# AISP

Network Survey Series

CAPACITY

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 equity. The following brief shares lessons from a February 2023 survey of 37 state and local data integration efforts in the AISP Network. All sites surveyed have some data governance and data sharing agreements in place, but vary widely in maturity, scope, purpose, and approach. Among the 37 survey respondents (19 states and 18 local efforts), there is representation from every major region of the continental U.S. and high representation of coastal states and cities.

This brief explores the people, relationships, and resources necessary for an effort to operate.

AUGUST 2025

Penn
Social Policy & Practice
UNIVERSITY of PENNSYLVANIA

Introduction

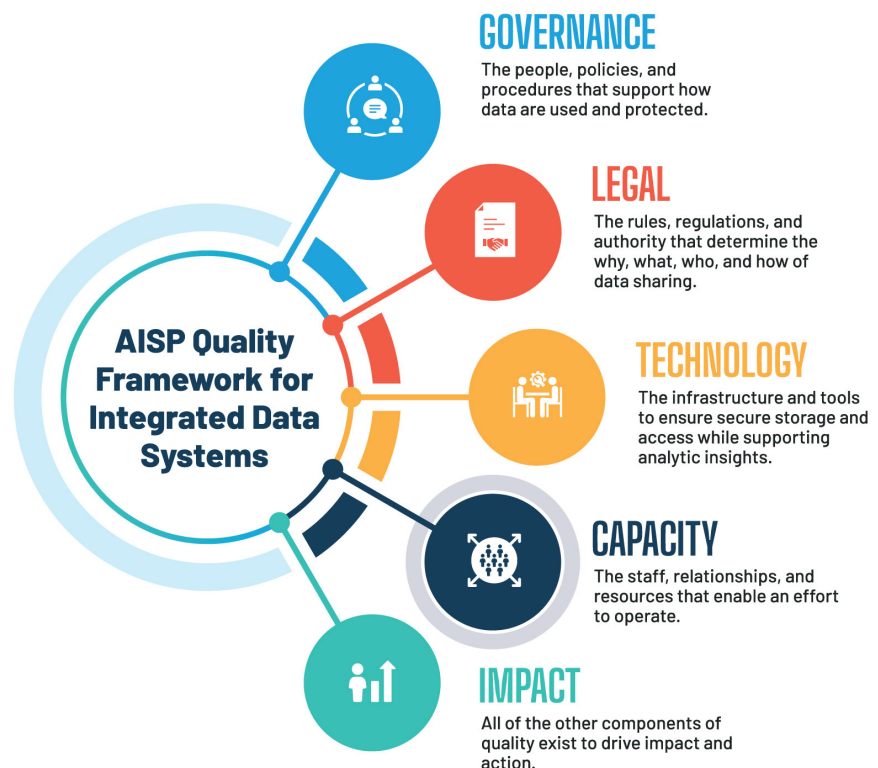## Cross-sector data sharing capacity
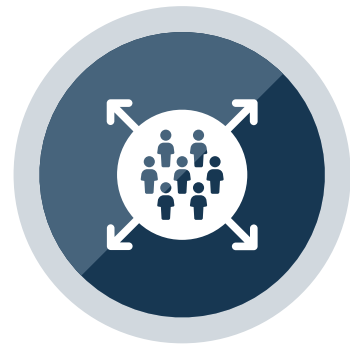
> Capacity: the staff, relationships, and resources that enable an effort to operate.
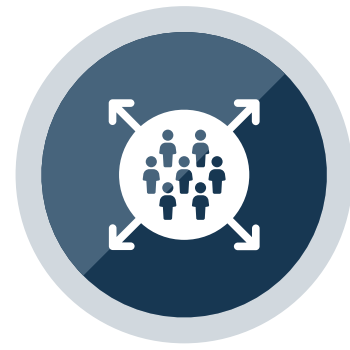
Capacity is the human, relational, and material resources that enable a data sharing effort to implement governance, execute legal agreements, build technical infrastructure, and, above all else, demonstrate impact. Building shared capacity requires effective leadership and planning, as well as sustained commitment from data partners. The process is iterative and often feels like a delicate dance. The goal is to keep growing the ability to meet partner goals without overpromising on what can be achieved with current data sharing capacity. Each successful proof of concept demonstrates value and can fuel advocacy for more resources.

