DQC Data Quality Campaign

disc Data Integration Support Center WestEd

JAN 2026

STATE IMPLEMENTATION GUIDE

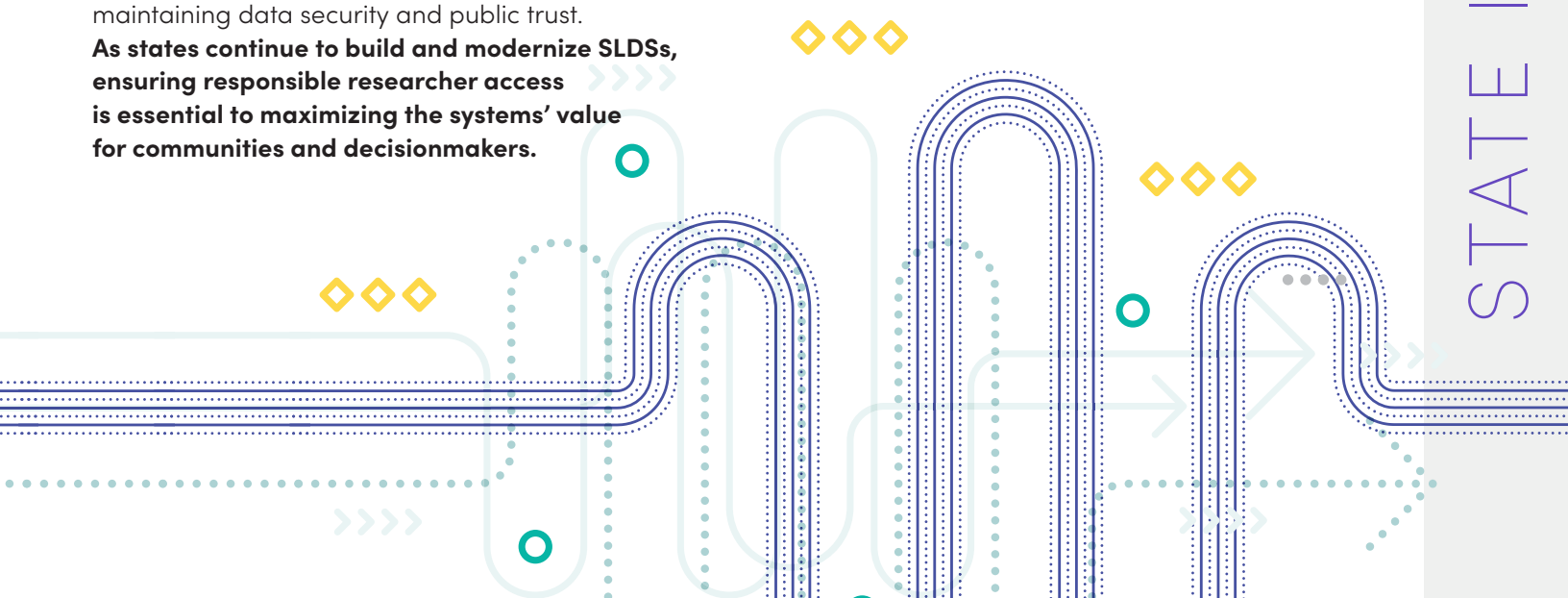
# Creating a Path to Answers: How State Data Leaders Can Empower Researchers

Statewide longitudinal data systems (SLDSs) can hold the key to understanding what works in education, workforce development, and beyond—but they can enable this insight only if researchers are able to access and use the data effectively. By linking individual-level data over time and across at least early childhood, K–12, higher education, and the workforce, SLDSs can provide a unique opportunity for researchers to use cross-sector information to produce trusted insights that help individuals, the public, and policymakers understand education and workforce transitions and outcomes.

As states work to optimize researcher access, they face common challenges, including managing limited staff capacity, navigating complex legal requirements, establishing appropriate security protocols, and balancing multiple stakeholder priorities. These challenges are solvable, and states across the country have developed innovative approaches that address these issues while maintaining data security and public trust. **As states continue to build and modernize SLDSs, ensuring responsible researcher access is essential to maximizing the systems' value for communities and decisionmakers.**

**Researchers seeking access to SLDS data** may be individuals or groups conducting a study that are affiliated with a data-contributing agency, a non-data-contributing agency, a postsecondary institution, or an external organization.

