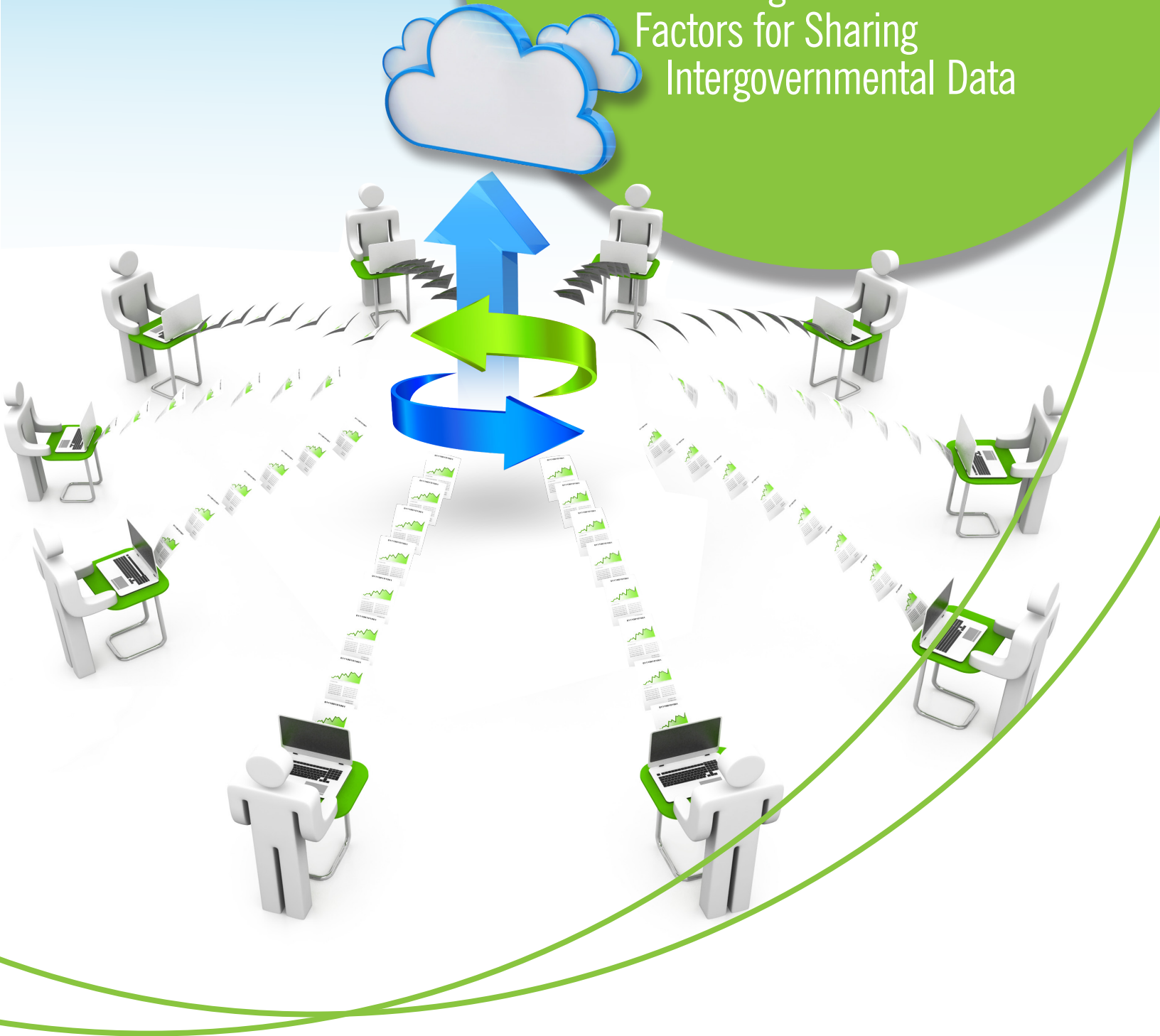


Silo Busting:

The Challenges and Success Factors for Sharing Intergovernmental Data



Jane Wiseman

Ash Center for Democratic Governance and Innovation
Harvard Kennedy School



IBM Center for
**The Business
of Government**

2020

Silo Busting: The Challenges and Success Factors for Sharing Intergovernmental Data

