Building + Sustaining State Data Integration Efforts: Legislation, Funding, and Strategies

AISP

Acknowledgments

"Building and Sustaining State Data Integration Efforts: Legislation, Funding, and Strategies" was created by Actionable Intelligence for Social Policy (AISP) at the University of Pennsylvania. The report was developed based on findings from our 2020 national survey and landscape analysis of data integration efforts across the US, interviews with AISP Network sites and domain experts, as well as from more than a decade of working with and learning alongside our Network and Learning Community Sites. AISP is supported by grant funding from the Bill & Melinda Gates Foundation and the Annie E. Casey Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the funders.

Many thanks to everyone who made it possible.

AISP Contributors

Sharon Zanti

Della Jenkins

Emily Berkowitz

Amy Hawn Nelson, PhD

TC Burnett

Dennis Culhane, PhD

We are grateful to our Network members who were interviewed as part of the report: Katie Breslin, Jessica Cunningham, Annice Correia Gabel, Scott Gaul, Frank Kohstall, Josh Martin, David Patterson, Kim Paull, Heather Rouse, and Lisa Tse.

We also recognize individuals who provided review of the report: Tyler Kleykamp, Jed Herrmann, Josh Inaba, Chris Kingsely, Becky Planchard, Kristen Smith, and Hayley Young.

Suggested citation: Zanti, S., Jenkins, D., Berkowitz, E., Hawn Nelson, A., Burnett, T., & Culhane, D. (2021). Building and Sustaining State Data Integration Efforts: Legislation, Funding, and Strategies. Actionable Intelligence for Social Policy. University of Pennsylvania.

Executive Summary

The economic and social impacts of the COVID-19 pandemic have heightened demand for cross-agency data capacity, as policymakers are forced to reconcile the need for expanded services with extreme fiscal constraints. In this context, integrated data systems (IDS) – also commonly referred to as data hubs, data collaboratives, or state longitudinal data systems – are a valuable resource for data-informed decision making across agencies. IDS utilize standard governance processes and legal agreements to grant authority for routine, responsible use of linked data, and institutionalize roles across partners with shared priorities.

Despite these benefits, creating and sustaining IDS remains a challenge for many states. Legislation and executive action can be powerful mechanisms to overcome this challenge and promote the use of cross-agency data for public good.

Legislative and/or executive actions on data sharing can:

- Require data sharing to address a specific state policy priority
- Mandate oversight and planning activities to promote a state data sharing strategy
- Grant authority to a particular office or agency to lead cross-agency data sharing

This brief is organized in three parts. First, we offer examples of these three approaches from states that have used legislation and/or executive orders to enable data integration, as well as key considerations related to each. Second, we discuss state and federal funding opportunities that can help in implementing legislative or executive actions on data sharing and enhancing long-term sustainability of data sharing efforts. Third, we offer five foundational strategies to ensure that legislative or executive action is both ethical and effective.

In order to set your data integration effort up for success:

- Seek to understand the full data sharing landscape
- Let purpose drive design
- Gather and incorporate meaningful contributions from community stakeholders and impacted populations
- Build standard processes and protocols for data access and use, but allow for flexibility as priorities shift
- Strategize for sustainability

Legislative + Executive Actions on Data Sharing

Legislation

Laws that promote or even require cross-agency data sharing are increasingly common as states strive to better leverage the vast amounts of information they collect to improve outcomes for residents. Because every government operates in a unique social, political, and fiscal context, stakeholders will need to assess whether the time is right to advance data sharing legislation and, if so, which model or models of legislation will work best given that context.

After 10 years working with over 20 state data sharing efforts, we have learned that highly prescriptive legislation on data sharing is unlikely to yield success, and may constrain a data integration effort's ability to meet changing needs and priorities over time. We recommend using legislative action to build upon existing relationships and support flexible, routine processes for agencies to jointly plan, implement, and modify data efforts as new needs emerge. The following examples demonstrate three discrete ways that states have used legislation, each with its own benefits and challenges. In practice, states may employ multiple models that are mutually reinforcing, particularly as an effort develops or a jurisdiction's context changes over time.