This brief provides state data leaders—the individuals spearheading states' longitudinal, cross-agency data work—with practical pathways to secure, transparent, and effective researcher access. Through research and input from state data leaders, the Data Quality Campaign (DQC) and the Data Integration Support Center (DISC) at WestEd have identified five best practices for enhancing researcher access to de-identified individual-level SLDS data.



## Why Individual-Level Data Matters for Research

Individual-level data—the detailed record of a person's journey through education and into the workforce—that is de-identified enables robust statistical analysis to better understand what is working (and not working) for learners and earners, which aggregated statistics simply cannot reveal.

Individual-level, or record-level, data allows researchers to:

- **Evaluate how specific interventions affect different student populations** over time;
- **Identify which pathways lead to successful workforce outcomes** for various demographic groups;
- **Understand why some students thrive** while others with similar backgrounds struggle;
- **Evaluate program effectiveness** by following actual participants rather than comparing group averages; and
- **Discover unexpected connections** between early education experiences and later career trajectories.

Without secure access to individual-level data, researchers can report that "60 percent of students succeeded," but they cannot explain which students succeeded, why they succeeded, or how to help the other 40 percent. This detailed understanding is critical for developing targeted policies and interventions that work for all students, not just the average.

# Best Practices for Advancing Researcher Access

Through thoughtful implementation of the following five best practices, SLDSs can become an even more powerful engine for continuous learning about what works in education and workforce development. These recommendations can help state leaders maximize the return on their state's data investment while centering security and privacy protections that preserve public trust.

1. **Communicate Consistently and Transparently.** Build trust through sharing clear, accessible, and up-to-date information with both the research community and the public at large.
2. **Engage Researchers in Decisionmaking.** Create effective and sustainable programs through inclusive governance and community engagement.
3. **Structure Research Programs So Value Flows in Both Directions.** Ensure that both researchers and states benefit from researcher access to data.
4. **Ensure That Data Use and Privacy Go Hand in Hand.** Keep data privacy and security top of mind throughout the process to maintain public trust while enabling appropriate use.
5. **Leverage Technology to Enhance Capacity and Efficiency.** Safely scale researcher access without overwhelming staff capacity.

Each of the best practices is an important component of advancing research access. As state leaders think about how to implement or improve researcher access in their own contexts, starting with the practice that is most relevant for them is important—but states should also think critically about how to eventually implement all of the best practices.

## Understanding Where Researcher Access Already Exists


In 2024, Strada Education Foundation identified 10 critical elements that contribute to the capacity of state data systems to strengthen the connection between education and opportunity. Among the elements is giving researchers access to individual-level but de-identified matched education-to-opportunity datasets. State leaders can look to Strada's State Opportunity Index to find other states that may be further along in offering robust researcher access.

## COMMUNICATE CONSISTENTLY AND TRANSPARENTLY

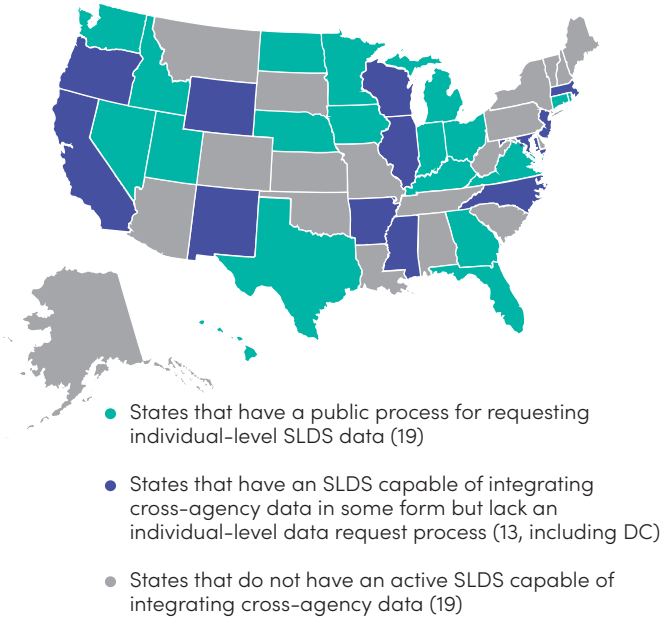
Clear, easy-to-access information about data availability and access procedures forms the foundation of any successful researcher access program. When researchers understand **what** data exists, **how** to request it, what it **costs**, and **how long** the process takes, both researchers and SLDS staff save valuable time and resources.