In this brief, we will explore different approaches to capacity across sites in our network, including how sites staff the technical and relational aspects of integration, as well as how they raise funds and distribute resources.

Capacity is the fourth of five components of quality for integrated data systems (IDS). For more on other components of quality, visit **https://aisp.upenn. edu/quality-framework-for- integrated-data-systems/**

AISP Quality Framework for Integrated Data Systems

GOVERNANCE
The people, policies, and procedures that support how data are used and protected.

LEGAL
The rules, regulations, and authority that determine the why, what, who, and how of data sharing.

TECHNOLOGY
The infrastructure and tools to ensure secure storage and access while supporting analytic insights.

CAPACITY
The staff, relationships, and resources that enable an effort to operate.

IMPACT
All of the other components of quality exist to drive impact and action.

Survey Analysis

STAFFING LEVELS

Capacity is all about maintaining the right mix of the people and resources necessary to achieve lasting impact. The staff of a data integration effort are the "doers" who carry out daily operations. This team should include diverse competencies to support both the relational and technical aspects of data sharing. The number and configuration of staff vary widely, depending on the structure, purpose, and **management model** of an effort.

Survey results show that, on average, IDS operate with lean staffing, doing a lot with a little. The median number of full-time staff employed across sites in the AISP Network is four. That said, there are extremes on both ends of the spectrum, with several sites reporting no dedicated staff (yet), and other sites reporting 40+ full time employees.

STAFFING MODELS

The structure of the organization(s) that host and lead an effort—also known as the management model—will affect the who and how of staffing. For an example, if the goal of an effort is to support care coordination across a variety of health and human service programs, the umbrella agency that manages the majority of those will be an easy choice for host. However, many efforts find that decisions about hosting are more complicated, particularly when many partners are contributing data.

To understand variation across efforts, we looked at how different models staff three core categories of activities:

- **Facilitating governance (e.g., partner engagement and procedural oversight, including development and implementation of legal framework)**

- **Conducting data analysis (e.g., research methods, analytic tools, and insights)**

- **Managing technical infrastructure (e.g., data storage, transfer, integration, and access)**

We asked efforts to identify what kind of organization staffs the three core functions, or whether those functions were split across partners. Over half of the efforts in the AISP Network are hosted by a lead government agency or executive office that staffs all three functions. Three sites are hosted and entirely staffed by a university, while four others use a university partner for one or two functions. Six relied on a public or private sector service provider to host their technical infrastructure. As of 2023, two sites had a nonprofit partner who staffed their data governance activities, but only one had a nonprofit partner who managed the technical infrastructure. We expect our next survey to reflect an increased number of non-profit led efforts, as the prevalence of this model is growing.

BENEFITS & CHALLENGES

Each management model comes with benefits and challenges that are highly context specific. We often recommend building capacity within government so that the infrastructure is a public good and the people who manage day-to-day activities are situated as close as possible to both decision-makers and those collecting data through program administration. That said, there are good reasons some network members choose different models and host organizations for their data sharing. We typically bucket them into four categories: executive-led, agency-led, university public partnership, and nonprofit-led.

Let's begin with an **executive-led model**, which refers to a government-led effort hosted within an executive office (e.g., office of the mayor, county supervisor, or governor) or office of management and budget. Given their size and influence over local or state affairs, executive offices are typically well-resourced. They may benefit from a more comprehensive, enterprise view of government programs and systems and thus offer a vision for alignment with other evidence-building and data-related activities. Staff often enter the work with a strong sense of purpose and are well-trained and well-tooled, regardless of the jurisdiction's overall capacity for data use. At the same time, an executive office is susceptible to elections, appointments, and changes in leadership and/or leadership priorities. There can also be challenges building trust with agency partners who may see the executive office staff as political actors or worry about their influence over spending.