Legislation that requires data sharing to address a specific state policy priority

Legislation on a specific social policy or population may direct that data be shared as a means towards achieving cross-agency aims and outcomes. For example, in 2015, the Massachusetts legislature passed a law known as Chapter 55, with the aim of developing a coordinated and evidence-based response to the quickly escalating opioid epidemic. This new law expressly permitted administrative data to be linked across 10 agencies to provide a more comprehensive analysis of the crisis. The results proved so insightful that the state legislature reauthorized Chapter 55 in 2016, doubling the number of data-contributing agencies and establishing a process by which external researchers could access data. Since then, the effort has informed major state-level policy and program changes, and Massachusetts has reported declining rates of opioid-related deaths.

> ***Key consideration:*** Embedding directives on data sharing in issue-driven legislation can drive rapid advancements in practice. However, data sharing capacity and collaboration may not be sustained as legislative attention shifts unless agencies are able to institutionalize trust, pivot to respond to new priorities, and attract continued investment in data infrastructure. See our Strategies section below for actions you can take to address these challenges.

Legislation that mandates oversight and planning to promote a state data sharing strategy and best practices

A second type of legislation is more explicitly focused on advancing a statewide data sharing strategy to drive evidence-based decision making. For example, in 2018, Connecticut legislators created the position of state Chief Data Officer in the Office of Policy and Management. This new position was tasked with facilitating the sharing and use of cross-agency data, and then reporting annually to the legislature and the public on progress. By starting with an inventory of existing data assets, data sharing agreements, and related legal issues across the whole landscape, Connecticut laid a strong foundation for future alignment.

> ***Key consideration:*** This approach, though top-down, should not be overly prescriptive. In Connecticut, the legislation did not mandate that any specific data be shared right away or dictate which agency would be responsible for data linkage. Rather, the legislation designated planning and oversight to a neutral office for an enterprise view of gaps and opportunities in order to benefit all involved.

Legislation that grants authority to a particular office or agency to lead cross-agency data sharing

Data integration efforts may also be formalized and functionally sustained via legislation that directly grants authority to an existing agency (e.g., Department of Human Services) or creates a new, standalone office tasked with linkage, analytics, and reporting. Such statutes can ease data access barriers by giving the designated agency authority to act as an agent or actor of other data sharing agencies. For example, in Indiana, the Management Performance Hub (MPH) within the Office of Management and Budget (OMB) is tasked by law to collect, analyze, and exchange government information in carrying out the powers and duties of the OMB and the executive state agency sharing the data. Likewise, the South Carolina Health and Human Services Data Warehouse, where data have been integrated since 1990, was codified through legislation "to ensure that the operation of health and human services agencies may be enhanced by coordination and integration of client information."

> ***Key consideration:*** Legislation that grants direct authority may be used to help streamline collaborative data sharing processes by clearly defining roles and dedicating staff to manage essential tasks related to cross-agency data governance, data linkage, and analytics. It should not, however, be used to centralize power or authority in a single office for unilateral control. All contributing data agencies, as owners and stewards of the data they collect, must maintain control over determining appropriate uses of the data. What's more, the decision about whether to create a new office or grant authority to an existing office or agency (and if so, which one) should be carefully considered. Notably, in South Carolina, data owner approval is required to both integrate data and also release project-specific information.

Legislatively designated authority for integration is even more effective when it is used in alignment with executive leadership and action. Indiana's MPH was first established through executive order in 2014, which granted MPH the authority to act as an agent of each agency they support in order to ensure data remains under the control of an approved entity. Then, in 2017, after exploring the benefits of data integration and developing trust between partner agencies, MPH's authority was codified into law (Chapter 26). MPH also created a standard form for data sharing agreements to streamline the process and reduce both legal and administrative overhead.

Executive Action

As Indiana's experience demonstrates, the role of executive leaders in advancing data sharing and integration is also crucial. Just as with legislation, executive action on data sharing can be specific to a given policy priority of the administration or more generally aimed at building capacity across agencies. And also like legislation, executive orders (EOs) may not be able to ensure that cross-agency data sharing occurs but can provide momentum in the right direction. EOs may also be put in place to formalize and standardize efforts already in motion to share and integrate data.

