

FINAL REPORT

OCTOBER 2021

Improving Data Infrastructure to Reduce Firearms Violence

Chapter 7. Practical Steps for Building State Capacity and Infrastructure to Use Data for Evidence-Based Decision Making

Nancy Potok, PhD and Nick Hart, PhD | NAPx Consulting and Data Foundation

Editors:

John K. Roman, PhD

NORC at the University of Chicago

Philip Cook, PhD

Duke University



Chapter 7. Practical Steps for Building State Capacity and Infrastructure to Use Data for Evidence-Based Decision Making

Nancy Potok, PhD and Nick Hart, PhD | NAPx Consulting and Data Foundation

Executive Summary

This paper provides a roadmap based on specific examples taken from successful federal and state experiences that can help states and local governments build capacity and infrastructure to use data for evidence-based decision-making. Each state may be starting from a different level of experience and capability, and each has a different legal and regulatory landscape, but the approaches provided here are adaptable and scalable. The paper provides examples from recent federal efforts as well as successful federal-state partnerships and highlights from states that have put important foundational building blocks in place. Each section of the paper provides key takeaways or lessons learned. An essential element of all successful models is that they start with a value proposition. Experience has shown that there must be a clear payoff for the agencies providing data and the public they serve, for improvements in data infrastructure and analytical capabilities to be sustained. The examples presented below show a clear value, rather than creating processes without a focus on solving specific problems and knowledge gaps faced by agencies.

Public agencies face complex, multi-dimensional challenges in understanding how government resources can be deployed to achieve the best outcomes for their constituents in areas such as health, safety, education, jobs, social equity, and housing. Many of these issues are interconnected and require a holistic view that crosses traditional organizational boundaries. Similarly, understanding how to measure and improve the welfare of children may involve healthcare, housing, education, and crime data. Agencies need to demonstrate that tax dollars are, in fact, achieving stated program outcomes for the public.

Improved data science and computer science methods are enabling more timely, relevant data analysis. However, changing the way in which government agencies share and use data is not a simple undertaking. The U.S. Commission on Evidence-Based Policymaking noted in its report to Congress (2017) that sustained, concerted efforts are needed to strategically build the capacity and infrastructure to analyze data in meaningful ways. Goerge (2019) noted six key factors that lead to effective engagement with and among state and local governments to

access and provide data, including providing short term tangible benefits and a vision that connects authentically to the real-world experiences of government leaders.

Building capacity to better utilize data is a continuous learning process that has already been providing value. The federal government and many states have taken significant steps to build the needed capacity to produce information that can be combined from many sources across agencies to better serve the public. This paper discusses the important elements of a successful data strategy and implementation of increased capacity to analyze and utilize data.

These include the following:

- Data leadership and strategy
- Legal and regulatory authority and policies (framework)
- Data governance
- Transparency of processes and uses
- Accessibility for data use
- Oversight and accountability
- Sustainability

The discussion for each element includes specific examples that illustrate how each has been put into practice at the federal and state-level. The examples can be used to model similar efforts in other states and are designed to stimulate thinking and action to promote evidence building with data sharing across agencies, states, and other government jurisdictions. Also included is a detailed look at the topic of firearms violence. The paper concludes with a list of ten actions that states can take to maximize their opportunities for success. These include:

1. Establish and empower state and local data leadership roles
2. Identify local knowledge needs and data gaps based on priorities
3. Review legal authorities to share, protect, and use data
4. Provide feedback to federal partners about data needs and priorities
5. Establish a Cross-Agency Governance Process and Data Strategy
6. Execute on Data Governance by Integrating Local Data Where Feasible to Prepare for Linkage and Sharing
7. Institute transparent approaches to using data
8. Partner with existing entities to promote rapid progress and pilot projects to demonstrate value
9. Establish responsive engagement processes and oversight mechanisms
10. Plan for and implement robust privacy safeguards

From “A Blueprint for a US Firearms Infrastructure”

“The development of a rigorous empirical research base to inform both citizens and policymakers requires a robust and sustainable data infrastructure. The most enduring data infrastructure is one that is comprehensive, flexible, and nonpartisan. Nowhere is that data foundation more needed than in the realm of firearms violence—reliable data are a critical bridge to effective policymaking that improves public safety by reducing the number of firearm accidents, suicides, homicides, and assaults.”

Roman JK. 2020. A Blueprint for a US Firearms Infrastructure. Bethesda, MD: NORC at the University of Chicago.

Introduction

States and localities face significant challenges in areas such as improving public health and safety, creating jobs and economic opportunity, tackling rising crime rates, addressing a lack of affordable housing, and measuring the long-term outcomes for students at public education institutions. The public depends on services delivered by a multitude of agencies at the federal, state, and local level. Complex, multifaceted, long-term challenges require approaches that are holistic and informed by information that crosses the traditional functional boundaries of agencies. Some state and local governments have taken steps to make better use of the data they collect from the public to improve service delivery and the policies behind them. These steps—often involving partners in academia, the private and non-profit sectors, and the federal government—have yielded positive results. This paper attempts to provide a roadmap of how state and local governments might learn from these successes: specifically, how to tackle systematically the seemingly overwhelming task of rationalizing and utilizing data to understand how to implement improvements and achieve better outcomes for the public.

Long-standing gaps in the national data infrastructure—including at the state and local levels—inhibit analytical capabilities that are long overdue and much-needed for contemporaneous policy debate, on topics ranging from the most efficient workforce development strategies to the most effective approaches to reducing firearm-related violence. An important illustration of this fragmentation and the gaps in data was identified by an expert panel commissioned by Arnold Ventures and convened by NORC at the University of Chicago. The expert panel issued three reports. One main finding in the [first report](#), on the state of firearms data, was that data on firearms violence—including health, public health, and crime data—are disordered and highly segmented, contributing to a lack of information that might otherwise inform public policy decisions.

At the federal level, many steps have been taken to focus attention on improving evidence-based policymaking and building capacity for data-informed decision-making. At the same time, several states have made great strides in improving access to and use of information not only to guide policy making but also to respond quickly to crisis situations, such as the COVID pandemic. Prioritizing better use of data provides an opportunity to modernize and update systems and analytical capabilities to reflect the public's expectations. Such a focus on data is an increasingly common practice among government officials.

Many states have learned hard lessons on what works and what doesn't as they have put in place capabilities to better use data. This paper attempts to consolidate and share examples of best practices that can save precious time for states and localities that want to make faster progress and avoid "reinventing the wheel." Each state is different. Laws, regulations, state government organization, local culture, and the maturity of existing capabilities may vary considerably among states. However, the information in this paper is general enough that it could be adapted in whole or in part to each state.

The paper provides an overview of the context for data modernization in 2021 and beyond, identifying successful national and state models that can be leveraged for further progress. It opens with background on federal developments or steps to advance the ability to use, share, analyze, and inform decision-making with data. Some steps can be replicated at the state-level, and some states have made significant progress with only a partial set of foundational steps in place. Next, this paper provides examples of successful state efforts and includes a brief overview of key characteristics for effective state and local capacity based on successful models. It discusses how those approaches can leverage federal investments and statutory authorities to promote data sharing and linkage. The paper then looks at possible examples of how states may use firearms violence-related data, with a roadmap for states to get started either building capacity or expanding existing capabilities. The paper concludes with key takeaways learned from the experience of existing intergovernmental efforts.

Finally, the paper emphasizes one key aspect of all successful models; that is, they start with a value proposition. Experience has shown that to marshal the resources needed to gain high-level attention, allocate sufficient funding and people, provide training, and change the culture in agencies where data silos have existed for decades, there must be a clear payoff for the participants, including the public. The examples presented here show a clear value for the agencies and the public they serve, above and beyond demonstrating compliance with mandates that may come from the federal government or the governor's office. In the examples, new processes focus on solving specific problems and knowledge gaps that agencies face.

What Can Be Learned from the Federal Experience

The federal government has invested significant resources over the last decade to improve the ability of agencies to use data in policymaking. Individual agencies recognize that they can learn more from better use of their own program data and by combining their data with data from other sources such as statistical agencies and the private sector. However, scattered individual projects have yet to result in needed systemwide improvements. There are significant legal and cultural barriers to sharing data across agencies, as well as a lack of resources to create a data sharing infrastructure. Over the past six years, the situation began to change in important ways. The sections below describe some of the major initiatives that have propelled increased capability for evidence building at the federal level.

Foundations for Evidence-Based Policymaking.

In 2016, the U.S. Congress passed a bipartisan bill to create a Commission on Evidence-Based Policymaking ([Pub. L. 114-140](#)). The Commission garnered bipartisan support because members of Congress, regardless of ideological viewpoint, realized that federal tax dollars were not being maximized, and there were few ways of knowing which programs were effective and why.

The bill established a 15-member Commission appointed by the President and Congressional leaders with consideration given to individuals with expertise in economics, statistics, program evaluation, data security, confidentiality, or database management. The Commission was required to submit a report and recommendations to Congress in 18 months, based on conducting a comprehensive study of data inventory, data infrastructure, database security, and statistical protocols related to federal policymaking and the agencies responsible for maintaining data. Several purposes for data were cited in the bill, such as to, “...*determine the optimal arrangement for which administrative data on federal programs and tax expenditures, survey data, and related statistical data series may be integrated and made available to facilitate program evaluation, continuous improvement, policy-relevant research, and cost-benefit analyses; make recommendations on how data infrastructure, database security, and statistical protocols should be modified to best fulfill those objectives; and make recommendations on how best to incorporate outcomes measurement, institutionalize randomized controlled trials, and rigorous impact analysis into program design.*”³² In addition, the Commission was to consider whether a clearinghouse for federal program and survey data should be established.