Jane Wiseman

Ash Center for Democratic Governance and Innovation
Harvard Kennedy School



TABLE OF CONTENTS

Foreword	4
Executive Summary	5
Introduction	8
The Value of Intergovernmental Data Sharing	13
Rapid Deployment of Intergovernmental Emergency Response	14
Improved Service Delivery	15
Better Allocation of Resources	16
A Seamless User Experience	18
Seven Challenges to Sharing Data Across Government	19
Challenge 1: People, Process, and Culture—Not Technology	20
Challenge 2: Resistance to Sharing Data	21
Challenge 3: Data Collected Without a Plan for Use, Quality	22
Challenge 4: Incomplete Data Sets	22
Challenge 5: Data Is Not Digitized	22
Challenge 6: Lack of Data Standards	23
Challenge 7: Fears About Legal Authority to Share	25
Vision of an Ideal Data Sharing Ecosystem	26
Examples of Successful Data Sharing Initiatives	32
Examples of Federal Data Sharing Initiatives	33
Examples of State Data Sharing Initiatives	34
Examples of Local Data Sharing Initiatives	38
Four Success Factors for Sharing Intergovernmental Data	40
Success Factors Related to Leadership	41
Success Factors Related to the Team	42
Success Factors Related to Processes	43
Success Factors Related to Data	44
Recommendations	46
Conclusion	51
Sources	52
Interviews with author	56
About the Author	57
Key Contact Information	58
Reports from the IBM Center for The Business of Government	59

FOREWORD

On behalf of the IBM Center for The Business of Government, we are pleased to present this report, *Silo Busting: The Challenges and Success Factors for Sharing Intergovernmental Data*, by Jane Wiseman, Harvard University.

The COVID-19 pandemic has clearly demonstrated the importance and value of being able to share data quickly between levels of government. Even with the stumbles that have occurred in standing up a national system for sharing pandemic-related health data, it has been far more successful than previous efforts to share data between levels of government—or across government agencies at the same level.

Ms. Wiseman offers a rich description of what intergovernmental data sharing can offer by describing a range of intergovernmental data sharing initiatives in various policy arenas, such as social services, transportation, health, and criminal justice.

She identifies common challenges that serve as barriers to more effective data sharing and relies on the insights developed from the various case studies to identify the key success factors for sharing intergovernmental data. She then offers a set of recommendations to guide government officials on ways they could undertake data sharing initiatives.

We hope this report provides leaders at all levels of government a useful set of actions that they could undertake in order to improve service delivery, make better decisions about resource allocation, and operate more seamlessly in serving citizens.



DANIEL J. CHENOK



TIMOTHY PAYDOS

A handwritten signature in black ink, appearing to read "D. Chenok".

Daniel J. Chenok
Executive Director
IBM Center for The Business of Government
chenokd@us.ibm.com

A handwritten signature in black ink, appearing to read "T. Paydos".

IBM Vice President
Global General Manager
for Government
tpaydos@us.ibm.com

EXECUTIVE SUMMARY

Effective intergovernmental data sharing is like the service at a fine restaurant—the better it is, the less you notice it. Yet, in both cases, while the result looks effortless, it takes significant time and toil behind the scenes.

The work of knitting together large government data systems so that they seamlessly connect and provide customer-friendly services to the public is difficult yet achievable and valuable. If this type of data sharing was easy, everyone would be doing it. Instead, there are only a handful of outstanding examples of success and a lot of barriers to achieving it.

The state of data-driven government has advanced rapidly over the past decade, but it has not yet achieved its full potential. Single-agency successes are prevalent, with rapid acceleration of the use of data to solve problems within an agency. Peer-to-peer data sharing across units of government is increasing. This horizontal sharing among peers in one government (city, state, federal agency) is often orchestrated either by the agency itself or by a shared data or IT organization, often to solve a particular and clearly defined problem. This type of data sharing remains far more common than vertical sharing across layers of government, say from city to county and state or to federal. And yet, solutions to the most complex and vexing public problems require data sharing that spans boundaries of government agencies.