### Actions for State Data Leaders

**1. Create and publish comprehensive researcher guidance materials.** More than half of states with an active SLDS have published documentation for their data request processes. Building a centralized, easily navigable online resource accomplishes multiple goals: it democratizes access to information, reduces repetitive inquiries to staff, and demonstrates a state's commitment to transparency. The most successful guidance materials include the following elements:

- **Data inventory and dictionary:** A catalog of all available data elements with descriptions, years available, data quality notes, and any access restrictions. These resources should include alignment to standards such as the Common Education Data Standards, where applicable.
- **Process transparency:** Visual and written guides showing each step from initial inquiry through data delivery, with realistic timelines for each phase; a summary of active and completed research projects with status updates and anticipated completion dates; a research library with links to publications, policy briefs, and presentations resulting from SLDS research; and regular reporting on request volumes, processing times, and system improvements.
- **Legal and administrative templates and rules:** Templates for memoranda of understanding, data sharing agreements, institutional review board (IRB) requirements, suppression rules, and other standard forms with clear instructions for completion.
- **Training resources:** Self-service materials explaining data structure, appropriate use cases, limitations, and technical requirements for access.

*19 of the 32 states with an active SLDS have a public process for requesting individual-level SLDS data*




- **Costs:** A transparent pricing structure with any variations by requestor type, data complexity, or processing requirements.
- **Decision criteria:** Information on eligibility requirements, the process for evaluating and sequencing requests, details on the approval process (e.g., decisionmaking structure, review committees involved, and typical approval rates by request type), and clear examples of in-scope vs. out-of-scope requests with guidance on alternatives for declined requests.

**As artificial intelligence (AI) tools become increasingly integrated into research workflows, states must establish clear guidelines** about their appropriate use with SLDS data. AI platforms (e.g., ChatGPT, Claude) may retain, train on, or expose data in ways that violate privacy agreements and legal requirements. Data use agreements and training materials should explicitly address if and how AI tools can be used with specific data to ensure that data privacy and security can go hand in hand with using data for research amid evolving technologies. These agreements should restrict researchers from uploading actual SLDS data to external AI platforms.

**2. Develop a researcher communications plan.** Effective communication goes beyond posting information—it requires thoughtful strategies for reaching different audiences through appropriate channels. A state's communications approach should balance comprehensive information sharing with internal staff capacity. To develop a successful researcher communications plan, state leaders should:

- **Determine a main point of contact for inquiries**, whether the contact is a dedicated researcher liaison, rotating staff member, or leadership team member.
- **Choose communications efforts that align with their capacity.** Keeping a frequently updated public website, a research portal, or a series of static downloadable guides current takes different levels of effort.
- **Develop materials with different levels of detail for different users** (e.g., quick reference sheets for experienced users, comprehensive guides for newcomers, technical documentation for data scientists) and create processes for keeping information current.
- **Ensure that the right audiences can access this information** by leveraging existing relationships—including partnerships with university research offices and state agencies as well as research–practice partnerships—to amplify communications and create opportunities for researchers to share feedback and/or ask questions.

## Best Practices in Action

**Washington's Education Research and Data Center (ERDC)** has a robust and transparent data request process that includes a guide on how to request data, expectations for how the data request process will work, a data dictionary and handbook, and available supports. A standout feature is ERDC's Public Data Request Log, which allows researchers to see recent requests, who submitted them, what data was requested, and the current status of each request. This visibility helps researchers track their own requests, avoid duplicating existing projects, and better plan their timelines. The result is a more efficient, researcher-friendly system that builds trust and streamlines access to valuable education and workforce data.

**Minnesota's Statewide Longitudinal Education Data System (MN SLEDS)** has a transparent data request process that outlines the timeline, duration, and sequence of each phase, from submission to review to project completion. The system informs researchers about when and how they will receive updates on the status of their requests, and MN SLEDS staff evaluate proposals using a scoring rubric aligned with clearly communicated expectations. This structure not only clarifies how proposals are assessed but also increases fairness and builds trust in the process. Additionally, MN SLEDS specifies levels of data access for different users, including external researchers. For instance, only MN SLEDS staff and designated staff from contributing agencies can access linked personally identifiable information (PII), while approved researchers are granted access to anonymized, linked data that is not PII. This transparent tiering results in a system that supports research while safeguarding privacy.