NYC's **Center for Innovation in Data Intelligence (CIDI) is housed in the Office of the Mayor where they primarily perform policy research and evaluations. Proximity to leadership helps ensure CIDI's work is aligned to the city's goals and needs. In South Carolina, the Health and Demographics Section of the Revenue and Fiscal Affairs Office hosts an integrated data system tasked with turning data into information and information into knowledge. This agency benefits from centralized analytic capacity close to state government operations.**

Another model is **agency-led**. This model has many variations. Across our network, agency-led models include those led by health and human services (HHS) agencies, centralized IT, or standalone data integration agencies created through statute. As outlined below, each of these has different benefits and challenges, though, across the board, this model benefits from the legal authority to hold protected data and receive on-going appropriations. Challenges posed by this model generally arise from the complexity and resource-competition inherent in an agency-led context.
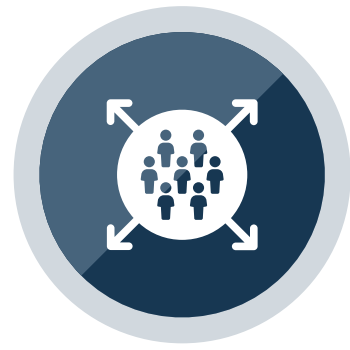
- HHS agencies are often the largest agencies with many programs consolidated underneath their leadership, which helps from a practical and legal standpoint. Given their high volume of service delivery, HHS agencies have proximity to both the people represented within the data and the practitioners implementing programs and policies. They are often well-staffed with trained analysts, due in part to their access to federal funding. That said, HHS agencies themselves may be siloed from other agencies, and are often constrained by competing demands and a complex web of funding streams.

- Conversely, standalone data integration and analysis agencies benefit from a clear purpose that enables staff to focus strictly on data sharing and use projects across the enterprise. However, their efficacy depends on resourcing—whether they actually have the human and technical power to deliver on their mandates—as well as their relationships with other agencies and perceptions of their legitimacy.

- Centralized IT agencies offer the benefit of advanced security and technical capabilities among staff. They may also be seen as a neutral agency among data partners. However, because centralized IT staff typically lack a human services background, it can be challenging to ensure this model stays aligned to programmatic data needs and community impacts.

The **Kentucky Center for Statistics (KYSTATS)** operates out of the state Education and Workforce Development Cabinet—a standalone agency within a commonwealth department that oversees education and workforce. KYSTATS is responsible for developing, maintaining, and utilizing Kentucky's longitudinal data system and the Labor Market Information Office, providing them with clear purpose, direction, and connection to those they serve.

The **Allegheny County Data Warehouse** in Pennsylvania is hosted by the county Department of Human Services' Office of Analytics, Technology, and Planning. It was initially created by consolidating publicly funded human services data from behavioral health, child welfare, intellectual disability, homelessness, and aging. Over time, it expanded to include data from other sources and has become a well-known national model. The primary purpose of the Data Warehouse is to improve services to clients, but it is also used to improve the ability of workers to perform their jobs and to support county decision making.

University public partnerships can be an attractive model given that they lend additional capacity to government. University partners operate as contractors to government but may come with in-kind resources, contribute specialized methods expertise, and offer flexible faculty, staff, and student labor. In some contexts, they are also well positioned to serve as a neutral third-party or "data Switzerland" so that no one agency is perceived to hold more power than another over data use decisions. At the same time, this model tends to be less responsive to urgent data needs within agencies and may experience tension between the priorities of publishing and practice. As the term "ivory tower" suggests, academic or university-based organizations often face barriers to developing trust and social license and must demonstrate their commitment to transparency and public priorities.