EOs often complement and are mutually reinforcing to legislation, as is the case in the state of Indiana. Supplementary legislation or funding provisions may be needed in order to sustain data sharing based on executive action in the long term. Multiple EOs may also be used in some jurisdictions. In Ohio, a series of orders were used to establish a statewide platform and expand on existing, institutionalized data sharing efforts at The Ohio State University. The InnovateOhio Platform (IOP) fundamentally shifted expectations and practice around cross-agency data sharing. The "Modernizing Information Technologies in State Agencies" EO of 2019 created IOP within the Department of Administrative Services with the express goal of "facilitating secure data sharing and analytics across State agencies and programs to better serve the public." For the first time, the EO required that Ohio agencies share data with IOP unless an agency head can identify in writing a specific legal prohibition to sharing. The EO also placed oversight responsibility with the lieutenant governor, signaling that data sharing was a high priority and codifying an executive commitment to the effort.

> ***Key consideration:*** **Gaining agency buy-in and developing trust are paramount to balancing the top-down nature of this policy mechanism and ensuring successful implementation of the order.**

The Ohio EO was successful in part because it built on existing integration efforts in the state. Since 2009, a collaboration between several agencies and The Ohio State University on the Ohio Longitudinal Data Archive (OLDA) has helped provide researchers with centralized access to linked administrative data on education and workforce outcomes. IOP was developed in collaboration with OLDA to scale up existing infrastructure, enable secure data exchange among even more public agencies, and facilitate a shift from university- to agency-based governance.

Executive leaders in Ohio also recognized that agency staff would be more willing to share data and more likely to utilize new data tools if data use aligned with agency priorities. A few months prior to the "Modernizing IT" EO, Governor DeWine passed two issue-specific EOs that required exploration into how data could support state response to high-priority social problems. In January 2019, EO 2019-01D created the RecoveryOhio Advisory Council and tasked a multi-sectoral group of experts, advocates, service providers, and other stakeholders to develop recommendations related to the state's mental health and opioid use crisis. The EO specifically challenged council members to identify "how federal, state, and local resources can be better coordinated or redirected to meet the needs of Ohioans," which helped highlight the need for improved data sharing architecture in response to an acute public health crisis. A second DeWine administration EO (2019-02D) created the Governor's Children's Initiative, directing all state agencies that serve children to align efforts and improve cross-agency coordination. Together, these orders established clear, shared cross-agency goals in service of the Ohio community, without being overly prescriptive or heavy-handed about how goals would be accomplished. These orders also inspired the Ohio Department of Health to become interested in joining the IOP's open data portal, Data.Ohio.Gov, which launched in December 2020 to publish nearly 200 datasets, and five additional agencies have since followed.

Funding Data Integration Efforts

Investments in cross-agency data capacity and integration yield substantial returns by allowing governments to identify what works, coordinate to decrease duplication, and target resources to maximize impact. A single integration project that identifies multimillion dollar cost efficiencies could fund a cross-agency data sharing effort for years to come. States may also see benefits in terms of their ability to procure new resources. But despite these clear fiscal benefits, the costs associated with starting and sustaining an IDS may present barriers. Below we describe potential funding opportunities to consider at both the state and federal level that can be used to support data sharing and integration.

State-Level Funding Opportunities

To start, if you are pursuing legislation or executive action on data sharing, **it is essential to attach funding to the effort through the legislation itself and/or the state appropriations process** so as to avoid unfunded mandates. Funding for IDS staff and to support agency capacity for data sharing is just as important as funding for technology infrastructure, and this is true whether staff are housed within an executive office, an agency, or a stand-alone entity tasked with data integration. The IDS staff and infrastructure will not replace agency expertise or agency source data systems, and will cost significantly less. Most state-level IDS operate with annual budgets of between $250,000 and $5,000,000. Think of them as a small but mighty investment in R&D that will help your state become more data-driven and act in alignment to improve outcomes.

Many IDS also charge fees to help cover their costs. Specifically, some efforts charge external researchers and evaluators for access to linked datasets or for custom matching requests. For example, a researcher might pay to link survey data to administrative records to find out what happened to study participants years later. Other efforts provide data access (for approved uses) to agencies that contribute data, while imposing costs on agencies that do not participate. Efforts may have a flat rate for access or may determine costs depending on the size or complexity of a request. Regardless of how fees are structured, the data request process and pricing should be communicated clearly. It is also important to consider whether fees will impose cost barriers for less resourced groups that could otherwise benefit from data access.