The Evidence Commission was required to consult with several federal agencies about their areas of responsibility. As part of its mission, the Commission held seven public meetings and three public hearings around the country, to gather valuable insights from the public. In 2017, when the Commission issued its [final recommendations](#) to Congress and the President, it

³² Evidence-Based Policymaking Commission Act of 2016, Pub. L. 114-140

provided a roadmap for improving data and analytical capabilities at the national level (CEP, 2017). The recommendations noted the need to bolster state and local government capabilities while supporting a broader ecosystem and infrastructure relevant for all levels of government. Many recommendations focused on leadership roles, data governance, data sharing, program evaluation, and planning processes to build evidence. Although the Evidence Act did not provide resources to implement the law's mandates, agencies have prioritized their activities and worked within existing budgets and staffing levels.

Value to the Government of Evidence Building

Without evidence, the federal government is an ineffective fiduciary on behalf of the taxpayer. Unfortunately, in many instances, federal decision-makers do not have access to the data necessary to best inform decisions. In such instances, agencies are unable to show the benefits or impacts of the programs they administer and cannot determine what, if any, unintended consequences are created by programs, or whether programs can be improved.

Ensuring policymakers have access to high-quality administrative data is essential for evidence-based policymaking. Administrative data already exists and tested protocols are in place to facilitate merging of information and ensure confidentiality, yet agencies fail to share data with each other in ways that might improve program outcomes.

Further, Statutory restrictions often prevent agencies from sharing data with researchers who may be in a position to help the federal government identify needed solutions.

The first step in creating a culture of evidence-based policymaking is to determine what data is available and how to best get the data to policymakers. The Evidence-Based Policy Commission Act of 2015 will bring together leading researchers, program administrators, and experts to conduct a thorough study of existing infrastructure and statistical protocols. These

individuals will consider various methods of ensuring that policymakers have the access they need while balancing personal privacy and data integrity interests, and make recommendations

on how to best approach the issue of federal data access.”

[From House Report 114-211](#) Accompanying the Evidence-based Policymaking Commission Act of 2015

Congress acted on the Evidence Commission's report in an unusually short period of time. About one year after the report was issued, the bipartisan [Foundations for Evidence-Based Policymaking Act of 2018](#) (Evidence Act) was signed into law (Pub. L. 115-435). The law put

into statute 11 of the 22 Commission recommendations and mandated several structural improvements not covered by the Commission. These improvements included creating a Chief Data Officer for each agency and mandating that agencies create publicly available data inventories and establish data governance and management mechanisms. The law also required agencies to create evaluation plans to examine the effectiveness of their program outcomes, undertake an annual learning agenda to address big questions affecting public policy and program efficiency, and put in place an evaluation officer to oversee the work and build evaluation capacity.

The recommendations from the Evidence Commission and the legal requirements in the Evidence Act are relevant to state and local governments. Together, they may offer a model that could be adapted to build capacity to generate and use evidence at the state and local level across virtually every policy domain.

The National Data Infrastructure and Evidence Ecosystem as Resource for States

Capacity for evidence building is multifaceted. It ensures organizations have available staff, resources, infrastructure, leadership, and processes to align the need for information with the ability to produce relevant, timely, and objective insights. However, simply allocating increased funding or staff to a data unit may not result in the type of capacity needed to support the cultural shifts, technical analysis, or legal and privacy considerations relevant for modern data analytics. The Evidence Commission outlined capacity in a governmental context at the national level along the dimensions of people, process, and the legal framework.

Within the U.S. federal government, extensive capacity already exists to support data sharing, linkage, and use for a range of analytical purposes. While this capacity is imperfect and brings with it ongoing internal challenges, operational components of this capacity include the federal statistical system; the performance, evaluation, and policy analysis infrastructure; and coordinated privacy and cybersecurity efforts, discussed below in turn.

Federal Statistical System. The decentralized statistical system in the U.S. is composed of 13 major agencies and dozens of smaller statistical units. Many of the agencies share a common data use and protection law—the CIPSEA—and others have their own unique legal authorities such as the Census Bureau (Title 13 U.S. Code) and National Center for Education Statistics (20 U.S. Code § 9573). Although it is decentralized, the system is coordinated by the Chief Statistician of the United States at the White House OMB, who maintains a community of practice for leaders in the system through the ICSP. Each agency and unit have its own resource constraints and priorities, but there have been some attempts to recognize the system more comprehensively in budgetary and staffing arrangements as part of the annual President’s Budget request. The statistical system also has multiple arrangements in place for bilateral and multilateral data sharing to capitalize on the capabilities and unique expertise in respective agencies and units. Some of the coordination of the system includes the ongoing development

of a common proposal application for researchers to apply for access to restricted data as well as the use of the [Federal Statistical Research Data Centers](#). The research data centers are managed by the U.S. Census Bureau and include over 30 locations where nongovernmental researchers can access highly restricted, de-identified data from multiple agencies to conduct approved statistical research projects. The infrastructure for accessing the data is part of the Census Bureau's internal information technology (IT) system.

Performance, Evaluation, and Policy Analysis Infrastructure. Federal agencies such as the Departments of Housing and Urban Development, Labor, and Education long ago established robust policy analysis offices and units to support the formulation of regulatory and administrative actions, as well as program evaluation activities. In the mid-1990s, as required by the Government Performance and Results Act of 1993 (GPRA), federal agencies bolstered capabilities to measure program performance and outputs. Building on performance measurement capabilities, the Evidence Act requires large federal agencies to maintain the program evaluation function to study program outputs and outcomes. Taken together, the performance, evaluation, and policy analysis infrastructure provide a venue for using administrative records, partnerships with the federal statistical system, and direct coordination with state and local governments to produce actionable insights and recommendations for policy actions.

Privacy and Cybersecurity Coordination. Federal agencies are required to maintain key leadership roles for privacy protections, such as the Senior Agency Official for Privacy and the Chief Information Security Officer. Increased attention in recent years to the need to bolster the cybersecurity infrastructure to safeguard systems and data has led to a proliferation of resources and staff. Cybersecurity is only one component of strong privacy safeguards, which also include confidentiality protections. Confidentiality is a subset of privacy and involves managing the risks of deidentification, deploying disclosure avoidance, and maintaining pledges of confidentiality. OMB also plays a central role through its privacy office, the Chief Statistician of the United States, and the Chief Information Officer, namely, to collaboratively and comprehensively deploy government-wide privacy safeguards that ensure information is both protected from potentially harmful uses while encouraging beneficial, authorized uses.

All federal activities are notionally coordinated at a policy level by the White House OMB, which is statutorily charged with coordinating data collection and management, IT systems, performance and evaluation, information policy, and privacy policy (NAPA, 2020). Technical details and standards may then be established by individual agencies, voluntary interagency councils, or voluntarily adopted based on industry or consensus approaches. Such collective approaches leverage the capacity within program offices and units, where administrative activities may supplement protections and capabilities to link or use data.

States will want to establish certain aspects of this capacity to address their unique contexts and needs, discussed in the roadmap below.

Federal-State Data Partnerships

Federal agencies can serve as partners and resources in aligning and bolstering state and local capabilities nationally. Indeed, there are multiple examples where state partners recognize the critical role the federal statistical and administrative data capabilities play in supporting state capacity (Abazajian and Kleykamp, 2020). Below, examples are described of both long-standing and newer initiatives designed to provide value to all partners.

The NCS-X Initiative

The NCS-X Initiative is a partnership between the U.S. BJS and the FBI. The goal of NCS-X is to transition all law enforcement agencies from reporting data into the older **Uniform Crime Reporting System (UCRS)**, which has been providing aggregate monthly crime statistics since 1930, to reporting into the newer **NIBRS**. **NIBRS** provides more information than the UCRS, such as the ability to provide circumstances and context for crimes like location, time of day, and whether the incident was cleared. Participation is voluntary for both systems, which contain data from cities, academic institutions, and state, county, tribal, and federal law enforcement agencies. The FBI and BJS provide incentives for governments and law enforcement agencies to participate. NCS-X leverages the existing infrastructure of NIBRS and is expected to increase the ability to monitor, respond to, and prevent crime by supporting production of timely, detailed, and accurate national measures of crime incidents. The NCS-X Implementation Team includes representatives from multiple organizations that provide technical assistance to NCS-X sample agencies and state UCR Programs, ranging from technical assessments and consultation to providing marketing assistance and facilitating agency interactions across states. One example of the types of assistance provided is the [NCS-X Playbook](#) for law enforcement agencies.

State Justice Statistics Program

The State Justice Statistics Program administered by BJS through the Office of Justice Programs is a grant program that supports the establishment and operation of Statistical Analysis Centers (SACs) in the states and territories. The SACs collect, analyze, and report statistics on crime and justice to federal, state, and local government and share state-level information nationally.