> **In Iowa, the university manages the governance, data analytics, and day-to-day operations, with priorities set by agency partners. This arrangement allows the data sharing effort to be nimble, staffing up or down as resources ebb and flow in a way that can be difficult in a state agency context. The IDS also relies heavily on student and post-doc support, which expands their capacity and helps build the pipeline of data users who are trained to use linked administrative data. Read more about their team here.**
>
> **In North Carolina, another staffing model relies on a slightly different type of university partnership. Beginning in 2019, North Carolina DHHS, the state's consolidated health and human services agency, hired a leading academic from Duke University to serve as the agency's chief data officer. From 2019-2024, they also contracted with AISP's Director of Training & Technical Assistance to develop and implement enterprise data governance. Though both academic partners have since left the data office, the collaboration continues to pay off. The data office they built and the NCDHHS Data Sharing Guidebook they created have transformed state processes and become an internationally recognized model.**

Finally, the **non-profit-led** model, while newer to the data sharing ecosystem, is becoming more common. Often relying on a collective impact non-profit such as a United Way, this model benefits from proximity to community and involvement with service provision, which may help build social license and connect directly to data-driven advocacy. Traditionally, very few non-profits have had the resources and technical infrastructure to manage advanced data security and linkage themselves. New technologies have begun to lower this barrier to entry. As with the university public partnership model, non-profits may also experience challenges establishing trust and sustainable partnerships with government agencies to broker data access. Legal frameworks are often most challenging in this context.

The **Baltimore Youth Data Hub** is hosted by Baltimore's Promise—a city-wide non-profit collaborative that uses data to improve outcomes for children and youth. They took on the role of hosting an IDS after a careful analysis of other options and many conversations with local partners, who felt neither government nor university partners made sense in their context. Baltimore's Promise has built trust, as well as a team of staff to do the work. The Youth Data Hub is now known as a leader in developing community-led data governance, having spent the last eight years ensuring their engagement with people represented in the data system would be robust, on-going, and well-compensated. Learn more about their work **here**.

Regardless of the management model, when asked about their greatest operational challenges, sites always put staffing near the top of the list. When we asked survey respondents specifically about staffing, the following themes emerged:
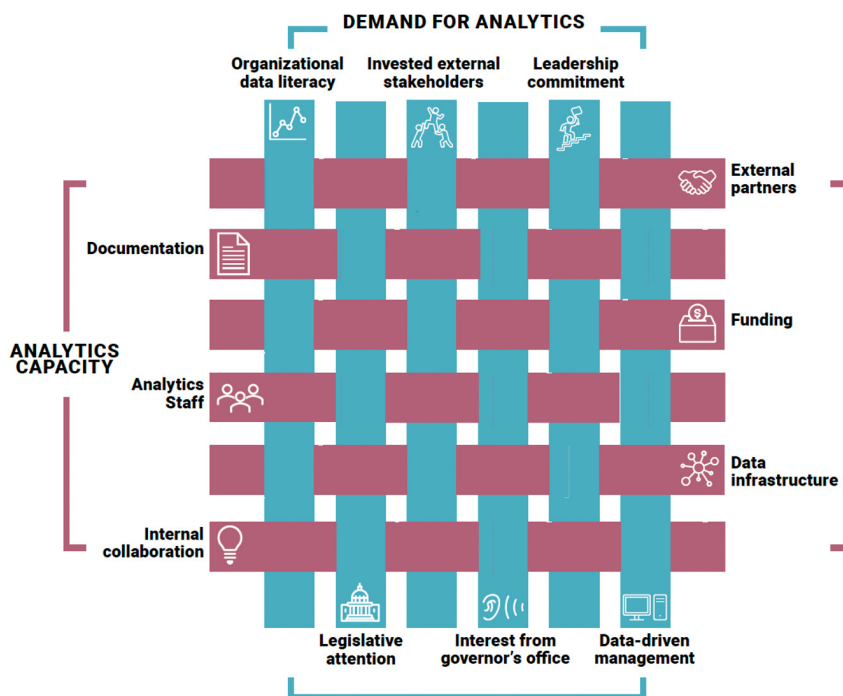
- **Retaining talented staff is a challenge; frequent burnout and turnover slow down forward progress**