Some data integration efforts also receive state funding through agency contracts for data and analytic services. For example, if an agency is conducting a needs assessment, program evaluation, or other analytic project that requires or could be strengthened by integrating data from other sources, the agency may choose not only to utilize the IDS for linkage, but also to contract out the analysis itself. This increasingly common arrangement allows many data integration efforts to sustain their own operations while directly serving the needs of agency data partners. However, it requires that IDS staff have content-specific expertise and work closely with agency partners throughout the process to ensure that data are used properly and findings are actionable.

In addition to state appropriations, fees, and agency contracts, **some IDS efforts utilize external support from foundations and university partners,** particularly during the early phases of their development. Foundation funding is often tied to ad hoc research projects and, while not typically an ongoing source for core IDS operations, can be used to get started, demonstrate proof of concept, and inspire state investment. Foundation-funded research projects may also support academic researchers who are lending their expertise to projects at no cost to the state. This type of partnership can enable efforts to flexibly expand or contract staffing depending on demand in a way that can be challenging for state governments. Some states have even succeeded in institutionalizing long-term partnerships with universities, relying on academic partners for ongoing technical and analytic support but maintaining ownership through strong data governance.

Federal Funding Opportunities

States may also look to federal legislation and grants for opportunities to support cross-agency data capacity, particularly when program objectives or evaluation requirements necessitate data sharing and integration.

The following are federal funding sources that states have leveraged to support IDS indirectly as a means of meeting their stated aims:

- Coronavirus Aid, Relief, and Economic Security (CARES) Act
- Family First Prevention Services Act (FFPSA)
- Health Information Technology for Economic and Clinical Health (HITECH) Act
- Medicaid's Mechanized Claims Processing and Information Retrieval Systems (90/10) rule
- Medicaid State Innovation Models (SIM) Initiative
- Performance Partnership Pilots
- Preschool Development Grant (PDG)
- Race to the Top Fund
- Race to the Top – Early Learning Challenge
- Substance Abuse and Mental Health Services Administration (SAMHSA) State Opioid Response Grants

The federal government has also made a direct investment in state data capacity through the State Longitudinal Data Systems (SLDS) program, which has allowed 49 grantee states, Washington, D.C., and five territories to construct new data systems for linking student data across programs. So far, most SLDS have limited linkages to early learning, K-12, post-secondary, and workforce data. However, in many states there may be an opportunity to build on existing SLDS capacity to add information about student health, access to human services, etc. (as KYSTATS and their partners continue to do in Kentucky, where data covers birth through workforce). In addition, state leaders are grappling with how to leverage the federal Workforce Innovation and Opportunity Act (WIOA) to support this work. WIOA asks states to target supportive services for priority populations with unique barriers to employment. Building cross-agency data sharing and integration capacity will therefore be critical for identifying these populations, connecting them with services, and evaluating their outcomes.

> ***Key consideration:*** Importantly, states that rely on federal program dollars to fund data sharing infrastructure may face a challenge in sustaining efforts as available resources shift, and must remain nimble to leverage new opportunities as they arise.

Funding Examples

Whether seeking state or federal funding for your IDS, we recommend finding options that are feasible in your context and that demonstrate the value of data sharing. It is also important to be flexible and opportunistic as you seek funding to scale. The following examples from North Carolina, Iowa, and Rhode Island show how states have done just that, braiding multiple funding sources to support cross-sector data capacity.

NORTH CAROLINA

The North Carolina Early Childhood Integrated Data System (NC ECIDS) was initially built with funding from the Race to the Top – Early Learning Challenge grant and has since been sustained by an annual state appropriation. The infrastructure was critical to North Carolina's success in receiving a federal Preschool Development Grant (PDG) in 2018, and NC ECIDS is now leveraging PDG funds to enhance their grant reporting and research request infrastructure, as well as to expand available datasets.