As an example, addressing homelessness can't be solved with just housing data but also requires data about an individual's situation and needs across employment and education, health and mental health or substance use, and criminal justice sectors. Yet, in each of those service delivery silos, most staff have no incentive to go outside of their sphere of responsibility to get at the root cause of the problem, nor do they typically have authority to access external data sources. As a result, data analytics projects that are cross-departmental require alignment across many factors and are both uncommon and inspiring.

With digital data being created at a dizzying rate with every mouse click, or swipe of a device to enter a building and credit card transaction, the world is awash in data but not yet keeping up with analysis of this data. How is data connected and used to drive action in government? Not nearly enough. And yet, there are some positive points of reference.

The Value of Intergovernmental Data Sharing. Intergovernmental data sharing can drive significant public value by enabling more efficient emergency response, improving service delivery, facilitating a better allocation of resources, and creating a seamless user experience. For example, the Commonwealth of Virginia was able to stand up its COVID-19 dashboard in a matter of days because it had already created a data sharing platform that integrated public safety, public health, and other data in response to its opioid crisis. Allegheny County built a data warehouse that now enables them to prioritize service delivery where it is most needed. Individual level data has led to development of risk modeling tools for child welfare and for homeless service delivery. The Mobility Data Specification speeds the allocation of micro-mobility transportation resources (e.g., scooters and bicycles) and improves multimodal transportation planning. And in Singapore, integrating data across systems and silos that enables a seamless user experience such as with its user-centric “Moments of Life” apps.

Challenges to Sharing Data. The biggest obstacles are not related to the technology, but to culture—there are core “people issues” that pose risk to any transformation initiative. Change management challenges for government data projects go far beyond normal human resistance to change. For example, staff time for those with the talent to conduct complex data analysis is limited, and turnover in the most in-demand public sector data analytics roles further constricts available expertise. Sharing data outside of a specific agency or outside of its intended purpose also goes against organizational culture often prevalent in government that avoids risk, focuses on narrowly defined tasks, and resists sharing of information for fear of being wrong. Often, there is a perception of a statutory barrier when in fact no barrier exists, and this can provide protective cover for those who simply don’t want to share. Sometimes, those afraid of sharing data will use privacy rules as an excuse to not share data, rather than admitting their anxiety (to themselves or others). Finally, in government, turf issues can significantly impede data sharing. While these challenges can be overcome, for example when a crisis creates urgency for collaboration, or when a disruptive technology compels government innovation, the barriers government data leaders face are significant.

Vision of an Ideal Data Sharing Ecosystem. The intergovernmental data sharing examples in this report demonstrate that a more seamless and customer-oriented government is possible. If progress could be achieved across government functions, an entirely new way of serving the public could leverage both the power of data sharing on the back end, and the front-end capability to make transactions easier on the public with digital services. Today, while many governments excel at data analysis and insight, or at developing user-centric digital services, few excel at both.

Success Factors for Sharing Intergovernmental Data. While the cases and examples of data sharing described in this report are diverse in many ways, they typically share four success factors:



Leadership. A leader with a clear and compelling vision motivates the team, and this is especially important with a data sharing project, which can require patience and persistence to achieve results



The Team. The staff working on an intergovernmental data sharing project need data skills, but they also need judgement and interpersonal acumen as well in order to broker data sharing agreements.



Process. The process of creating intergovernmental data sharing platforms can take a long time and requires patience and persistence.



Data. Data analysis and data sharing efforts are only as good as the underlying data, so data quality issues must be addressed.

The most important success factor is leadership. Without dedicated, talented, and inspirational people leading and managing these initiatives, none of the excellent examples described in this report would be possible. The vision of leaders fuels the persistence and patience of all team members.



Recommendations. Based on the findings from the relevant literature, expert interviews, and the case studies and examples described here, the author offers four recommendations to advance intergovernmental data sharing, accompanied by detailed actions, that include the following:

- **One:** Congress and the president should create a policy and governance framework. They should define a broad vision, with incentives to act, and a strong data governance structure. This would include actions such as establishing an “ask once” goal for data collection, rewarding agencies that link their data sets, and creating intergovernmental data councils.
- **Two:** Congress and the president should establish funding and capacity building mechanisms to support implementation of increased data sharing across all levels of government. This would include actions such as supporting data literacy efforts in federal agencies and among federal leaders, funding for data sharing projects, and resources to improve data quality.
- **Three:** The nonprofit and philanthropic sectors should proactively support intergovernmental data sharing efforts. This would include actions such as providing incentives to innovate and link different sources or types of data at the state and local levels, and supporting information exchange networks.
- **Four:** Agency managers and data leaders at all levels of government should champion data sharing efforts. This would include actions such as articulating and creating a shared vision for data sharing, establishing shared data standards and protocols, and sponsoring communities of practice for data enthusiasts.