- **Salary bands make it difficult to compete with the private sector when hiring for technical roles**

- **Research and grant funding cycles often drive staffing, which can make continuity and sustainability a challenge**

- **Staff are often stretched thin across multiple roles and responsibilities, with some respondents reporting that none of their staff are entirely dedicated to the IDS work**

- **Professional development and peer networking opportunities may be limited, as IDS teams are often small and highly specialized**

STAFFING STRATEGIES:
BUILDING BOTH CAPACITY & DEMAND
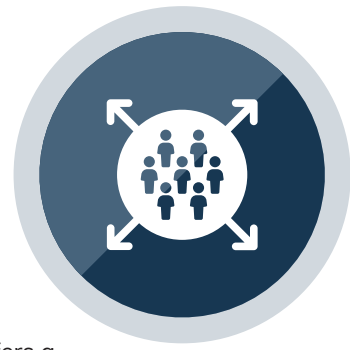
When it comes to staffing challenges, we highly recommend starting with a resource from our partners on the TANF Data Collaborative. ***Strengthening Analytics in Government Agencies: A Toolkit for Sustainable Data Use*** offers a variety of capacity-building strategies for data efforts helmed by government agencies, whether focused specifically on TANF data or data sharing and use more broadly. This resource uses the metaphor of interwoven strands to describe how data use is strengthened by a mix of data consumers and capacities, with the red strands representing key components of data or analytics capacity, and blue strands representing key components of demand for data.

**Figure 2.1. Data Use is Strengthened by a Mix of Data Consumers and Capacities**

Strengthening Analytics in Government Agencies: A Toolkit for Sustainable Data Use

The Toolkit offers several resources to help with staffing a data team or data sharing effort, including sample language for analytic job descriptions, sample interview questions, guidance on identifying the right training program for your team's needs, and a documentation checklist to help teams weather turnover. Explore the Toolkit **here**.

One key tenet of the toolkit is that demand for data is as important as our ability to link data. Without demand, our investments in data capacity languish. In Indiana, the **Management Performance Hub**, a standalone, cabinet-level government agency tasked with supporting data integration and use, has trained more than a thousand state employees on just that. The Data Proficiency program uses a progressive set of online trainings and badge levels to motivate in-house personnel to get more comfortable with data use. The free program is also available online for anyone interested in leveling up their data skills. Explore it **here**.

DATA SHARING LEADERSHIP:
THE ROLE OF CHIEF DATA OFFICERS
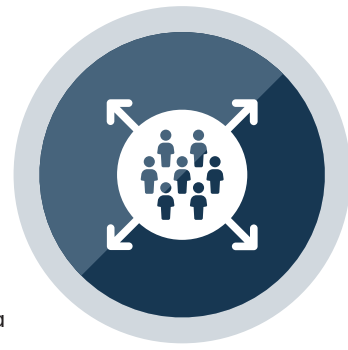
To support well-trained staff, a data sharing effort needs strong leadership to guide a positive culture for and commitment to evidence use. As former Indiana Chief Data Officer Josh Martin likes to say about data use in government, "Not everyone is a data person, but everyone has a data job." Chief Data Officers (CDOs) can play a key role in unifying the components and activities of a data sharing effort to make meaningful impacts. More specifically, CDOs can help ensure an IDS is tied to a larger state strategy and aligned with complementary efforts in their context. Chief Data Officers hold many responsibilities, with cross-sector data sharing often being one of many strategic priorities. CDO roles and how they are situated vary and may impact the kind of support they provide to data sharing efforts.

One in three sites who responded to the survey reported that a CDO was involved in their data sharing effort. Some CDOs who are involved are in a leadership role with decision-making authority to direct the activities of the effort, while others serve a key role in governance and security activities specifically. A handful of sites note that their CDO is aware of or provides guidance to the effort but is not currently hands on, and several expressed a desire to forge a deeper relationship. For more about the role of CDOs, check out the **State Chief Data Officers Network** operated by Georgetown's Beeck Center.