The National Directory of New Hires

The National Directory of New Hires (NDNH) is a system maintained by the Department of HHS as a repository of quarterly earnings, UI, and new hires records. States provide data to HHS under federal law and can access insights about noncustodial parents, to support implementation of child support enforcement programs. Shared use of the data provides actionable and programmatic uses of information and supports research and evaluation activities that generate summary statistical insights. For example, the Department of Housing and Urban Development used NDNH data to analyze long-term outcomes in the Family Options

Study (HUD, 2016); the statistical uses complemented an administratively collected data asset, while providing valuable insights to inform federal and state policy (Fletcher, 2019).

The State Longitudinal Data Systems (SLDS)

The SLDS were established in 2005. SLDSs have demonstrated value to policymakers and educators in improving the ability to generate real-time, actionable information about student performance as well as capabilities to study long-term educational outcomes. Every state has an SLDS and can link basic educational data, and states are investing additional resources to connect education, workforce, and human services data to reduce barriers to service delivery (Garg et al., 2021). Data linked for administrative purposes produce tremendous benefits for research and evaluation activities, including the ability to study important outcomes and identify which educational strategies are most effective for groups of students. For example, the systems provide resources to state administrators and local practitioners for ensuring they can address major educational and workforce gaps, including learning disruptions from the global coronavirus pandemic. These systems are operated at the state-level; however, the major federal investment in development and operation places SLDS at the nexus of federal-state infrastructure. States receive substantial federal resources to develop education data capacity that generates benefits for both the state and federal governments.

The Vital Records System

The Vital Records System is operated and maintained by states, with substantial federal investment from the HHS NCHS and the Social Security Administration (SSA). Vital statistics are collected to provide insights about morbidity and death, to help government agencies prevent improper payments and monitor health issues that may be leading indicators. The Vital Statistics Cooperative Program—the oldest federal-state data sharing partnership—is a partnership between NCHS and 57 state and local jurisdictions; the Program includes detailed microdata about individual characteristics and causes of death (Rothwell, 2017). States provide data, and in exchange, NCHS provides state funding under contract, and training and technical assistance to standardize data quality. SSA receives a subset of the data from states, including the record of death and date of death for individuals. States are the data provider to the federal government, and the benefits include resources to fund and operate the vital records system as well as electronically coded records based on provided information.

The Health Information Technology for Economic and Clinical Health Act (HITECH)

The HITECH, enacted as part of the American Recovery and Reinvestment Act of 2009, authorizes funds to aid in the adoption and use of EHRs and linkages through Health Information Exchanges (HIE). The Centers for Medicare & Medicaid Services (CMS) and the HHS Office of the National Coordinator for Health Information Technology (ONC) have grant funds for states to close the gaps in interoperability, infrastructure, and other activities. Federal Medical Assistance Percentages funds at the 90/10 matching level support HIE activities such as EHR adoption, linking laboratory or other data sources for Medicaid eligible through HIE, and

supporting hardware and software EHR/HIE linkages at the provider site. CMS can provide funding for state administrative activities related to development of core HIE services (e.g., designing and developing a provider directory, privacy, and security applications, data warehouses), public health infrastructure, electronic Clinical Quality Measurement (eCQM) infrastructure, and provider on-boarding. CMS administration and enforcement of Health Information Portability and Accountability Act Administrative Simplification regulations also promotes interoperable data exchange by means of standards and operating rules. In addition, ONC has been advancing the Standards and Interoperability Framework, the State HIE Cooperative Agreement Program, the Direct Project, the Nationwide Health Information Network Exchange, and the ONC Health Information Technology Certification Program. States may use the 90/10 matching level HITECH administrative funding to update existing HIE infrastructure to align with ONC interoperability and security guidelines and to meet requirements for exchanging data with federal agencies.

QCEW

The QCEW is conducted by the BLS primarily from state UI programs, supplemented by two BLS surveys. It is a quarterly count of employment and wages reported by employers covering more than 95 percent of U.S. jobs available at the county, MSA, state, and national level, by [detailed industry](#). BLS pays the states for these data and works closely with state workforce agencies to review and enhance the QCEW data before they are released. This product has been produced by BLS for decades. The **Local Employment Dynamics Partnership**, a voluntary federal-state partnership started in 1999, built on these efforts. States have agreed to share historical and ongoing administrative records of UI earnings data and QCEW data with the Census Bureau. The Census Bureau then produces a longitudinal data infrastructure from which new statistics can be produced about the dynamics of local employment and the locations of jobs and workers.

Census-ERS–FNS Joint Project

The Census-ERS–FNS Joint Project is a long-term joint research project to acquire administrative data on U.S. Department of Agriculture (USDA) food assistance programs—the SNAP and the Special Supplemental Nutrition Program for WIC—from states and link the data to Census Bureau surveys. The linked data provide insights on how program participation affects participants and who is eligible for assistance but does not participate and why. State SNAP and WIC agencies in participating states send their confidential microdata to the Census Bureau in exchange for state-specific analyses and reports. ERS researchers participate in joint research projects. Census has been able to link Veterans Administration and DOD data to gain insights into veterans decision-making. The project overcame several barriers to share records between two federal agencies and multiple states.

TANF Data Collaborative Pilot Initiative

The TANF Data Collaborative Pilot initiative sponsored by the HHS ACF is designed to build TANF state agency capacity to improve program performance through applied data analytics. The initiative includes eight pilot sites supported for 30 months with funding, intensive training, and technical assistance to support state and local efforts and build strategic partnerships. The selected states were California, Colorado, Michigan, Minnesota, New Jersey, New York, Utah, and Virginia. ACF's outside partners include the MDRC Center for Data Insights, Actionable Intelligence for Social Policy at the University of Pennsylvania, the Applied Data Analytics program of the Coleridge Initiative, and Chapin Hall at the University of Chicago. The Fall 2019 data analytics training program included TANF receipt data and QCEW data for both Illinois and Indiana. The program examined questions such as what characteristics increase an individual's risk of returning to TANF, what factors increase an individual's likelihood of not finding stable employment after leaving TANF, and what factors increase an individual's likelihood of not finding any employment after leaving TANF.

The examples above are just a few of many federal systems, or federally supported programs, operated in partnership with or by states. Each highlights the critical role that intergovernmental cooperation has in the national data infrastructure and evidence ecosystem. In addition, the examples present different models of how data are collected, managed, and used to support insights relevant for administrative decision-making, policymaking, and societal benefit.

Federal Data Modernization Efforts Relevant to States

In recent years, several federal agencies pursued data modernization efforts that were intended to be cross-cutting, with resources and capacity that benefited state and local governments.

Medicaid, Nutrition, and Human Services Interoperability Modernization

Historically, human services, health, and nutrition assistance programs in most state and local governments had distinct administrative systems that could affect coordination when individuals applied for benefits; in addition, the lack of coordination affected the ability to analyze data across programs. In 2011, HHS and USDA announced a [major shift in policy](#) to encourage states to modernize systems and to facilitate interoperability across federally funded health, human services, and nutrition programs. The policy at the time was a waiver to federal rules requiring a certain match on technology and data investments, meaning that states could leverage the federal government paying for 90 percent of upgrades outside of Medicaid programs. Federal waiver of the rules for certain matches as an exception to federal government-wide guidance (i.e., OMB Circular A-87) gave states an incentive to upgrade systems for more efficient eligibility and enrollment systems expected to realized benefits for program administration and research activities. In practice, the waivers led states to better integrate Medicaid, SNAP, TANF, and other human services systems.

Waiver authority was available for states over a period of more than five years, meaning the funding was largely predictable and sustainable during that period. The waiver supports successful efforts to avoid duplication across traditional program boundaries and to improve access to benefits. Examples of [successful activities](#) include the use of electronic data matching to verify eligibility, conversion of paper-based systems to electronic records, the use of data management and analytics capabilities, and funding for state data hubs. Waiver of the federal rules allowed for blending of funding across the traditional program boundaries to support major systems upgrades and modernization, which in turn resulted in operational efficiencies and demonstrated benefits for the programs. Coordinated guidance across agencies and traditional program silos can create incentives for redesign of state systems.

Public Health Interoperability Modernization

In 2020, the Coronavirus Aid, Relief, and Economic Security Act appropriated \$500 million for [public health data modernization](#) through the U.S. Centers for Disease Control and Prevention (CDC). The investments were intended to support federal, state, and local governments in strengthening the data collection, reporting, and analytical infrastructure for health surveillance and monitoring of diseases across the country. The focus has been designed around coordinating systems, modernizing the infrastructure to enable more rapid data collection, and building better partnerships and analytical capabilities. One major theme in the modernization effort is encouraging interoperability of systems for a seamless exchange of data. Coordination across systems should enable CDC to provide real-time information to the public, effective evaluation of health outcomes, and information for decision-makers in setting public health policy. Modernization also includes efforts to improve public health data standards, such as a strategy to reduce reporting burden from hospitals and states and more automated reporting for systems that involve lab results. In less than a year, the CDC is realizing progress, with new datasets published as open data, pilot projects to build pandemic-ready infrastructure to support states, and reduced time to apply data for contact tracing and other analytics.

The CDC data modernization efforts directly benefit the NCHS, specifically around vital records. Prior to modernization the vital records system was disconnected, error-prone, and sometimes duplicative. Under the CDC's modernization investments, the implementation of greater consistency in processes and data standards is facilitating bi-directional data sharing where both the states and federal government benefit from high-quality data. The vital statistics system and electronic case reporting receive 15 percent of total funding, with plans to allocate across all reporting jurisdictions to increase interoperability in 2021 and 2022. Similarly, the application of data science tools at NCHS is accelerating CDC's ability to produce relevant and timely health statistics. Other planned activities for NCHS include establishing a virtual data enclave to facilitate simpler, secure access to restricted data, for researchers to partner with NCHS in rapidly addressing major public health questions.