INTRODUCTION

Today, there is a growing understanding of the power of data integration to solve vexing and entrenched public problems and the value of seeing the whole person across their many touchpoints in public service systems.

Policy analysts and data stewards are crafting ways to link data across systems so that policy-makers can track services delivered and results achieved by individuals and families over time and across multiple stages of their lives.

Over the past decade, data and analytic capacity in federal, state and local government has advanced rapidly. Notably, since the 2019 enactment of the Foundations for Evidence-Based Policymaking Act of 2018,¹ federal agencies subject to the law appointed chief data officers, evaluation officers, and statistical officials and have begun to implement their required actions under the Federal Data Strategy, such as creating inventories of data assets and creating learning agendas to improve data literacy in their agencies.

However, while impressive progress in advancing the use of data in government has been made, there is still a gap between what could be and what is. Most data sharing successes are single agency projects in data analysis, visualization, or insight. Few case studies exist on how data sharing across departments, across levels of government, or with the private sector can achieve public benefit by adding increasing value with each data set shared or compared. Tapping the vast stores of data held by different parts of government across departments, and across levels of government can dramatically improve success in achieving the goals of the Federal Data Strategy and Action Plan, which specifically call out the goal of using data as a strategic asset.

REPORT OBJECTIVES, SCOPE, AND METHODOLOGY

The objectives of this report are to (1) describe the complexity and challenges of connecting data across government functions and organizations to devise more effective and efficient services to meet public needs, (2) shine a light on the success cases, and (3) based on these and other case studies, define the key factors for successful intergovernmental data sharing. The report limits its scope to linking data created by government for government insight or action, with a respectful nod to academics who have done this work outside of government.

The report is based on interviews with leading government data experts and a series of case studies of intergovernmental data sharing. None of the case studies profiled were without challenges, yet with vision and persistence they have succeeded in creating public value by sharing intergovernmental data.

1. Public Law No: 115-435, January 14, 2019.

Government Collects and Generates Vast Quantities of Data. Government collection of data is not new—it dates back to the earliest civilizations. Long before the Greek and Roman civilizations, the Babylonians² were conducting a regular census of not just people, but of commodities³ such as livestock, butter, honey, milk, and vegetables. Statistical data collection for our national census began in 1790. In France, by 1825, detailed crime statistics were being collected by the French Ministry of Justice, describing every arrest and conviction by region.⁴

Most public sector activity generates detailed administrative data about individuals and about transactions—whether it is the information provided in a building permit or tax form, the information shared in an application for unemployment benefits, or the details of daily life for an inmate in a jail or prison. The City of Chicago, for example, produces seven million⁵ rows of data every day, ranging from library transactions to potholes filled to police reports.

While government does collect a great deal of administrative data, it is often not used to generate insight or create value. As noted by former New York City agency heads Robert Doar and Linda Gibbs, administrative data is widely available and yet an untapped tool for governments seeking data insight:

To be sure, the issue is not that government agencies lack data (indeed they often have more data, more accurately collected than anyone else) but that the repositories of these data are highly protected and bureaucratically controlled. Much of the country's administrative data—collected by government entities for program administration, regulatory, or law enforcement purposes—is underappreciated, underdeveloped, and underused.⁶

Academic Researchers Have a History of Using Disparate Data Sets to Gain Insight.

Researchers have a long history of linking one source of data to another to find deeper insight than is available in one data set alone. Several university researchers are doing exemplary work in this area, including Harvard Economist Raj Chetty's work combining Census and income data in a longitudinal examination of America's failure to provide economic mobility low income and minority youth. At the University of Michigan, economist Michael Mueller-Smith is connecting Census and justice system data for a number of states. Nonprofit organizations such as Measures for Justice are also compiling and combining data and publishing it for public value.