In Connecticut, two CDOs have contributed leadership to state data sharing efforts during their tenure with impressive results. The state's first CDO, Tyler Kleykamp, led open data efforts and supported legislation to study, design, and implement interagency data sharing. In 2019, Scott Gaul led the first report on interagency data sharing, as legislated by Public Act 19-153 and oversaw the transition of the P20 WIN system from CT State Colleges and Universities to the Office of Policy & Management, supporting the expansion of data governance activities, data holdings, and the modernization of technical infrastructure.

Data use for state initiatives are formally proposed through the CT State Data Plan, which can be accessed here.

IDS BUDGETS

IDS vary in terms of cost based on purpose, management model, and scale, but they all require sustained funding over time. While interest and investments in data sharing are increasing as more agencies and funders recognize the substantial returns, many efforts continue to work under substantial resource constraints, doing a lot with a little. Resource constraints will likely increase in subsequent budget years, as federal budget cuts impact state and local governments and their IDS.
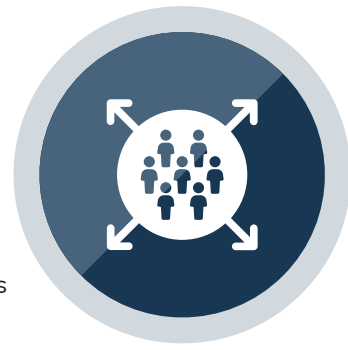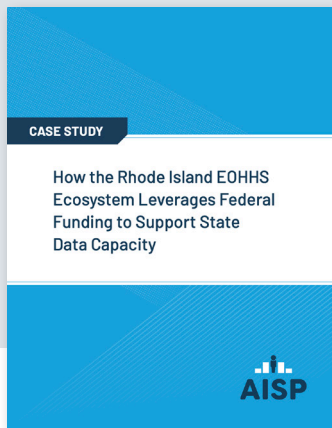
Annual Budget Ranges By Geography
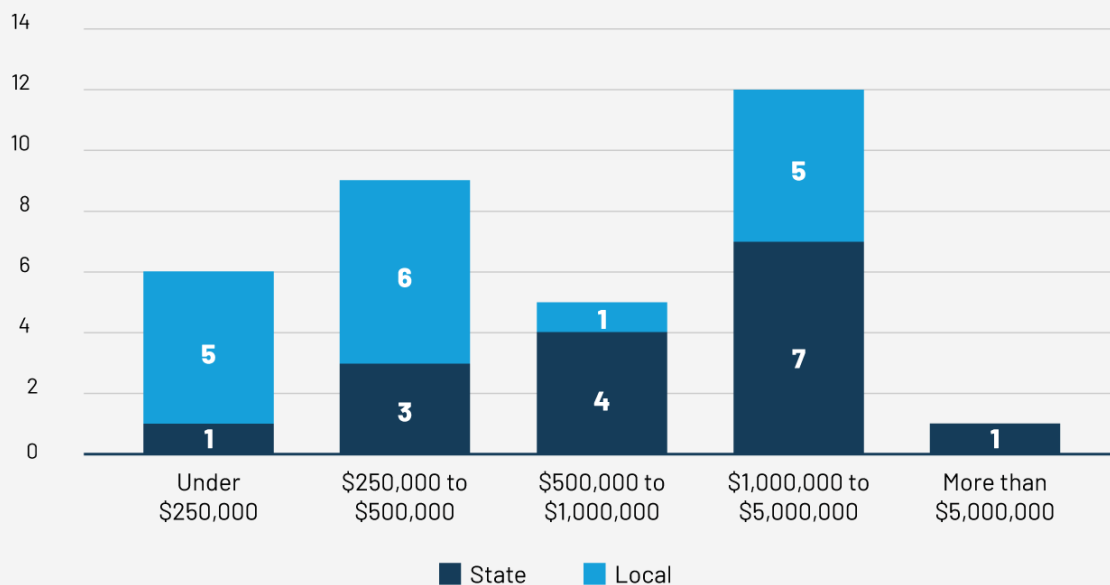
Though there is a wide range, many state data sharing efforts report budgets between $1M and $5M annually. Local data sharing efforts are also varied, in part due to major differences in size. Five local efforts reported budgets under $250,000, six spend between $250,000-$500,000 annually, and another 5 spend over $1M. There are also five efforts in the network who report having no budget. Some of these are still in the planning stage and devote staff time towards data integration within current job responsibilities. Others reporting no budget describe their data integration as integral to their operations, often as part of a shared services model.
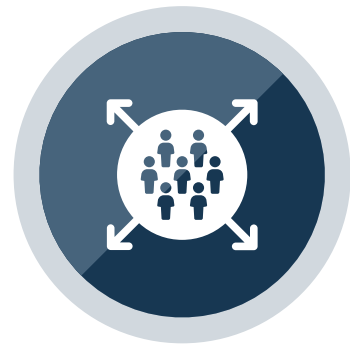
FUNDING FOR DATA SHARING