Data Modernization in the American Rescue Plan

In March 2021, President Joe Biden signed the American Rescue Plan (ARP), a \$350 billion proposal the Administration presented to Congress to support economic recovery in the wake of the global pandemic. Throughout the Administration's engagement in developing and advocating for the proposal, there were strong signals about the need for state and local governments to apply data and technology to effectively implement ARP investments. In May 2021, the U.S. Department of the Treasury issued guidance to state, local, and tribal governments requiring certain data reporting about programmatic activities, specifically linked to certain requirements and expectations in the Evidence Act. For example, the guidance and corresponding federal regulation explicitly authorizes states to allocate funding received under ARP "to make improvements to data or technology infrastructure and data analytics, as well as program evaluations" (U.S. Department of Treasury, 2021). It is too early to say how states in general are using this funding flexibility, in coming months this will become clearer because the federal government requires a reporting category for data and evidence investments.

Some agencies also have announced targeted resources to support data infrastructure in states, in addition to the general ARP guidance from the Treasury Department. For example, the Economic Development Administration's Good Jobs Challenge includes \$500 million to support data and capacity that leads to scaling evidence-based interventions and for conducting evaluations (Economic Development Administration, 2021). In addition, given the guidance on implementation from the Treasury Department, additional investment opportunities for states will likely be forthcoming.

Building State Capacity for Evidence Building

At the same time that federal initiatives have been underway in recent years, many states have built robust data capabilities, both within the state and in collaboration with other like-minded states. Both state and local activities have focused on enhancing capacity for evidence building and informing operational decision-making in real-time across many topical areas. Some efforts have increased transparency for the public about how government operates. Many states have realized that fragmented and siloed operational or programmatic structures can inhibit data sharing, linkage, and use. Even without legal barriers to sharing, organizational culture, scarce resources, and bureaucratic red tape can create frustrating roadblocks that can seem insurmountable. Yet some states have succeeded in overcoming these roadblocks. This section offers examples of pathways and foundational activities that can help states succeed.

One valuable resource is the annual [Invest in What Works State Standard of Excellence](#) published by Results for America, a national non-profit that works to promote data and evidence use. This voluntary self-assessment or framework reviews basic capacity characteristics across participating states, using a standard that includes explicit criteria about data governance and leadership, data policies, data infrastructure and data use (RFA, 2021). The framework

presented in the 2020 assessment identified 169 examples of promising systems and programs across 35 participating states (RFA, 2020).

We expand on the RFA framework to include other key aspects of capacity that are relevant at any level of government based on successful models, such as transparency, legal framework, and incentives for data sharing. Our approach is intended to reflect the collaborative nature of intergovernmental cooperation on data infrastructure, including the joint needs of states and federal agencies in using state-collected or managed data. Together, the key characteristics to weigh and consider in building, maintaining, and adapting state capacity are briefly described below.

Key Considerations for Increasing State Capacity to Build Effective Data Infrastructure

- **Leadership and Strategy:** Have priority areas been identified by decision-makers to include in high-value data sharing and analytics projects? Have senior leaders expressed support for improving data infrastructure and have they identified, resourced, and empowered a responsible senior-level individual (e.g., chief data officer) for the state? Are the governance approach and planned action items articulated in a publicly accessible data strategy?
 - **Legal and Regulatory Authority and Policies:** Have experts reviewed appropriate state legal authorities to ensure sufficient capabilities exist in law to enable data collection, sharing, analysis, protection, and responsible use? Does the state have policies that outline expectations to use data for informing decisions, sharing data across organizations when appropriate, and safeguarding sensitive information?
 - **Governance:** Does the state have effective mechanisms to govern data practices and quality, including developing a data inventory and promoting accessibility of information? Are data sharing practices supported with common memorandum of understanding and with procedures generalizable to state agencies?
 - **Transparency of Processes and Uses:** Are mechanisms available to communicate openly with the public and stakeholders about data practices and uses, including the value and benefits of using data?
 - **Accessibility of Data:** Are procedures in place to enable applications and efficient approvals to qualified users, for access to government-managed data? Are open data assets and de-identified data available when possible, for public access?
 - **Oversight and Accountability:** Do external stakeholders routinely participate in processes that provide a diverse range of perspectives about data governance and use? Are legislative oversight procedures clear? Have routine processes been established to audit or periodically review state data infrastructure and practices?
 - **Sustainability:** Are resources available to provide for sustained systems, procedures, and personnel with expertise to collect, manage, protect, and use data? Is there a mechanism for training employees for ongoing collaborative work with their data?
-

Data Leadership and Strategy

Data leadership is key to advancing more effective use of data. While many states have created the position of a Chief Data Officer, leadership can come from multiple sources. Often leadership arises from agency heads who understand that they need better data to improve their programs and the programmatic outcomes for the public. These leaders can have substantial influence and garner support from their state legislature and governor's office when the value of having better information is made clear. Value has often been clearest when focused on topics, such as tackling unemployment or reducing crime.

Data leadership requires having a strategic vision. At the federal level, a key driver of change has been the [Federal Data Strategy](#), developed by White House OMB officials including the Deputy Director for Management, the Chief Statistician of the U.S. and the U.S. Chief Information Officer. The work of developing the strategy was accomplished by a team of representatives from agencies across government and with important input from the public, gathered through roundtables, listening sessions, and solicitations for public comment. The Federal Data Strategy provides a government-wide framework that includes principles, practices, and an annual action plan. As stated on its website, "*The mission of the Federal Data Strategy is to fully leverage the value of federal data for mission, service, and the public good by guiding the Federal Government in practicing ethical governance, conscious design, and a learning culture.*"

The Federal Data Strategy is a model for how top-down leadership can drive change across government. Strategy development was incorporated into the President's Management Agenda. This agenda is taken up by the President's Management Council, comprising the Deputy Secretaries of the Cabinet agencies. Quarterly progress reports for agenda items are published on a website and delivered to the Council, which is chaired by the Deputy Director for Management at OMB. The cross-agency effort has high-priority and visibility. Deadlines were established, and progress was reported weekly. Funding was made available for the effort. The strategy was developed in about 18 months, including the first year action plan.

The action plan itself was organized into three categories: 1) actions taken by all agencies; 2) actions taken by the community of practice; and 3) shared solutions actions undertaken by one agency on behalf of other agencies.

The excerpt from the Federal Data Strategy below shows the types of actions taken by every agency and highlights many of the data standards of excellence discussed here. For example, agencies are required to set up a data governance body, assess the maturity of their infrastructure, and publish data inventories. Additionally, a community of practice was required to establish a cross-agency Chief Data Officer Council and focus on improvements in key areas such as financial management and geospatial data. Shared solutions included in the action plan include developing a data ethics framework, creating a toolkit for agencies on best methods for protecting privacy and confidentiality of sensitive data, and establishing a central data policy

coordinating committee. Each of the items in the action plan is supported by either a statute, an executive order, or official OMB guidance.

Agency Actions			
	Foundations for Evidence-Based Policymaking Act of 2018 and Associated OMB Guidance	Executive Order on Maintaining American Leadership in Artificial Intelligence	Improving Implementation of the Information Quality Act (M-19-15)
1. Identify Data Needs to Answer Priority Agency Questions	✓		✓
2. Constitute a Diverse Data Governance Body	✓		✓
3. Assess Data and Related Infrastructure Maturity			✓
4. Identify Opportunities to Increase Staff Data Skills	✓		
5. Identify Priority Data Assets for Agency Open Data Plans	✓	✓	✓
6. Publish and Update Data Inventories	✓		

Federal Data Strategy, First Year Action Plan, www.strategy.data.gov

For states, leadership from the top can help create momentum and motivation remove barriers to data sharing and collaboration. Having a governor and state legislature drive innovation and change can free up resources and help cross-agency coordination. An executive order from the governor or enacted legislation can be helpful in focusing attention on high-priority policy areas and garnering resources to pay for people, training, and infrastructure.

One model of this top-down approach is Ohio. In 2019, the Ohio governor consolidated state data systems under an [executive order](#), establishing the **InnovateOhio Platform**. Through senior-level leadership provided by the Lieutenant Governor’s office, the platform was established as a shared resource, to link datasets intentionally across programs. The executive order established certain characteristics, many identified below, such as an expectation for data sharing across agencies unless prohibited in law. The Lieutenant Governor takes an active role in implementing the infrastructure, including serving as its director and [supporting public messaging](#) about the value of the system, the efficiencies achieved, and the benefits constituents directly receive from better data sharing.