The **Utah Data Research Center (UDRC)** publishes a clear and easily accessible data dictionary that includes descriptions, field values, and years collected for variables, along with potential limitations and notes about the variables that may be of use to a researcher. Researchers can use this information to determine if the UDRC can answer their research questions, helping ensure that research requests align with data availability and utility.

The **Hawai'i Data eXchange Partnership (DXP)** offers a Data User Training Course that all requestors of de-identified, individual-level data must complete. The course details the data request process, describes how to responsibly handle DXP data, and includes quizzes that users must pass to obtain a certificate of completion. Data requestors must include this certificate in their research request. This process provides researchers transparency around the research process and confirms that they are aware of data use, privacy, and security expectations.

## ENGAGE RESEARCHERS IN DECISIONMAKING

States have limited time and capacity. But investing some of that time and capacity in robust engagement with researchers to inform state data efforts ensures that states are building systems and tools that researchers can trust and use right from the start. Going back to the drawing board and rebuilding data systems, processes, and tools is much more expensive, frustrating, and time consuming than taking the time and energy on the front end to build state data systems on a strong foundation of engagement. Investing in that foundation can help state data efforts go further in the long run.

**Engaging communities** is a critical step toward ensuring that data efforts work for their intended audiences, including researchers. For more on engaging communities, see DQC's resource, *Community Engagement Is Crucial for Successful State Data Efforts*.

### Actions for State Data Leaders

**1. Establish relationships with postsecondary institutions.** Partnering with intermediaries, such as universities or colleges, can help state leaders reach audiences more effectively. Universities and research organizations offer more than just potential data users—they can provide infrastructure, expertise, and networks that can amplify an SLDS's impact. Strategic partnerships help distribute the workload of researcher engagement while building sustainable capacity for high-quality research. Elements of these partnerships may include:

- Working with universities to **cocreate research agendas** and identify research priorities that address both academic interests and state policy needs.
- **Leveraging university infrastructure** and resources such as secure data enclaves, IRB processes, or technical support to reduce SLDS operational burden.
- **Collaborating on federal grants, foundation funding, or other resources** that benefit both the SLDS and research institutions.
- **Building future research capacity** through student pipelines, including graduate assistantships, dissertation support, or practicum opportunities.
- **Sharing research findings** through university networks including academic conferences, publications, and policy forums.

**2. Bring researchers into the decisionmaking process.** Researchers who help shape the SLDS become invested partners rather than external critics. Creating formal roles for researchers in the process ensures that user perspectives inform critical decisions about data access, security requirements, and system improvements. Incorporating feedback from and making decisions with researchers hones the strategic direction of state data efforts. This engagement not only will build trust, but it also will improve data efforts by ensuring that the systems, tools, and resources states develop are valuable for this group. Engagement may include:

- Establishing a standing **research advisory committee** of 8–12 researchers representing different institutions, methodologies, and research domains to advise on strategic decisions.
- **Reserving a seat on the state's governance board** specifically for active researchers who can represent user perspectives.
- **Creating technical working groups** for specific issues (e.g., data quality, privacy methods, emerging technologies) that include researcher expertise.
- Bringing researchers into the SLDS for 6- to 12-month **fellowship assignments to work on priority projects** while providing different perspectives on data access and system improvements.

### Best Practices in Action

**The California Cradle-to-Career Data System (C2C)** and the **Nebraska Statewide Workforce & Educational Reporting System (NSWERS)** allocate seats on governing and advisory bodies to postsecondary institutions and researchers. The C2C's Data and Tools Advisory Board includes several researchers and university staff. Similarly, the Executive Council of NSWERS (charged with governing the system) includes representatives from the Nebraska State College System, the University of Nebraska, and Nebraska's community colleges. This type of inclusion provides researchers with space to give feedback on the research request process and be involved in shaping the vision and direction of the system.

## STRUCTURE RESEARCH PROGRAMS SO VALUE FLOWS IN BOTH DIRECTIONS

Successful researcher access programs create value that flows in both directions. Researchers gain access to unique longitudinal data that enables high-quality research while states receive evidence-based insights that improve programs, validate investments, and guide policy decisions. The key is designing research processes that systematically capture this value while maintaining sustainable operations.