Most efforts rely on multiple funding sources, including state appropriations, federal funding, philanthropic grants, and fees for data access. This is good practice, as diverse funding sources enable data sharing efforts to manage and sustain operations in the event of a revenue loss or as priorities shift.

CASE STUDY

How the Rhode Island EOHHS Ecosystem Leverages Federal Funding to Support State Data Capacity

AISP

This AISP case study offers a deep dive on how the Rhode Island EOHHS Data Ecosystem has blended federal funding streams to support their state data capacity, drawing on Medicaid SIM, PDG, HITECH, and SAMHSA, among others. Explore the case study here.

Of those surveyed, seven sites still rely on just one funding source: of these, three reported that they are entirely funded by a public agency line item, two relied solely on foundation grants, one was funded by a single federal grant, and another was funded exclusively by contracts with agencies to provide data services.

As efforts have sought to diversify their funding portfolios, federal funding has increasingly become a more common source for supporting state data infrastructure. The National Academy of Public Administration recently released a field guide to funding IDS and evaluation capacity, which both offers guidance on what is allowable with federal sources and outlines how some states are moving to think about data infrastructure and integration as a shared service. Explore the field guide here and access the webinar recording from our conversation about this topic here. Once again, AISP Network members are leading the way! Of course, it must also be noted that these same innovative leaders will likely soon feel the impact of federal changes in funding for programs and data systems, so future surveys may capture change in this area.

Looking Ahead
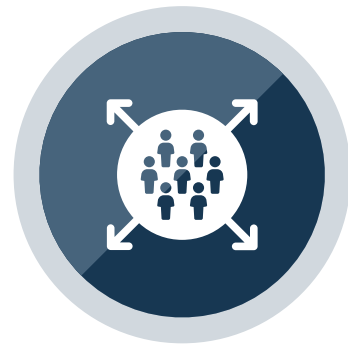
The success of a data integration effort is driven by the people who do the work, day in and day out. No matter the size or shape of an effort, it is first and foremost a relational endeavor and one that needs to be sustainably resourced and effectively staffed. No matter where you are on your data sharing journey, we hope this brief offered useful insight into the many ways to establish and sustain capacity.

WONDERING WHERE TO GO NEXT?

For guidance on communicating with policy and decisionmakers—those whose support we need to resource data efforts—check out AISP's 2025 brief titled **"Integrated Data Systems for Policymakers."** This piece highlights real-world examples from network sites and offers guiding principles to help policymakers become data leaders. The aim is to help any IDS gain buy-in and build capacity.

If this document raised questions for you about IDS approaches broadly, check out our **network map**.

A NOTE ON THE DATA

To improve data quality, initial survey results have been supplemented with document review and qualitative research. If you have questions, please reach out to the AISP team at **aisp@sp2.upenn.edu**.

Suggested citation: Berkowitz, E., Jenkins, D., Hawn Nelson, A. (2025). Network Survey Brief: Capacity. Actionable Intelligence for Social Policy, University of Pennsylvania. **www.aisp.upenn.edu**