifies them for public benefits)
Lower quality match rates across data sources may be acceptable		The quality of match rates across data sources must be at the highest level
Data are compared to other similar reports at the federal, state, and institutional levels, with public explanations about why misalignment may occur	Analysts may use new data points that have missing or lower-quality data to address research questions	Information about individuals must be accurate or offer opportunities to correct errors
Technical Architecture		
Data warehouses for source data		Mechanism for receiving and sending information (such as an API)
Master data management solution to link records		
Data warehouse for linked data set		Data warehouse
Visualization capacity	Ability to upload additional information to the data system and merge it with the data set	Visualization capacity
Public website	Secure environment where authorized users can access approved information	Public website
	If using a virtual environment, provide a variety of analytical tools and programming languages to authorized users	

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Security		
Security requirements ensure that the data system and the specific information within it are protected from unauthorized access		
Access controls limit the people who can view the information provided by source agencies that is used to match records		
The managing entity has an incident response plan should data be breached		
Data providers review information before it is released in public dashboards to ensure the tools comply with federal rules	If information is provided to an external entity, the managing entity ensures that the data have been destroyed after the study is complete	Security protocols for various levels of access have been developed based on an individual's role (for example, a counselor might be able to see information on many students, while a parent can only see information on their own child)
Governance		
Sustainability		
Legislation and relationships are in place to ensure that the data system persists, with enough flexibility for the system to evolve		
Ongoing funding from public sources and sufficient funding for the work assigned to the managing entity		
Develop use cases and case studies that clarify the value of the system		
Document what changed as a result of people having access to the data		
For a policy audience, support descriptive analyses that help to clarify the value of specific investments	For a policy audience, support complex analyses that help to clarify the value of specific investments	For a policy audience, document that proportion of people in the state benefiting from the tool

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
For philanthropic audiences, use public tools to develop descriptive analyses on priority populations	For philanthropic audiences, focus research on priority populations	For policy, philanthropic, and data provider audiences, focus service delivery tools on priority populations and clarify how the information contributed to improved outcomes
For data providers, document how the public visualizations support required and priority analyses	For data providers, document how the research request process supports required and priority analyses, while reducing the workload of individual data providers to fulfill those requests	
Managing Entity		
The organization has the following capacities, either on staff or through contracts: • receive, match, store, and display information from data providers • address alignment and consistency of information • create descriptive statistics • create data visualizations • domain knowledge in content areas • engage policy makers, data providers to identify shared agendas for information to compile and visualize	The organization has the following capacities, either on staff or through contracts: • receive, match, store, and display information from data providers • domain knowledge in content areas • support researchers and evaluators to understand available data • support researchers and evaluators to navigate the request process • construct data sets that align with research request • conduct research including predictive and advanced analyses	The organization has the following capacities, either on staff or through contracts: • receive, match, store, and display information from data providers • create data visualizations • understand service delivery processes (such as how to apply for the FAFSA) • support local entities to improve underlying data quality

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Legal Framework		
Data sharing agreements among data providers that allow for data to be compiled, stored, and displayed by the managing entity	Tiered data sharing agreements among agencies that allow data to be linked by the managing entity in a standardized but flexible way and shared with authorized external parties	Data sharing agreements provide consent for the managing entity to link and share information with specific types of people at specific entities (such as counselors at a high school)
Legal agreements do not change very often	Legal agreements focus on the purpose for data use, specific terms and conditions governing data access, and roles and responsibilities of data recipient	Data sharing agreements may be supplemented by legislation that authorizes data sharing
Data providers can make minor modifications to the data they share without re-executing legal agreements	Data sharing templates streamline the process of granting access to external parties	Data providers may ask individuals/parents to provide consent to share data for use in service delivery tools as part of standard processes (such as enrolling at an institution)
Governance Structure		
The governance structures include representatives from the data contributors, including leadership to guide priorities and appropriate subcommittees to address technical topics such as privacy, security, and data definitions.		The governance structures include representatives from the data contributors, including leadership to guide priorities and appropriate subcommittees to address technical topics such as privacy, security, and state processes.

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
The governance structure establishes joint priorities about which data tools get developed, with buy-in from the data contributors.	The governance structure manages the process for setting the research agenda and approving and prioritizing data requests.	The governance structure empowers the group to set joint priorities about which public services are enhanced with linked data, with buy-in from state and local entities.
The governance structure provides opportunities for input from the public about the information that is linked and visualized, determine ways to make it more accessible, and identify potential additional visualizations.	The governance structure provides opportunities for input from external entities that conduct research and evaluation and the public about the research agenda.	The governance structure provides opportunities for input from the public about the efficiencies that support service delivery.
Policy Requirements		
Laws and regulations provide permission to share data across agencies		
A governance structure is written into law, with enough flexibility for the system to evolve		
Laws and regulations provide the necessary authority for the managing entity to fulfill reporting requirements	Laws and regulations provide the necessary authority for the managing entity to fulfill approved research requests	Laws and regulations provide the necessary authority for the managing entity to provide individual level data to support service delivery
Privacy		
The design of the data system is based on the privacy strategy and privacy policies ensure that the identity of individuals represented by the information in a data system are protected		
Staff at the managing entity must complete training on data security and ethics		