### Actions for State Data Leaders

**1. Define clear goals and success metrics for the research program, including the research agenda.** Without explicit goals, researcher access can become a compliance exercise rather than a strategic function. Defining what success looks like—and how it gets measured—enables better resource allocation, clearer communication with key audiences, and stronger justification for continued investment. Determining desired goals and outcomes is a crucial step and can help determine how much flexibility a state has when designing processes, allocating resources, and managing competing demands. Understanding the "why" behind researcher access will help state leaders both communicate about the process *and* make the case for ongoing support and investment. States can accomplish this work in the following ways:

- **Establish a research agenda that outlines top research priorities for the SLDS**, ideally multiagency priorities aligned to the state's education and workforce goals and determined by a cross-agency governing body.
- **Develop a process for evaluating whether and how researcher access is meeting the state's goals.** For example, track research requests and outputs to measure the impact on policy and practice.

**2. Incentivize and prioritize research that aligns to the research agenda.** When demand exceeds capacity, states need to determine how they will evaluate requests and in what order they will process those requests. Prioritizing research that aligns with a state's agenda can help ensure consistency. Some states may even restrict research requests to those that align with the state's research agenda. In either case, state leaders can publish these priorities clearly to ensure that researchers submit appropriate requests. While states may not have the capacity or authority to address every request they receive, building partnerships to expand research capacity can help address policy or practice questions that may not be aligned to top SLDS research priorities.

**Thirteen states** require every data request to align with state or SLDS priorities.

### Should a State's SLDS Implement Cost Recovery? It Depends.

Although some states have implemented cost-recovery models to offset costs associated with researcher access, especially for large or complex projects, this approach is not a one-size-fits-all solution. Whether cost recovery is right for your state depends on several key factors:

- **Is it legally allowed?** A state's enabling legislation may prohibit or explicitly allow the collection of external funds. Understanding the legal framework is a necessary first step.
- **Can the agency collect funds?** Not all SLDS-managing entities have the legal or administrative capacity to collect money from external sources. In some cases, states must create new processes or route funds through partner agencies. For example, in Washington, researchers pay the vendor that hosts the ERDC's secure enclave directly, offsetting the ERDC's costs and bypassing the need for the state to handle payments.
- **Is it worth the return?** Revenue from cost recovery varies widely by state and scale. In recent years, Maryland has focused on cost recovery, which generates between $10,000 and $30,000 annually. Kentucky, with a longer-term focus on cost recovery, sees $150,000–$200,000 in revenue annually.

If a state chooses to pursue cost recovery, transparency is essential. Publishing clear cost structures and encouraging researchers to include SLDS access fees in grant proposals can help streamline this process and build trust.

**3. Set clear requirements for deliverables that benefit all parties.** State leaders should set clear parameters for ensuring that research generates usable insights while also meeting researcher goals. Well-designed requirements that include the following elements create value without stifling innovation:

- **Specific outputs during and at the conclusion of the study.** Researchers could be asked to provide final presentations, reports, executive summaries, or other outputs that describe the insights gained from the research in plain language for a broad audience, which agency staff will post on the agency's public-facing website. They could also be required to provide open source code and technical documentation.
- **Ongoing updates, documents, and communications with the relevant SLDS governing committee throughout the research process.** The timing and format of the deliverables should be defined in the data sharing agreement, but updates on progress throughout the process ensure that states can speak to the value of research projects as they are taking place.
- **Citation of the SLDS as the data source in any materials.** All materials should provide attribution to the SLDS.

## Best Practices in Action

The **Maryland Longitudinal Data System (MLDS) Center** structures its research agreements such that all research requests result in a product owned by the MLDS Center. Examples of products include presentations, reports, research summaries, policy briefs, dashboards, and data tables. This process ensures that all research projects provide value to the state, the MLDS receives proper attribution and maintains a research portal showcasing the value of the data system, and project findings and outcomes are made available to the MLDS's audience.

The **Texas Education Research Center** requires all external researchers to produce policy briefs that link research findings to Texas policy or practice within 60 days of project completion. Recent topics have ranged from teacher retention to computer science education to the impact of homelessness on K–12 and postsecondary outcomes. This requirement ensures that all external research requests are policy relevant and produce timely, actionable insights for the state.