Another model involves leadership shared across states, driven by agency leaders and engaged partners. One of the most successful examples of multistate collaboration is the [Midwest Collaborative for Evidence-Based Policy Making: Transitions in Education and Workforce](#). [The Collaborative focuses on](#) development of core data analytics projects and competencies for a group of Midwestern state workforce and education agencies, working with academic partners, that began convening in 2018. Several workshops and training sessions for state agency employees were held to identify a roadmap and strategy that would work for multiple states. The need was clear: members of the public were crossing state lines to live, work, and go to school, and to understand effective workforce and education strategies, states needed to share data with each other. The states used a secure shared cloud platform supplied

by a non-profit partner, the Coleridge Initiative, which also provided training to state employees on how to work with their shared data. The platform and training allowed the collaborative work to move forward more quickly, without having to coordinate multistate infrastructure investments. Hands-on work with the data also inspired new ideas about how to improve data quality and interoperability. One important aspect of the Collaborative is that they have implemented tiered access to protect privacy and confidentiality of sensitive data. The Governors and their staff, as well as the workforce boards, did not want synthetic or artificial data, so Coleridge produced confidential summary tabs of data that are password protected. Data have been supplied by the Missouri Departments of Employment Security, Labor, Corrections, and Higher Education; the Indiana Departments of Workforce Development and Higher Education; the Ohio Longitudinal Data Archive; and the Illinois Department of Employment Security. Kentucky, Tennessee, and Michigan are planning future training and work sessions. Funders have included several philanthropic foundations. A few of the topics addressed through the workshops and training sessions have included the earnings and employment outcomes of credentials and degrees, remedial skills training, vocational training, or apprenticeship training; how well postsecondary student majors match to in-demand occupations; and earnings across demographic groups and what factors explain observed differences.

Legal and Regulatory Authority and Policies

Laws and regulations provide formal mechanisms for agencies to promote data access and use. The federal Evidence Act, for example, includes the requirement for OMB to issue regulations to implement several provisions, including the expanded sharing of data across agencies for evidence building activities. Many states implement federal laws through cooperative federalism, in addition to state-specific authorities; for this reason, state request for improved clarity and consistency in interpreting federal laws and regulations are common. Federal authorities offer examples of processes that can be leveraged at the state-level.

Similarly, important elements were incorporated in the Ohio executive order referenced in section 3.1 above. The Ohio executive order cited a section of Ohio state code that had been revised to allow data sharing between agencies ([Section 125.32, Ohio Revised Code](#)). This change in the law enabled “...an enterprise data management and analytics program to gather, combine, and analyze data provided by one or more agencies to measure the outcome of state-funded programs, develop policies to promote the effective, efficient, and best use of state resources, and to identify, prevent, or eliminate the fraudulent use of state funds, state resources, or state programs. Participating state agencies may use data gathered under the program for these purposes.” State agencies were required to provide data for use under the program. Ohio’s approach is similar to that of the Evidence Act, which has a presumption of accessibility to data for statistical agencies.

Presumption of Accessibility to Data Assets

“ACCESSIBILITY OF DATA ASSETS.—The head of an agency shall, to the extent practicable, make any data asset maintained by the agency available, upon request, to any statistical agency or unit for purposes of developing evidence. “(b) LIMITATIONS.—Subsection (a) does not apply to any data asset that is subject to a statute that— “(1) prohibits the sharing or intended use of such asset in a manner as to leave no discretion on the issue; or “(2) if enacted after the date of the enactment of this section, specifically cites to this paragraph. “(c) REGULATIONS.—The Director shall prescribe regulations for agencies to carry out this section. Such regulations shall— “(1) require the timely provision of data assets under subsection (a); “(2) provide a list of statutes that exempt agencies from the requirement under subsection (a) pursuant to subsection (b)(1); “(3) establish clear and consistent standards, to the extent possible, for complying with section 552a of title 5 (commonly known as the ‘Privacy Act of 1974’) and any other applicable law requiring the protection and confidentiality of individually identifiable information; and “(4) require a transparent process for statistical agencies and units to request data assets from agencies and for agencies to respond to such requests.”

Foundations for Evidence-Based Policymaking Act (Pub.L 115-435) Title III, Part D, Section 3581

Ohio passed legislation enabling data sharing, reinforced by the governor through creation of a specific platform through which agencies must share their data. The federal Evidence Act mandated all agencies to share data only with the principal statistical agencies, but other agencies with programmatic data are not prohibited from sharing data with each other, to conduct statistical activities and conduct evidence building and evaluations. Federal data sharing remains decentralized, but the Evidence Act establishes an advisory committee to determine whether a centralized national data service should be established to combine sensitive data for statistical analytical purposes (including program evaluation).

Similarly, the governor of Indiana issued an [Executive Order](#) that mandated data sharing and established the governor’s Management and Performance Hub. The hub is discussed in more detail in section 3.4 below. Both the Ohio and Indiana initiatives, driven by Executive Order and statute, have been highly successful.

One important element of success for the Midwest Collaborative has been the use of common agreements and a clear legal framework. Having these in place is critical for sharing data across agencies and states. Further, bringing in other important state officials such as a Chief Privacy Officer or Chief IT Security Official can move forward a statewide effort that may involve multiple regulations around different types of data.

The Evidence Act mandates that OMB promulgate regulations and guidance around the data sharing enabled by the law. These regulations include establishing standards for agencies to: 1) categorize the sensitivity of their data and assign appropriate levels of accessibility; 2) determine whether a less sensitive data set could be produced that would increase accessibility; 3) conduct risk assessments for a data asset prior to release; and 4) be transparent about their processes and make the information easy to understand and publicly available. Although these requirements would apply to statistical agencies, they are generalizable to any agencies that make potentially sensitive data available to the public or to researchers.

Data Governance

Data governance may be defined as “the specification of decision rights and an accountability framework to ensure the appropriate behavior in the valuation, creation, consumption and control of data and analytics” (Gartner, 2021). Put another way, data governance defines the roles and responsibilities of all participants involved in the collection, processing, use, and dissemination of an entity’s data and includes establishing policies, procedures, metrics, and accountability mechanisms.

The federal Evidence Act creates the position of Chief Data Officer (CDO) for each agency and specifies that this is to be a nonpolitical appointment. The CDO is required to be selected on the basis of qualifications such as “...*demonstrated training and experience in data management, governance (including creation, application, and maintenance of data standards), collection, analysis, protection, use, and dissemination, including with respect to any statistical and related techniques to protect and de-identify confidential data.*”³³

Establishing and implementing data governance policies can be challenging in a government environment, where ownership of the data can be spread among multiple agencies. Only about 50 percent of states have appointed CDOs or their equivalent. Several states (e.g., Arkansas, Connecticut, and Texas) have adopted the idea of Agency Data Officers to support governance across departments. By establishing a CDO in each agency and assigning responsibility to that person for data governance, the states align with the Evidence Act in identifying data governance as important to using data to achieve better outcomes. The Evidence Act establishes a CDO Council composed of all the agency CDOs. The function of the Council is to “...(1) *establish Governmentwide best practices for the use, protection, dissemination, and generation of data; (2) promote and encourage data sharing agreements between agencies;*(3) *identify ways in which agencies can improve upon the production of evidence for use in policymaking; (4) consult with the public and engage with private users of Government data and other stakeholders on how to improve access to data assets of the Federal Government; and (5) identify and evaluate new technology solutions for improving the collection and use of data.*”³⁴

³³ Pub.L 115-435 Title II Section 3520

³⁴ Ibid Section 3520A

Likewise, at the state-level, governance boards or steering committees are needed, to allow for sharing resources and knowledge in addition to data.

At the federal level, some of the most significant barriers to data sharing have yet to be addressed, but some states have moved forward. For example, putting in place a common agreement or Memorandum of Agreement for data sharing can save years of effort compared with the process of each agency negotiating unique agreements each time data are to be exchanged or shared. In addition, having a common application process for requesting permission to access sensitive data for specific projects can streamline the approval process considerably. There is no common federal Memorandum of Agreement mandated for agency use. Every agreement is negotiated separately, and the process of securing an agreement can be even more complicated if multiple agencies are involved. In 2014, OMB offered a complete model agreement for a standard Memorandum of Agreement for data sharing in guidance to agencies ([M-14-06](#)). However, adoption was not mandatory, and agency legal departments continued to craft unique agreements around each project.

The Commission on Evidence-Based Policymaking identified another challenge, namely, that, “Cumbersome and idiosyncratic data access procedures create confusion, impose unnecessary costs, and are a barrier to evidence building, without always providing significant privacy benefits.”³⁵ Many federal agencies provide direct access to researchers and have developed unique application and approval processes to grant that access. Such an approach becomes complicated and time consuming if researchers want to access data from multiple agencies. The Evidence Act required OMB to set up a single portal for applications for data from statistical agencies; however, this provision did not address the need to streamline the business process for approvals once an application is received.

States have established data governance structures to address some of these challenges. For example, North Carolina has used centralized data governance to increase productivity and effectiveness through the [North Carolina Government Data Analytics Center \(GDAC\)](#), which is a central organizational structure for comprehensive data management. As part of the state Department of IT, GDAC services are closely aligned with the Chief Information Officer (CIO) for systems modifications and analytical needs. GDAC’s approach facilitates collaboration with state agencies to define priorities, obtain access to relevant data assets, and provide for integration of analytical capabilities. For example, the [Criminal Justice Law Enforcement Data Services](#) support offered by GDAC facilitates an integrated criminal justice data system. GDAC also collaborates across agencies for operational activities related to compliance and fraud, and coordinates the state integrated data system for education and workforce data. Finally, GDAC’s authorization in [state law](#) led to the creation of a data governance program that includes the use

³⁵ U.S. Commission on Evidence-Based Policymaking (CEP). *The Promise of Evidence-Based Policymaking: Final Report of the Commission on Evidence-Based Policymaking*. Washington, D.C.: GPO, 2017. p35

of a standardized Memorandum of Agreement for sharing data across governmental jurisdictions, with strong privacy protections.