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Individual identities are protected in public tools using a suppression protocol, which describes how data will be aggregated (such as requiring at least 10 people to be included in a data set)	Researchers and evaluators must complete training on data security and the ethics of human subject research	
The managing entity conducts a disclosure avoidance process to ensure that individual identities cannot be inferred	Data providers review requests for data access and ensure it is an allowable use	
Data providers review results before they are released to ensure that individual identities cannot be inferred	Data requests are evaluated by an Institutional Review Board to protect the interests of the people whose information is in the data set	
	Individual identities are protected in the data included in research and evaluation reports using a suppression protocol, which describes how data will be aggregated (such as requiring at least 10 people to be included in a data set)	
	The managing entity conducts a disclosure avoidance process to ensure that individual identities cannot be inferred	
	Data providers review results before they are released to ensure that individual identities cannot be inferred	

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Use		
Data Visualization		
Dashboards, query tools, and reports that allow members of the public to explore specific topics that have been prioritized through the governance process	Reports that allow members of the public to benefit from the research conducted using the data system	Dashboards, completed forms, and case management information that support service delivery
	Information on who has requested data, the status of those research requests, and information on why requests were denied	
Data Access		
Information is presented in a manner that makes it easy to understand	The managing entity makes data dictionaries available to the public, with clarification on which data points are available, from which data providers, and for what timeframe	Information is formatted in a way that is aligned with the related service delivery process
Information has technical notes and plain language explanations that support users at multiple levels of data literacy to understand the information presented	The managing entity makes methodologies for research studies publicly available to support consistency across analyses	Information is exchanged between systems so that it can be imported into the service delivery process, rather than requiring individuals to enter information (such as the courses a student took in high school being uploaded into a college application)
	Researchers and evaluators have convenient access to the information to conduct their studies (such as through a secure virtual interface)	Information is easy to understand and described in plain language

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Community Engagement		
Provide information that supports data literacy, including:	Regular communications that provide transparency to the public about what is being done with the data and how it is being protected, including sharing findings from research studies	Targeted communications to institution leadership, advisors, and service providers about the availability of information and its potential use
Explaining the nature of the information provided in each visualization	Targeted communications with interest holders that address misconceptions and differences of opinion about the allowable use of data	Trainings for people who will use service delivery tools
Explaining why figures may differ from similar public reports		
Clarifying the factors that shape data timeliness		
Suggestions for ways to use the information		
Regular communications that provide transparency to the public about what is being done with the data and how it is being protected		
Targeted communications with interest holders that address misconceptions and differences of opinion about the allowable use of data		

Public Reports & Dashboards	Research & Analytics	Supporting Individuals
Items to Track Continuous Improvement		
Whether dashboards and reports are being generated in a timely manner for accountability purposes	The timeliness of fulfilling data requests	Impact of the service delivery tools on the individuals served (such as increased rates of application, college-going, or benefit uptake)
The degree to which data are being accessed and by whom	The degree to which data are being accessed and by whom	The quality of data used in service delivery tools and ways to improve the source information
The breadth and depth of training and interest holder engagement	The security status of the data warehouses that hold sensitive information	The security status of the data warehouses that hold sensitive information
The security status of the data warehouses that hold sensitive information	The security status of the secure environment used by external parties to access data	

More About Our Team

Actionable Intelligence for Social Policy

Actionable Intelligence for Social Policy (AISP) supports the ethical use of individual-level administrative data for social policy change and advocates for the expansion of resources and infrastructure that makes this possible. We help foster cross-sector collaborations, build the relationships and trust that enable and sustain data sharing, and center racial equality. We are proud to convene a national network of 44 state and local integrated data systems, including a range of models that span education, health and human services, housing, and more.

Data Quality Campaign

The Data Quality Campaign (DQC) works to support states as they change the role of to ensure that data works for everyone navigating their education and workforce journeys. Today's most important education and workforce problems require data and people to work together across early childhood, K–12, postsecondary, and the workforce. That is why DQC's strategy is about supporting people to change policies, systems, and culture. Making data work for people is not about technology; it is about managing change, getting buy-in, and helping people connect dots. To support state policy and implementation, DQC offers our expertise and support to local, state, and federal leaders on policies and practices that help states implement data ecosystems that provide value for the people they serve; recommends state and federal actions to promote, support, and incentivize improvements to state data systems; and acts as a critical

friend to leaders, working across the field to support people on messaging, building trust, and coalition building.

Education Commission for the States

Education Commission of the States (ECS) collaborates with education policy leaders nationwide to address issues by sharing resources and expertise. The mission of ECS is to advocate for attaining educational excellence for all and to help state leaders identify, develop and implement public policy for education that addresses current and future needs of a learning society. ECS is proud to serve both the people who develop and implement education policy and the students who directly benefit from effective policy change. Every day, ECS provides education leaders with unbiased information and opportunities for collaboration because informed policymakers create better education policy.

WestEd

The Data Integration Support Center (DISC) at WestEd provides expert planning and user-centered design, policy, privacy, and legal assistance for public agencies nationwide. DISC supports efforts to integrate data across key sectors while protecting privacy and ensuring the usefulness, fidelity, and transparency of the data systems. DISC provides comprehensive support and expertise to public agencies to modernize integrated data systems through customized assistance in focus areas that are critical to the development, sustainability, and maturity of integrated data systems. Many of these services are provided at no cost to public agencies.