## ENSURE THAT DATA USE AND PRIVACY GO HAND IN HAND

Centering data privacy and security ensures trust among all those who have a stake in the education, workforce, or other data that is linked within an SLDS. Privacy and security are not just compliance checkboxes, they are foundational elements that enable sustainable researcher access. Building privacy and security into every stage of the research process creates a culture of responsible data use that protects individuals while enabling valuable research that informs better policy and practice.

### Actions for State Data Leaders

**1. Adopt a framework for security and privacy that meets the needs of the SLDS, its users, and the legal landscape.** State frameworks must balance three competing demands: protecting individual privacy, enabling meaningful research, and complying with legal requirements. The most effective frameworks are clear enough for researchers to understand but robust enough to withstand scrutiny. State leaders should pursue the following steps:

- **Adopt comprehensive data use agreements and/or data use licenses** to set expectations, establish enforcement mechanisms, and demonstrate commitment to data privacy. Avoiding overly restrictive terms will ensure that agreements are not limiting research without adding meaningful protection.
- **Define access roles and responsibilities.** Any framework should clearly outline who can access the SLDS, under what conditions, and with what confidentiality or nondisclosure conditions.
- **Ensure legal compliance.** States should align procedures with the established legal framework for the SLDS.
- **Coordinate IRB requirements.** States should determine when IRB approval is necessary for research involving SLDS data and document these requirements clearly.

**2. Build privacy expertise and a culture of privacy across the organization.** Privacy protection is part of everyone's responsibility, and creating a culture of privacy requires investment in people and processes. States should establish mandatory training for researchers that includes appropriate handling of sensitive information, policies, and procedures for the SLDS as well as incident response protocols. A designated privacy officer can provide immediate guidance and identify potential risks early.

**3. Enable "bring analysis to the data" approaches.** The most secure data never leaves a state's control. Modern approaches, such as secure enclaves, allow researchers to conduct sophisticated analyses without downloading sensitive datasets to their local machines, dramatically reducing privacy risk while enabling complex research.

**4. Layer protections.** States should use a mix of technical, contractual, and procedural controls to provide appropriate protection as risks and capabilities evolve. They can minimize risks by using technical controls (e.g., encryption, access logs, automated monitoring) where possible and complement these controls with contractual terms that are reinforced with policies and procedures.

## Understanding Secure Enclaves

Secure enclaves are defined by the Massive Data Institute as "virtual computing workspace[s] that enable authorized users to access sensitive data and securely conduct analysis." These virtual computing workspaces allow authorized researchers to conduct complex analyses without ever downloading or transferring sensitive information outside of the SLDS.

Key characteristics of secure enclaves include:

- **Controlled access:** Researchers must authenticate through multiple security layers and can access only data specifically approved for their projects.
- **Self-contained analytical environment:** All necessary tools—statistical software (R, SAS, Stata), programming languages (Python), and specialized applications—come preinstalled within the enclave, eliminating the need to import external software that could introduce vulnerabilities.
- **Scalable capacity:** Cloud-based infrastructure can expand to accommodate multiple concurrent users and computationally intensive analyses without compromising performance.
- **Data containment:** Raw data never leaves the enclave. Researchers cannot download, email, or transfer datasets to external systems, preventing data breaches through lost devices or compromised networks.
- **Output review:** All analysis results undergo disclosure review before release, ensuring that no individual-level information can be extracted or re-identified.
- **Complete audit trail:** Every action is logged, from login to data queries to output requests, creating comprehensive records for compliance and monitoring.

## Best Practices in Action

**Maryland**, **Virginia**, and **Washington, DC**, are working to leverage the Coleridge Initiative's secure data enclave, the Administrative Data Research Facility (ADRF), to securely link individual-level, de-identified records that span education and workforce data across the multistate region for the first time. Funded through the Democratizing Our Data Challenge, this emerging effort will enable these states to better understand the outcomes of residents who cross state lines following high school and help state leaders make more informed decisions about programs and policies. By using the ADRF, the two states and Washington, DC, will be able to collaborate and share data confidently knowing the environment is secure.