Researchers are often by design at a distance from their subjects, so that they remain objective. But there is also a precedent for researchers, and more recently data scientists, becoming embedded in government operations. Washington, D.C. now has a scientific team, the Lab@DC with data scientists, evaluation specialists and policy analysts both in a central team and also embedded in agencies such as the police department. This type of engagement has a history. When the National Institute of Justice created its Research Action Partnership Grant program in the mid-1990s, it provided funding to create interactive and collaborative relationships between

2. Office for National Statistics, "Census-taking in the ancient world," January 18, 2016, <https://www.ons.gov.uk/census/2011census/howourcensusworks/aboutcensuses/censushistory/censustakingintheancientworld>.

3. PRB, "Milestones and Moments in Global Census History," September 4, 2019, <https://www.prb.org/milestones-global-census-history/>.

4. Hannah Fry, Your Number's Up, *The New Yorker*, September 9, 2019

5. Wiseman, Jane. "Lessons from Leading CDOs," January 2017.

6. Doar, Robert and Linda Gibbs. "Unleashing the Power of Administrative Data: A Guide for Federal, State, and Local Policymakers," Results for America and American Enterprise Institute. October 19, 2017.

researchers and practitioners. Many of those partnerships still exist today and are driving insight and reducing crime and victimization. Other initiatives to bring outside ideas into government, such as academic fellowships, executives on loan and interagency personnel agreements exist, but they remain the exception not the rule. But with only a handful of these efforts and 90,000 units of local government across the country, it is still too little.

Intergovernmental Data Sharing Is the Most Complex Type of Data Sharing. This report focuses on the most complex form of data sharing—intergovernmental data sharing, which spans the boundaries of a government organization, connecting to sources outside of its own span of control. While the term “data sharing” can mean many different things, for purposes of this report, a typology of the levels of complexity of data sharing methods is detailed in Table 1. Data sharing is not new—complex data sharing involving a point-to-point query has been in place for many years. More complex is the creation of a sharable platform for data, and the most complex is the sharing across levels of government or even beyond the boundaries of government.

Table 1. Intergovernmental Data Sharing

Data sharing Level of Complexity	Query (Complex)	Platform (More Complex)	Intergovernmental (Most Complex)
Purpose	Transactional data exchange	Create and publish shared data platforms	Compare across sources to create policy insight
What it is	<ul style="list-style-type: none"> Provide communication via direct exchange of data typically between two points, either via one way or two-way sharing using established data standards and schemas. 	<ul style="list-style-type: none"> Create common data platforms such as data warehouses, data portals. Provide role-based access to the public and internal users. 	<ul style="list-style-type: none"> Bring together data in a secure platform that can be accessed by authorized users for insight and analysis, and are specifically designed to create actionable intelligence.
What it does	<ul style="list-style-type: none"> Provide one-time answers to yes-no or other binary inquiries. 	<ul style="list-style-type: none"> Enable comparison and description by combining sources. Data standards and data governance allow increasing sophistication of analysis. 	<ul style="list-style-type: none"> Enable deeper insight by comparing individual-level data across sources, and across time.
Selected Types of Uses/ Examples	<ul style="list-style-type: none"> States query national data on income when assessing applicants for benefit programs Local police officers run checks of drivers licenses against national databases of those with outstanding warrants. Child support agencies can query the National Directory of New Hires for wage information for noncustodial parents. 	<ul style="list-style-type: none"> Open data portals for cities, states, and federal agencies Performance dashboards, including Performance.gov Results dashboards, such as College Scorecard GIS mapping portals Typically across one jurisdiction or department, but can span boundaries. 	<ul style="list-style-type: none"> Human services Integrated Data Systems (IDS), and Allegheny County Human Services data warehouse in particular Mobility data shared via the Mobility Data Specification Statewide Longitudinal Data Systems

Examples of data sharing that span organizational boundaries in government may be vertical or horizontal in nature (see Table 2). Horizontal data sharing is among peers; for example, departments in the same city or state government. Vertical data sharing spans levels of government; for example, integrating city, county, state, and or federal government data into one analytics platform. Vertical data sharing may also include third-party data from sources external to government, including private sector sources as well as public input received via social media sentiment mining, crowdsourced and real-time feedback, surveys, and the like. The key benefits of both horizontal and vertical data sharing are shown in the table below.