Similarly, the [Kentucky Center for Statistics](#) (KYSTATS) links data from across 15 state agencies to conduct analysis on education and workforce programs. KYSTATS is transparent about their capabilities, which include a [common portal for requesting data access](#), data use restrictions, a common data sharing agreement, and a data dictionary. Each of these practices streamline the efforts from data requestors—often researchers—to learn what data are available for access and use and how to request access; KYSTATS reflects comprehensive data governance activities. Further, in 2019, Kentucky began streamlining its data sharing across agencies in the [Enterprise Data Management Policy](#), which is implemented in conjunction with the state [Enterprise Privacy Policy](#).

Arizona's Department of Administration also developed and implemented an enterprise [memorandum of understanding](#) for data sharing, now used by dozens of state agencies. The approach aligns with Arizona's statewide [policy for interoperability on government data](#) and a policy that outlines requirements and expectations for data governance in the state.

Transparency of Processes and Uses

Transparency is key to establishing and maintaining public trust in how the government uses data. There are several aspects to transparency to incorporate into building capability and infrastructure. However, transparency starts with the assumption that the public has a right to know how information is collected, used, and shared and how society benefits from these activities.

Statistics Canada, the centralized national statistics office of Canada provides a model for data transparency. Statistics Canada maximizes the use of administrative program data in producing statistics that supplement its surveys and censuses. The agency has created a set of [principles](#) with strong legal, policy, and organizational safeguards. These principles include strong privacy protections and consider the effects on privacy related to redirecting program data collected on individuals to new purposes, particularly when combined with other data. In addition, Statistics Canada has pre-approved certain types of data linkages with low risk of privacy violations. Further, Statistics Canada provides a summary of all linkages on its website.

Many states do not have open data portals. However, the COVID pandemic has demonstrated the value of providing important data to the public and to researchers. One approach is demonstrated by [Mississippi LifeTracks](#), a state longitudinal data system that enables analysis of administrative data from multiple state agencies. One primary purpose of LifeTracks is to assess education and workforce outcomes. The system received substantial initial funding through federal government investments from the Department of Education and receives annual state appropriations. When establishing LifeTracks, Mississippi engaged with stakeholders to prioritize transparency in the design and uses of the infrastructure, for example, transparently

sharing information on its public website, to promote and increase accountability. The LifeTracks website includes reports and projects completed from the system's data, demonstrating value while sharing relevant information.

The Indiana [Management and Performance Hub](#) is another example of a useful portal that provides transparency on the workings of state government and access to data. It includes the following features: a data hub with more than 160 secure and de-identified datasets of actionable data covering subject matter areas within state government; a searchable data catalog with definitions to available fields from datasets on the Indiana Data Hub; access to a secure collaboration environment that expedites research and analysis by bringing research teams, code, and data together; and a statewide data proficiency program for state employees.

Similarly, Maryland enacted the [Open Data Act](#) (State Government, Section 10-1501 through 10-1504) and as a result, established an [Open Data Council](#) with 37 members, including Cabinet agencies, five county representatives, and five private citizens. The Council Chair is the Secretary of the Department of IT, with the Director of the Governor's Office of Performance Improvement serving as the Vice-Chair. Maryland's [Open Data Portal](#) offers access to several useful data sets about demographics and business/economic activities in the state, government data to provide transparency of government operations, the ability to provide feedback, and mapping information.

Accessibility for Data Use

Accessibility encompasses many functions and activities. It may mean providing information to potential data users through data inventories, clear instructions on how to request data, and data management policies that emphasize data quality and interoperability. In addition, states may collaborate to eliminate traditional barriers to access. The examples from Mississippi, Indiana, and Maryland above offer lessons on accessibility. Below, we highlight other state examples to enhance access to data for both the public and state employees.

COVID UI dashboards, initially created by the state of Illinois and the same non-profit organization supporting the Midwest Collaborative, are now being used by nine states in the Collaborative. The dashboards combine data from multiple sources, including up-to-date UI claims (initial and continuing) with crosstabs summarizing the education levels, age, race and sex of unemployed workers by county and industry, how much of their earnings have been lost, and estimates about how quickly their benefits are being used. The crosstabs can be linked to state TANF and SNAP records to determine the impact on low-income families. A philanthropy has provided funding to expand the existing infrastructure to enable the states and researchers to access, use, and build tools to inform decision-making about the loss of jobs in the pandemic; these tools may then be used at the state and sub-state-level across the country. The project will develop a standardized approach to employment and benefit use histories, industry characteristics, and outcome measures. Models are being developed on the impact of different sector-specific or educational interventions on getting successful jobs. Models forecast the

impact of the pandemic on individual and family participation in state income transfer, nutrition, and social service programs.

These innovations would not have been possible without a focus on increasing accessibility to key data. The Evidence Act recognized the need to improve accessibility by including provisions to encourage data sharing within a strong privacy framework. One particular provision, for example, was based on the Evidence Commission's recommendation that the federal government should change the default to encourage data sharing in conjunction with privacy safeguards, to realize the strategic value of data collected by the government when possible.

Oversight and Accountability

Approaches for oversight and accountability are expected to vary by state, reflecting political priorities and structures and aligning with expectations from the state population and relevant stakeholders. Multiple mechanisms exist for enabling and encouraging oversight activities that promote accountability for the public from government's use and management of data. In general, such processes should include explicit participation, advice, and feedback from a diverse range of stakeholders related to all aspects of the data pipeline, including data collection, data protection, and data use. When Congress passed the federal Evidence Act, it included an [Advisory Committee on Data for Evidence Building](#) as one construct to proactively support new strategies moving forward and to provide strategic advice and consultation from experts about implementing key data provisions in the Evidence Act. The federal government also has established processes for publishing public notices about data collection requests required by the PRA and permissible uses under the Privacy Act of 1974.

Many states have established advisory boards and committees to support feedback on data priorities. In Kentucky, the Education and Workforce Statistics Board provides oversight and direction for the state longitudinal data system with integrated education and workforce data. The Board was established in [state law](#) with an advisory board to provide direction and oversee compliance with expectations, use limitations, and privacy protections.

Other approaches for oversight could include periodic annual reports to the legislature and public or an independent evaluation or audit process.

Sustainability

A secure and stable funding source for state data activities is typically necessary to ensure continuity and the capability to meet future demands on the infrastructure. Because data activities may span multiple years, predictability in funding is an important characteristic for sustained capacity. Planning for and building capacity as an ad hoc exercise can be costly and inefficient. Instead, states should consider how resources and funding are provided, to enable a sustained presence for key priorities. A 2018 review of state data capacity noted that state

officials cited staffing and expertise as the single greatest challenge for using data in the state (Pew, 2018). Over a quarter of the respondents also identified funding as a major limitation.

In the District of Columbia, the capacity to organize and use data was bolstered through the creation of the [Lab at DC](#). Funded initially through a philanthropic contribution from Arnold Ventures, the Lab was eventually supported primarily through a sustained investment from DC government, following successful completion of multiple projects that demonstrated the value of the endeavor. While the Lab did not address all issues or data capacity in the District of Columbia, it did support capacity-building efforts and achieved a sustainable funding model to enable staff and expertise to support a range of priority projects.

There are multiple approaches for developing sustained capacity and resources to implement state data activities. As discussed above, funding may be available from federal agencies and programs to support a broad range of initial infrastructure and development activities. In addition, states may choose to identify philanthropic contributions for one-time or targeted priorities and activities that build capacity. However, sustained investment will likely require direct state appropriations, which could be supplemented by user fees for access to restricted data, for example.

One approach for enabling sustained resources may involve setting aside a portion of program implementation funds to ensure data infrastructure is adequately resourced to support program implementation. The Evidence Commission recommended this approach at the federal level, and the funding mechanism is equally relevant for state data capacity (CEP, 2017; Fatherree and Hart, 2019). Another successful approach is a shared resource model, where infrastructure and data, as well as training, are shared across government.

Finally, resources should be used to support individual- and organizational-level training on an ongoing basis. Staff at all levels of an organization should be exposed to resources for data literacy and use, including senior leaders and managers. Staff working with data collection or curation should also receive frequent updates and training to ensure the latest privacy and cybersecurity protections and practice are deployed.

A Roadmap for State Data Capacity and Infrastructure

Every state and local government has a role to play in supporting evidence-based policymaking in its own context, to improve services and operations for local populations. For a state government seeking to build increased capacity and infrastructure to use data across agencies and programs, there are foundational building blocks. These building blocks can support state officials in implementing evidence-informed decision-making, to understand outcomes and to better describe characteristics or the economy, people, and programs in a state. Building this capacity and infrastructure is challenging, can be resource intensive, and always requires clear leadership.

Building blocks for the success of state data initiatives include the key characteristics of capacity outlined in section 3 above. These characteristics can be developed in tandem, though some may need to occur before others, based on specific state attributes, needs, and constituencies. The roadmap presented below for state capacity suggests approaches that can work well, though states should expect to modify steps based on their own situations. Roadmap steps may not be linear or sequential; many may require repetition or may be mutually inclusive to be developed concurrently with other steps.