## LEVERAGE TECHNOLOGY TO ENHANCE CAPACITY AND EFFICIENCY

Resource constraints are a reality for most SLDS teams; technology can multiply impact without multiplying staff. Strategic automation and self-service options enable small teams to handle growing research demand while maintaining quality and security. The key is choosing solutions that match a state's technical capacity and provide genuine efficiency gains.

### Actions for State Data Leaders

**1. Leverage expertise across agencies and government.** States already have technology infrastructure and expertise; state leaders should tap into their existing resources rather than build from scratch. Centralized resources often provide better security, reliability, and cost-effectiveness than standalone solutions.

**2. Automate high-volume, repetitive processes and implement workflow management.** State leaders should consider where technology can take some of the burden off staff, especially in the early stages of research requests, including processing submission forms, performing initial screenings, and tracking requests. Combining process automation with workflow management creates efficiency throughout the research request lifecycle and will ensure that requests do not get lost, timelines stay on track, and researchers maintain confidence.

**3. Develop self-service alternatives to custom data requests.** Self-service options satisfy common needs immediately while preserving staff capacity for complex, high-value requests. Data downloads in files that can be manipulated allow exploration of aggregate, de-identified data. States can also create prebuilt, de-identified, and documented datasets addressing common research topics to allow researchers to access information on demand, further freeing staff time.

### Best Practices in Action

The **Indiana Management Performance Hub's (MPH)** research portal prompts researchers to consider MPH's inventory of publicly available data prior to submitting a formal data request. One such source, the Indiana Data Hub, provides access to more than 180 secure, de-identified, aggregate datasets. Each dataset comes with a data dictionary, source information, and an update history. These datasets create opportunities for researchers to explore P–20W outcomes and answer their research questions without the assistance of agency staff.

The **Kentucky Center for Statistics (KYSTATS)** publishes dashboards, such as the High School Feedback Report, that include links to download the underlying aggregate data and access detailed technical documentation. This documentation includes data sources, elements, and formulas, equipping researchers with the information they need to confidently conduct their own analysis and interpret the results. By making aggregate data easily accessible to researchers, KYSTATS eliminates the need to fill any requests that can be answered with the publicly available data.

# Conclusion

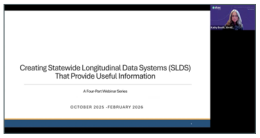
As decisionmakers and the public want to make informed decisions about education and workforce journeys, trusted and actionable research is more important than ever. Researchers can play a powerful role in a state data ecosystem by producing trusted analyses and insights that help a range of individuals, including policymakers, school leaders, and students, better understand education and workforce outcomes. Robust, clear access to the individual but de-identified data within an SLDS is key to ensuring that researchers can play that role in every state. State leaders have the opportunity and responsibility to ensure that researchers can use data to strengthen the connection between education and opportunity.

# Acknowledgments

DQC and DISC are grateful to the state data leaders in the inaugural cohort of the State Data Leaders Network, who generously contribute their time and expertise to discuss the current state of SLDSs and potential opportunities for improvement. Their insights informed this resource.

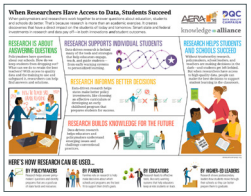
# Additional Resources

### DISC



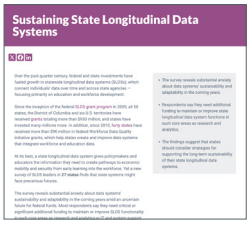

- *Maximizing the Power of Research and Analytics*: In this webinar, state leaders explore how modern SLDSs can securely expand access to linked data and enable more meaningful analyses that help leaders understand long-term outcomes and the impact of critical public investments.

### DQC


- *Use Case: Provide Researchers Access to Data*: This resource outlines real-world use cases that demonstrate why researchers need access to data to develop crucial insights for the field.
- *When Researchers Have Access to Data, Students Succeed*: This infographic, created in collaboration with the American Educational Research Association and Knowledge Alliance, shows how access to quality data and the training to use and safeguard it empowers researchers to help find answers and solutions to questions about education.

### EDUCATION COMMISSION OF THE STATES (ECS)




- *Sustaining State Longitudinal Data Systems*: ECS surveyed SLDS leaders in 27 states to determine the current state of their systems, as well as anticipated challenges. All 27 states identified research and analytics as a very important or important function of their SLDS, emphasizing the importance of a robust researcher access program.