IOWA

Building from the legislative mandate established under Early Childhood Iowa, faculty at Iowa State University (ISU) helped stand up I2D2, an IDS staffed by ISU but governed by participating state and local agencies. The partnership between the state and university capitalizes on ISU's land-grant mission and expertise in data management, analytics, and security. This partnership also includes a governance approach that centers state and local priorities to ensure data are used to advance state early childhood program collaboration and effectiveness. While it was initially supported by several foundation grants and university start-up funds, like NC ECIDS, Iowa's I2D2 leveraged the PDG to expand their data infrastructure, enable a more comprehensive statewide needs assessment, and better coordinate services for children aged birth to five. I2D2 also receives state funding from agency contracts to support data management and analytic projects. State contracts have included comprehensive state and local needs assessments to inform improvements in home visiting and family support services; a study of connections between home visiting and child removals due to family substance use that is currently informing a statewide recovery network effort; and understanding unduplicated counts of children across preschool programming, such as universal pre-k, a targeted program for low-income families, child care assistance, and Head Start.

RHODE ISLAND

The Rhode Island Data Ecosystem has relied on multiple federal grant opportunities for IDS development, including Medicaid SIM, PDG, HITECH, and SAMHSA. In 2020, they also moved to partially fund their IDS using CARES Act dollars, and pivoted their focus to cross-agency pandemic dashboarding and analysis. Some analytic projects aimed to improve the state's initial pandemic response, such as evaluating staff and resource capacity in congregate care facilities where risk of COVID-19 mortality is higher than among the general population. Other projects were more focused on long-term recovery, including helping state agencies identify areas for cost savings and efficiencies to weather the resulting state budget shortfalls and cuts.

Strategies to Maximize the Efficacy of Executive + Legislative Action on Data Sharing

In the previous sections, we focused on how states can use legislative and executive action to promote data sharing, with options for funding efforts from start-up to scale. In this section, we share key preliminary strategies and steps that will set up your effort for success and ensure that legislative or executive action on data sharing is both ethical and effective. These strategies have been developed over many years of work with state and local governments and their partners, and will help you navigate the inevitable obstacles along the road to cross-agency data integration and use.

Seek to understand the full data sharing landscape. Learning about other data integration efforts, differentiating goals, and aligning work wherever possible can reduce unnecessary duplication of efforts and interagency competition. Here are a few steps to help you identify and learn from data sharing efforts in your community/jurisdiction:

- Identify who in the jurisdiction's leadership and within various agencies is invested in integrated data work: Who has supported this work in the past? Who is involved now? Who has authority over various initiatives and data holdings? Which key stakeholders have influenced or will influence the legislative process?

- Leverage what you learn about these key players to guide your thinking about where to strategically locate the IDS (e.g., executive office, agency, university partner) as well as who needs to be involved from the start. Sites tend to choose to integrate data within the agency that already holds the most statutorily protected data that will be needed for analysis, often a health and human service agency.

- Learn from the experiences of other data sharing and integration efforts in your jurisdiction, particularly those that were not successful. This knowledge can help you avoid prior missteps and navigate potential challenges that may surface (e.g., data quality concerns, disentangling issues of capacity versus trust, unstable political environment, resource constraints and competing priorities, public mistrust of data use).

Let purpose drive design. Data infrastructure should only be built, and data should only be shared, with a clear purpose. Developing cross-system relationships and establishing clear, shared motivations for data sharing will center the use rather than the technical components, and will prevent the effort from becoming just another technology procurement project. Here are a few steps to get you there:

- Spend time upfront ensuring that expectations are aligned across participating offices and agencies. Establish a clear mission or purpose statement to refer back to when making decisions about how to gather and use data. While disagreement may be unavoidable, defining a clear purpose at the outset can be useful in navigating roadblocks and difficult decisions.

- Create spaces for cross-agency learning, and include both analytic and front-line staff. Analytic staff have an important perspective as end users of integrated data, while front-line staff know why, when, and how data were originally collected. Both perspectives are essential in creating infrastructure that actually does what you want it to do.

- Design with flexibility in mind. The field of data integration is growing and changing rapidly, as are the needs of agencies and the people represented in the data. Your technical approach should allow you to be nimble and responsive to the shifting environment. Flexible *design* means avoiding reliance on one vendor or single source procurement. Flexible *infrastructure* means drawing on different tools, at different times, for different aspects of the work.