Table 2. Benefits of Horizontal and Vertical Data Sharing

Type of data sharing	Characteristics and key benefits
Horizontal	<ul style="list-style-type: none"> Peer to peer integration of data at the back end can create a seamless user experience on the front end (e.g. user centric design for government web sites that provides navigation based on user task not agency boundaries). This is most frequently for data already in a standard format such as GIS data. Integrating data at the person level across multiple agencies can improve personalization of service delivery and identify service gaps.
Vertical	<ul style="list-style-type: none"> Integration across levels of government and with third parties simplifies the user experience (e.g. city road closure, DOT data shared with Waze) Insights can be mined from all levels of government and from third party data sources to develop predictive profiles that help target resources where most effective

Data Sharing in Government Is Not New. The ability of curious public servants to devise innovative methods to connect data across the silos of government is not new, and over the years, two models have been developed to operationalize such systems.⁷

In the 1980s, leading urban geospatial data experts realized they could link location-based data across sources and created integrated mapping tools in cities like San Diego⁸ and Louisville.⁹ As described later in this report and in a separately published companion case study,¹⁰ in 2003 the Allegheny County, Pennsylvania, Department of Human Services, combined state and county data in their data warehouse.¹¹ Yet, this idea has not yet caught on widely. Nor has the idea of connecting disparate data to create a full picture of a policy problem always been welcome.¹² The aim of this report is to draw attention to existing successes and build momentum for additional valuable data sharing efforts.

7. These are described in more detail here: <http://www.businessofgovernment.org/blog/two-models-successful-intergovernmental-data-sharing>.

8. SanGIS.org, About SanGIS, 2012, <http://www.sangis.org/about/history.html>.

9. <https://www.lojic.org/>.

10. Wiseman, "Better Government Through Data: Using the Allegheny County Human Services Data Warehouse to Design More Effective Results," available at <https://scholar.harvard.edu/janewiseman/publications/better-government-through-data-using-allegheny-county-human-services-data>, completed as part of the research for this report.

11. Allegheny County Department of Human Services, "Allegheny County Data Warehouse," July 2019.

12. Note: the author of this paper was dis-invited from participating in a 2007 interagency working group at a police agency, after presenting a preliminary analysis that demonstrated hypothetical integration of education, substance use treatment, medical and foster care data to get a "whole client" view of a typical frequent utilizer of police services.

AREAS FOR POTENTIAL ADDITIONAL RESEARCH

The field of data sharing is ever evolving and a number of questions that need to be addressed are beyond the scope of this report, which focuses specifically on intergovernmental data sharing. Topics for potential future research and the development of case studies include:

- **The role of bias and ethics.** Issues related to bias or ethics need further explored, especially when predictive models affecting individuals are involved.
- **The impact of automation and artificial intelligence.** Automation and artificial intelligence have potential to improve data loading, data cleansing, data quality and reduce labor-intensive data processes, but are they affordable to localities? Can partnerships be developed with local universities?
- **Privacy and personally identifiable information.** Might the European data privacy model—as reflected in the European Union’s General Data Protection Regulation (GDPR) serve as a framework for U.S. approaches, especially as more is learned from its implementation?
- **Handling the storage cost of data and its ownership.** What are the cost implications for storing data sets centrally, such as the creation of data warehouses, vs. providing access to distributed, shared data sets? Related is the issue of who owns the master data, and how possible contentious issues related to this are resolved.
- **Addressing potential national security concerns.** When federal agencies share and merge data across federal agencies, can they be sufficiently separated in order to not pose a national security concern. That same data combined, could become classified. How should this be included in a robust intragovernmental data governance model?