As roadmap steps are considered in a state, a foundation should be established and expanded over time, with a clear project or focus in mind. Establishing priorities and areas of immediate attention will help build success stories and momentum for further success. For this reason, planning from the outset to scale approaches may prove fruitful for many states.

The key characteristics and examples identified in this white paper identify several clear opportunities for states to engage in evidence-based policymaking while building additional capacity and leveraging both national and peer capabilities. Roadmap steps reflect considerations in the federal Evidence Act along with state-specific considerations, as follows:

- 1. Establish and empower state and local data leadership roles.** Anyone can be a leader when it comes to improving data quality and access; however, recognized senior-level leaders within states and agencies signal the effort as a priority to colleagues and partners. Leaders are instrumental in securing resources, personnel, and space on the policy agenda for improvements. States can establish a recognized CDO, ideally recognized, resourced, and empowered by elected officials and civil servants alike. In virtually every successful data initiative, a clear leader emerged to champion implementation of the project and to align interests, resources, and expertise to drive change. A recognized leader such as a CDO can also be instrumental in fostering new partnerships, for example, with industry, nonprofits, and research institutions in a state. Designated data leaders can coordinate with other C-suite officials such as the CIO to support effective implementation of systems necessary for robust data analysis and data governance. Initial leadership can come from agency heads who have a clear, immediate need for better information.
- 2. Identify the Local Knowledge Needs and Data Gaps Based on Priorities.** All states will have limitations in resources and must prioritize how to implement data initiatives for modernization and capacity-building. There are multiple planning processes relevant to address priority data issues. A basic consideration is how to best identify where informational needs exist and how to devise a plan to address these needs. One approach is to develop a learning agenda, or a strategic plan for research and evaluation that asks first about the questions to be addressed, then identifies existing assets and areas where new data collection or data sharing may be necessary (Newcomer et al. 2021). To the extent possible, aligning such efforts with strategic planning or other administrative priorities will encourage leadership and development of resources about the selected topics.
- 3. Review Legal Authorities to Share, Protect, and Use Data.** Identifying and understanding existing legal authorities in states is critical to determining what, if any, legal modifications

may be necessary to share and protect data while planning for uses. Such a review should determine whether data sharing is considered the default and whether unintentional prohibitions or limits exist on data sharing or use. Once a data champion has been identified and priorities established, conducting a legal review may be assigned to agencies, the CDO, a research institution, or an independent commission of experts. Reviews of legal authorities should be shared with relevant executive and legislative actors. This will encourage resolution of provisions that may unintentionally inhibit data sharing and areas that may need stronger privacy protections. Further, sharing such reviews may encourage recognition of how existing authorities may be used to simultaneously support program implementation and bolster data infrastructure or capacity.

4. **Provide Feedback to Federal Partners about Data Needs and Priorities.** States operate numerous programs that are federally funded or mandated. For this reason, states should be encouraged to actively share insights and perspectives about data needs and gaps, especially where there may be a role for the federal government to support states with resources, technical assistance, training, or interstate data coordination with standards and analysis. For example, states could participate in the federal agency learning agenda process required by the federal Evidence Act. One aspect of learning agenda feedback may be requesting support from federal partners to coordinate where there are data gaps across states. Another area might be making requests and clarifying areas where blending funding across multiple federal funding sources could help states develop and operate increasingly integrated systems to improve operations and research. States may encounter areas where federal law, regulation, or guidance impedes state-level priorities, sometimes unintentionally. In addition, state should be encouraged to use the formal learning agenda process to recognize where improved guidance from federal agencies could support state efforts.
5. **Establish a cross-agency governance process and data strategy.** State-level data governance encounters issues unique to each individual state, yet the presence of a clear, consistent, and coordinated governance process provides coherence to data activities for civil servants and partners. States can ensure that the capability to coordinate among senior officials in key agencies occurs frequently. Some states may choose to set up committees of data officials across agencies while others may choose a more centralized approach within a statistics unit or administrative office. Governance processes designed in different ways may each be paired with development of a data strategy to gather stakeholder input and feedback. The Federal Data Strategy received extensive feedback, including from industry and researchers; state-level strategies would likely receive similar attention.
6. **Execute on Data Governance by Integrating Local Data Where Feasible to Prepare for Linkage and Sharing.** Government agencies and institutions should know what information they have access to and the relevant quality of that information. Establishing robust data governance practices is one approach to requiring an inventory of available data assets along with relevant characteristics as metadata. Having data governance processes in place can simplify implementation while ensuring data assets are approached with future potential data sharing in mind. Further, integration is increasingly possible across education, workforce, criminal justice, and public health systems to support improved insights about

outcomes and to encourage individuals eligible for government benefits and services to apply for and receive them.

7. **Institute Transparent Approaches to Using Data.** While some states rapidly developed approaches for various initiatives that prioritized transparency, this was not a consistent theme across many organizations. Moving forward, as more individuals in states are subject to federal law and practices, it will be critical to ensure that transparent approaches are available to demonstrate the value of data management and use. Transparency can be encouraged using websites and portals to share insights about available data assets and particular projects, including how data are used for a benefit. It can also be championed by governors and other senior officials who speak to how the approaches benefit individuals in each state. Sharing information with the public or stakeholders should be viewed as a requirement for advancing the use of data collection, management, and use.
8. **Partner with existing entities to promote rapid progress and pilot projects.** The need to develop quick wins and success stories cannot be overstated. States could partner with other states, existing research institutions, or nonprofits such as the Coleridge Initiative to promote rapid progress in modernizing data infrastructure at the state-level. Pilot projects are especially valuable because they can be relatively low cost, reach a range of audiences, and provide clear narrative about the value proposition for better using data in practice. Many federal programs include waiver authorities that can be leveraged to test new ideas and support data infrastructure, for example, through Medicaid and some human services programs funded in part by the federal government.
9. **Establish responsive engagement processes and oversight mechanisms.** Data activities should be reviewed periodically by an independent organization to ensure that intended goals and objectives are being achieved. States can support this approach by explicitly recognizing the need for responsive engagement from knowledgeable individuals. In addition, states can encourage the institution of certain oversight mechanisms in legislative bodies or with voluntary advisory committees. For example, states could opt to establish a new independent advisory committee charged with assessing progress in implementing data governance procedures and with addressing underlying concerns about state data.
10. **Plan for and implement robust privacy safeguards.** Advancing data uses and data sharing at the state-level can responsibly occur alongside data access by ensuring that privacy safeguards are robust. The Evidence Act suggests that privacy and data access improvements can be reinforcing and concurrent. States must plan for strong cybersecurity approaches and to increase training, cultural responsiveness, and awareness of deidentification techniques and other privacy safeguards.

Roadmap steps can be implemented at the same time. However, identifying leaders who in turn can secure sustained resources may be a relevant starting point in most states that are beginning to build more robust data infrastructure and capacity. Similarly, publishing a comprehensive data strategy or learning agenda may generate many ideas for how states can most effectively participate in such initiatives.

Focus on Improving Collection, Access, and Use of Localized Firearms Data

The expert panel commissioned by Arnold Ventures and convened by NORC at the University of Chicago to study firearms data observed fragmentation and gaps in the data ecosystem at the state and national level. Such fragmentation can be reduced by incorporating the roadmap from the previous section, regarding issues specific to firearms data. Indeed, some state and local jurisdictions have begun to improve capacity and infrastructure related to criminal justice topics; these efforts could be expanded upon to improve availability and access to information about firearms.

Allegheny County, Pennsylvania, which includes Pittsburgh, incorporates data from a wide range of programs into a portal for easy access and use by both the public and decision-makers. The portal includes data from criminal cases and the local jail as part of a larger warehouse of information used to identify strategies for early interventions and prevention among individual receiving services or with active criminal cases. Even with the initial infrastructure in place, the portal has gaps in coverage with more than 100 other police departments within the county. Summary tabulations are available to the public as open data through the county's QuickCount website, which enables the ability to monitor trends and view some additional detail for analysis by program, gender, age, ethnicity, and race.

In Maryland, a consortium formed the Baltimore Neighborhood Indicators Alliance with input from nonprofits, foundations, community members and the city government. The consortium brings together key data about the community, with visualizations and open data to support city accountability on specific indicators. For example, the portal includes a range of information for the city available by neighborhood on crime rates, shootings, gun-related homicides, arrest rates, and other relevant indicators.

Each of the local jurisdictions described above has practices in place to support access, governance of data, and transparency, in line with portions of the roadmap above. However, each also has room for improvement in addressing the needs for data linkage and sharing to strengthen the insights available using existing data. They offer examples of a starting point for other state and local jurisdictions that aspire to establish enhanced capacity for evidence-based decision-making.

National infrastructure may provide supports, efficiencies, and an economy of scale in supporting state and local jurisdictions in developing such an infrastructure on many topics, including firearms data. This is especially the case when a jurisdiction recognizes a need to integrate public health information while deploying strong cybersecurity safeguards and confidentiality protections. Roman (2020) has identified benefits for considering integrated approaches to firearms data and designing a data infrastructure that could facilitate improved data sharing capabilities in a secure environment. Leveraging the federal and more local

capacity, including available resources and systems, could accelerate approaches for bolstering the national, state, and local data ecosystem more rapidly and effectively in coming years.