Gather and incorporate meaningful contributions from community stakeholders and impacted populations. Those most impacted by policies and by data sharing – e.g., children and families receiving services, adults applying for public benefits, or front-line staff working directly with clients – should be engaged throughout the process and have a say in how data are used. Many cross-agency efforts struggle with this, in part because alignment between leadership priorities and community demands is not always obvious and because authentic community engagement is resource intensive. Here are steps you can take to lead the way:

- Assess public perceptions of past and present data sharing efforts, and be sure to take note of prior harms as well as disproportionalities in how communities experienced those harms. This process has led some data integration efforts to explicitly adopt a racial equity lens in their work. For more on centering racial equity in data integration, see our toolkit, *A Toolkit for Centering Racial Equity Throughout Data Integration*.

- Make a realistic plan to consistently engage the community. Trust and collaboration are built over time (and often with missteps and course corrections along the way). Regular opportunities for bidirectional communication can help establish trust and repair past harms. Remember – you don't have to do everything right away, but it's important to do something today. For more on how to facilitate community conversations and public engagement, specifically around issues of privacy and confidentiality, see our toolkit, *Nothing to Hide: Tools for Talking (and Listening) About Data Privacy for Integrated Data Systems*.

- Document ways in which community engagement has strengthened data efforts to make the case for sustained investment in these activities. We recommend tracking the number of stakeholders engaged, the growth of community partnerships, and the impact of analysis on outcomes identified as important by community members, in addition to those identified as executive or legislative priorities.

Build standard processes and protocols for data access and use, but allow for flexibility as priorities shift. A key benefit of developing IDS is the standardization of processes, guidance, and templates that lead to a streamlined approach to data use across the jurisdiction. However, finding the balance between standardization and flexibility can be challenging. The following steps can help:

- Think through who should oversee decisions about provisions that have cross-sector impact. Some considerations – like privacy, security, anonymization procedures, and other compliance-driven actions – are often formalized within policy, statute, or rule. Other considerations, like metadata standards (e.g., standard gender categories), legal frameworks, and data review processes, are better suited to flexible protocols and processes that can be continually improved upon.

- Approach data use policies with an eye towards growth, because new priorities, opportunities, and challenges will inevitably emerge. When writing policy and legislation, consider both the current context and the vision for future data work.

- If your ideal legislation is not politically feasible right now, work with other stakeholders to identify strategies and opportunities to maintain progress and growth. The field of data integration is ever-changing, so course corrections are expected.

- Don't let change and transition catch you by surprise. Think through your project management approach and standardization to ensure that data security, quality, and documentation are not compromised when staff changes occur, when there are new high priority data requests, or when community conversations surface an emergent concern or question. Codifying these practices early on not only establishes the value of a systemized approach but also helps to develop a sense of shared responsibility around managing and maintaining data.

- Overall, be flexible and creative within the current policy context and don't let the "perfect" stand in the way of the "good."

Strategize for sustainability. There are numerous ways to grow and strengthen a data sharing and integration effort. Here are a few approaches that efforts have taken to enhance sustainability:

- Seek out small and quick wins during early stages of development in order to build momentum and support for long-term sustainability.

- Cultivate an organizational culture that supports data use and literacy alongside the development of policy and legislation to help the current effort move forward, weather inevitable political changes, and build support for evidence-based decision-making long term.

- Establish a Chief Data Officer (CDO), which is increasingly common at the state level, to adopt a holistic and focused approach to data use by and for agencies. As referenced in the example above from the state of Connecticut, CDOs are particularly helpful in developing a broad view of data-driven efforts across the state, differentiating and aligning various streams of work, and delivering the "right data to the right people."

- Document the process of IDS development and lessons learned with regard to governance and to legal, technical, and political considerations. This documentation is valuable institutional memory that will support inevitable staff transitions and stewardship of resources. The goal is to fail fast and course correct. Such shifts are not possible without strong process documentation.

- Above all, prioritize privacy and security throughout the policymaking process, and never compromise the trust of stakeholders or the privacy of those whose data are being shared or integrated. This is essential to creating a sustainable and ethical data sharing and integration effort.

There's no clear blueprint that will work to advance data sharing in every context. Nevertheless, the strategies and examples outlined in this brief aim to support policymakers and other IDS stakeholders in the process of building and institutionalizing data capacity. This process is not linear, and a single decision point rarely makes or breaks a data sharing and integration effort. Rather, impactful and sustainable data sharing efforts build and change over time by employing these strategies opportunistically and iteratively to maximize data use for public good.


AISP