The Value of Intergovernmental Data Sharing



Intergovernmental data sharing can drive significant public value, incrementally adding value with each additional data set shared or compared. The value of data sharing efforts is often hard to measure, as government does not typically quantify operational efficiency metrics (such as labor time saved and the like). Yet, there are a handful of ways to concretely demonstrate value in:

- Rapid deployment of intergovernmental emergency response
- Improved service quality
- A better allocation of resources
- Creating a seamless user experience

Following are examples of each:

Rapid Deployment of Intergovernmental Emergency Response

The Commonwealth of Virginia leverages its Data Trust to provide real-time COVID-19 information. Virginia's ability to create an executive decision-making dashboard for COVID-19 in a matter of days is an excellent example of both the power of a crisis to compel data sharing, and the value of incremental growth of data sharing efforts. In the decade before the pandemic struck, the Commonwealth of Virginia's Chief Data Officer Carlos Rivero led the development of a safe, secure, and legally compliant information sharing environment that establishes consistent requirements through a standardized data sharing agreement process, the Commonwealth Data Trust. This process allowed for the collaboration between the Virginia Department of Criminal Justice Services and the chief data officer (CDO) to develop and pilot a data sharing platform that could respond to the opioid crisis in the Northern Shenandoah Valley. The shared data platform that grew out of the opioid crisis now serves as an extensible common platform for turning data into insight, and harnesses state, local, and federal data along with private sector data sources as well.

The urgency of the opioid crisis enabled the breaking of data silos in one of the most heavily impacted communities and the data effort sought to compare across data sources to identify age of first use, and the time from first use to first criminal justice system involvement. With cross-data set insight, appropriate interventions were crafted to prevent onset of opioid use among youth and an additional \$1 million in grant funds were garnered because of the deep understanding of the problem enabled by the data insight.

Significant effort went into creating the data sharing protocols to assure privacy and security of the data, working collaboratively with a statewide Data Sharing and Analytics Advisory Committees. Developing the standards that enabled data to be compared across the many source systems took a concerted effort, including months of cross-agency meetings and dozens of "road shows" by the state's CDO and his team. Creating data standards didn't start from scratch but rather built on the established National Information Exchange Model (NIEM) protocols developed through a multiyear national public-private effort.

What started as a way to integrate data for the opioid response has now expanded to incorporate broader justice data analysis, workforce development, overall state performance management, and now the response to the COVID-19 crisis. Data included in the COVID-19 dashboard is updated frequently, in some cases, as often as every 15 minutes. This gives state leaders near real-time information about hospitals in need of supplies and pharmaceuticals, hospitals, and regions that have surge capacity, and locations with the largest outbreaks of COVID-19 cases.

Rivero noted that the speed of deployment of the new dashboards for COVID-19 data relies on the foundation set by the work on the opioid crisis as well as incremental growth since that time. “Had we not previously had the technical, legal, and governance infrastructure in place, the expansion that took us just days to complete would have taken months. We were prepared; and that preparation allowed us to best support our constituents and communities during a time when it’s needed the most.”

Improved Service Delivery

Allegheny County prioritizes service delivery with data insight. In Allegheny County, Pennsylvania, (which includes the city of Pittsburgh) human services case workers serving elders, individuals with disabilities, children in the child welfare system, and individuals experiencing homelessness have integrated, real-time client level data that lets them tailor services to best meet client needs, and in some cases to proactively anticipate needs. For example, the Allegheny Family Screening Tool (AFST), a predictive analytics tool built using the data warehouse improved the accuracy of screening-in for children in need of services, and reduced racial disparities in case opening rates between black and white children.¹³ Launched in 2016 and updated in 2018, this tool helps frontline workers make child welfare call screening decisions by calculating a risk score, integrating and analyzing hundreds of data points from across multiple data sources. The risk score predicts the long-term likelihood of out-of-home placement and adds some data insight to the human judgement necessary when making a decision about whether to investigate a call about potential child maltreatment.