Many states collect substantial data from local health and public safety departments, relevant to addressing firearms-related violence. States need not wait for the federal government to lead the way on data linkages. As demonstrated by the proactive and successful efforts of state consortia examining workforce and education data, sharing data in a secure environment that protects privacy and confidentiality of sensitive data is an effective approach to tackling these complex issues and does not require federal involvement.

Key Findings:

This report offers a set of key findings for states, based on the lessons learned and experiences described above.

- Bipartisan support is important. There is value in having better information to understand outcomes that individuals from different political or ideological perspectives can agree on, such as better service delivery, more effective use of taxpayer dollars, increased emergency preparedness, economic development and job creation, improved health and safety, and full employment.
- A short-lived commission created by a legislature (or governor) could make recommendations for legislation. The commission should have members with technical expertise, not be overtly political, and should take input from a variety of public interests as well as relevant state and local agencies. It is important for groups with specific interests—such as privacy advocates, academic researchers, communities whose data may be shared across agencies because they participate in government programs, and others—to share their interests and concerns as part of the process.
- If a commission is created by legislation, key sponsors of the legislation should be kept apprised of progress and be committed to advancing legislation that would enact at least some of the commission's recommendations right away.
- States need not replicate the precise bureaucracy and infrastructure that the federal government has created, to create useful, actionable information from their data. States may leverage existing infrastructure and capitalize on their resources, systems, and data assets and even include some of those maintained by federal agencies. Capacity is built through central coordinating and sharing mechanisms. Significant infrastructure elements must be strategically aligned to focus on priority work that extends beyond a single agency, to build sufficient capacity to address complex challenges.
- Establishing standards and methods that can be adopted by collaborating agencies will have a synergistic effect to increase capacity. Fewer resources will be needed in the long run for sharing data, analytical methods, and infrastructure improvements. Ideally, standards could be adopted by consensus.

- Cultural change is difficult. Agencies have traditionally been focused on their own missions. Impetus toward culture change must come from the top but also must provide value to each agency engaged in creating a more collaborative culture.
- Intergovernmental partnerships can accelerate innovation and progress through mutually beneficial exchange of data and resources. Successful partnerships leverage the creation of value for all partners.
- Successful partnerships require sustained commitment over long periods of time, with continuous improvements as the partnership matures.
- Partnerships do not need to wait for all pieces to be assembled to begin. Often federal partnerships begin with a vision shared by one or two states and build from there, as the value of the partnership becomes evident.
- Targeted federal investments can provide resources to support data modernization and linkage at the state and local level. This approach may require knowledge of available flexibilities and federal laws.
States and localities can make rapid progress on modernizing systems, standardizing data, and creating a data infrastructure by ensuring that federal funding is invested in alignment with a data strategy that will result in significant improvements to data quality, sharing, and analytics.
- States should maximize federal investments by taking a cross-agency, cross-functional approach to making improvements that comply with federal requirements. Such an approach can help remove state and local barriers to increased interagency data sharing.
- Agreement on core characteristics is necessary to support comprehensive capacity for data use at the state-level. Many states have established aspects of strong capacity that can serve as models for states seeking to improve data infrastructure.
- Building and improving capacity can be nonlinear and iterative, meaning there is likely no single path to success. Approaches should be tailored to meet relevant state priorities and constituent expectations or needs.
- Successful efforts have focused on specific topics, such as public health, education, and workforce issues. Such efforts have placed agency employees into environments where they can have hands-on training and gain experience working collaboratively across agency lines to improve data quality, create immediate insights and demonstrate value that can lead to sustainable processes. This type of collaboration can be expanded to firearms violence-related topics.

Conclusion

As states continue to pursue data infrastructure and capacity improvements, they should leverage federal investments and opportunities from existing state-federal partnerships to enhance existing capabilities. Identifying one or two priority topics can be useful in setting up structures and processes for institutional and cultural change, yet all of these activities described in this report must be considered scalable when designed. The capacity to use data is not built in the abstract; it should be specific to align with existing priorities, staffing capabilities, and resource availability.

If approaches are successful, states can facilitate data leadership for policy priorities in a way that solves real problems. They can ensure that data are accessible for operational and evaluative purposes. They can plan for processes that align evidence building and use. They can develop substantially enhanced and expanded privacy-protective data infrastructure. Most critically, states that have lagged in devising approaches and identifying resources to build further infrastructure cannot afford to wait longer to recognize the strategic value of using the data they have already collected.

In the twenty-first century, all Americans need access to reliable, valid information for decision-making at every level of government. Aligning some approaches at the state-level with recent progress and momentum from the federal government is an example of how resources can be more efficiently and effectively aligned to deliver the best services and policies possible for the public.

References and Resources

Abazajian, K. and Kleykamp, T. *Leveraging Data for Economic Recovery: A Roadmap for States*. Washington, D.C.: Georgetown University Beeck Center, 2020.

Doar, R. and Gibbs, L. *Unleashing the Power of Administrative Data: A Guide for Federal, State, and Local Policymakers*. Washington, D.C.: Results for America and American Enterprise Institute, 2017.

Economic Development Administration. (2021). Good Jobs Challenge. Accessed at: <https://eda.gov/arpa/good-jobs-challenge/>.

Fatherree, K. and Hart, N. *What is the Role of States in Evidence-Based Policymaking?* Washington, D.C.: Bipartisan Policy Center, 2017. Accessed at: <https://bipartisanpolicy.org/wp-content/uploads/2019/03/What-Is-the-Role-of-States-in-Evidence-Based-Policymaking.pdf>.

Fletcher, A. What Will it Take to End Family Homelessness? The Family Options Study. In Hart, N. and Yohannes, M. *Evidence Works: Cases Where Evidence Meaningfully Informed Policy*. Washington, D.C.: Bipartisan Policy Center. Accessed at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3766880.

Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act). Public Law 115-435. January 2017. Accessed at: <https://www.congress.gov/115/plaws/publ435/PLAW-115publ435.pdf>.

Garg, K., Guidera, A. and Hart, N. *Improving Data Infrastructure to Meet Student and Learner Information Needs*. Washington, D.C.: Day One Project, 2021. Accessed at: <https://www.dayoneproject.org/post/improving-data-infrastructure-to-meet-student-and-learner-information-needs>.

Gartner <https://www.gartner.com/en/information-technology/glossary/data-governance>, 2021

Goerge, Robert, 2019. "[Engagement with State and Local Government](#)," [SocArXiv](#) v6tc2, Center for Open Science.

National Academy of Public Administration (NAPA). *Advancing Information Policy and Use at the Office of Management and Budget*. Report of the Select Task Force: Organization of Information Policy and Use at OMB. Washington, D.C.: NAPA, 2020. Accessed at: https://napawash.org/uploads/OMB_Information_Policy_and_Use_Working_Group.pdf.

Newcomer, K. Olejniczak, K., & Hart, N. *Making Federal Agencies Evidence-Based: The Key Role of Learning Agendas*. Washington, DC: The IBM Center for The Business of Government, 2021. <http://www.businessofgovernment.org/report/making-federal-agencies-evidence-based-key-role-learning-agendas>.

Pew Charitable Trusts. How States Use Data to Inform Decisions. Pew Charitable Trusts, 2018. Accessed at: https://www.pewtrusts.org/-/media/assets/2018/02/dasa_how_states_use_data_report_v5.pdf.

Results for America (RFA). 2020 Invest in What Works State Standard of Excellence. Washington, D.C.: Results for America, 2020. Accessed at: <https://2020state.results4america.org/>.

RFA. 2021 Invest in What Works State Standard of Excellence: Criteria and Sub-Criteria. Washington, D.C.: Results for America, 2021. Accessed at: <https://results4america.org/wp-content/uploads/2021/05/2021-Invest-in-What-Works-State-Standard-of-Excellence-Criteria-and-Subcriteria-Final.pdf>.

Roman JK. 2020. The State of Firearms Data in 2019. Bethesda, MD: NORC at the University of Chicago.

Roman JK. 2020. A Conceptual Framework for a Firearms Data Infrastructure. Bethesda, MD: NORC at the University of Chicago.

Roman JK. 2020. A Blueprint for a US Firearms Data infrastructure. Bethesda, MD: NORC at the University of Chicago. 2 Roman JK. 2020.

Rothwell, C. The National Vital Statistics System. Presentation to the U.S. Commission on Evidence-Based Policymaking on January 13, 2017.

U.S. Commission on Evidence-Based Policymaking (CEP). *The Promise of Evidence-Based Policymaking: Final Report of the Commission on Evidence-Based Policymaking*. Washington, D.C.: GPO, 2017.

U.S. Department of Housing and Urban Development (HUD). *Family Options Study: 3-Year Impacts of Housing and Services Interventions for Homeless Families*. Washington, D.C.: HUD Office of Policy Development and Research, 2016. Accessed at: <https://www.huduser.gov/portal/sites/default/files/pdf/family-options-study-full-report.pdf>.

U.S. Department of the Treasury. (2021). Compliance and Reporting Guidance: State and Local Fiscal Recovery Funds. Washington, D.C. Accessed at <https://home.treasury.gov/system/files/136/SLFRF-Compliance-and-Reporting-Guidance.pdf>;

Treasury. (2021). Coronavirus State and Local Fiscal Recovery Funds Interim Final Rule. 86 FR 93 pp 26786. Washington, D.C. Accessed at: <https://www.govinfo.gov/content/pkg/FR-2021-05-17/pdf/2021-10283.pdf>.