Allegheny County has also recently begun using a risk model to help prioritize homeless services, a scarce resource which can be provided to about half of those in need in the county. The goal was to speed the process of establishing priority for services by creating a standard measurement of need for services, leveraging existing client level data in the data warehouse, rather than asking individuals in need to provide the data themselves again, given that it was already in the system, and that the self-assessment process is not only time-consuming but can be fraught with inaccuracy owing to stigma and to the fact the individual may be in crisis. The risk prioritization system uses existing client data to automatically model future adverse outcomes, such as worsening mental health, incarceration, or emergency medical services. Identifying the individuals at greatest risk of these outcomes allows them to be given priority for homeless services, improving their health and potentially lowering overall cost for the county.¹⁴

To combat the opioid crisis, Allegheny County examined overdose death autopsy records and learned that over two-thirds (68.4 percent)¹⁵ of those who died had a prior interaction with the county’s human services agency. Deeper analysis showed that 18 percent¹⁶ of the fatal overdoses having been incarcerated in the prior year, and 49 percent¹⁷ had been incarcerated at some point in the past. Next, they examined how long it had been between the fatal overdose and the last contact with Allegheny County’s Department of Human Services (ACDHS).

For those who had contact with ACDHS in the year prior to their overdose, most had that contact in the 90 days before their fatal overdose, and many had contact with ACDHS in the 30

13. https://www.alleghenycountyanalytics.us/wp-content/uploads/2019/05/Impact-Evaluation-Summary-from-16-ACDHS-26_PredictiveRisk_Package_050119_FINAL-5.pdf.

14. https://www.alleghenycountyanalytics.us/wp-content/uploads/2020/09/9-2-2020-AHA-methodology-report-without-comments_pdf.pdf.

15. Karen Hacker, et al., “Linking Opioid-Overdose Data,” Public Health Reports, <https://journals.sagepub.com/doi/pdf/10.1177/0033354918803938>.

16. Ibid.

17. Ibid.

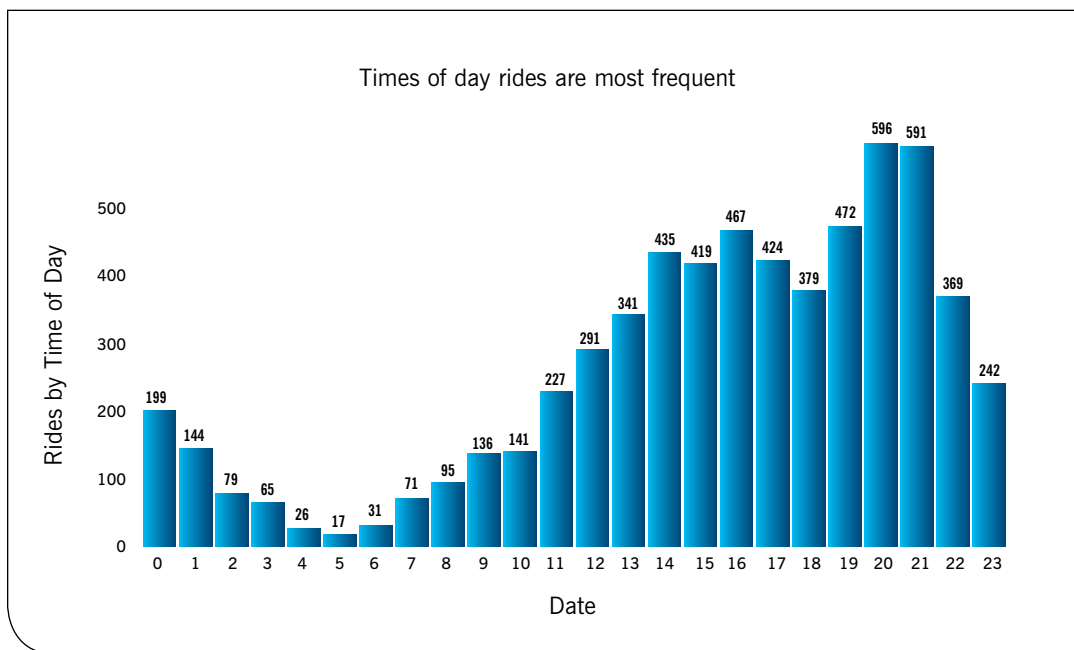
days prior to their fatal overdose. Based on this insight, ACDHS saw an opportunity to reach people as they exit service to protect them from a fatal overdose in the 30- and 90-day period following service.

Drug overdose risk rises for individuals who begin using again after a period of abstinence as their tolerance has decreased and they may not realize it—this can make it risky to go back to their prior drug dosage. This can be exacerbated if there has been any change in potency of street drugs during the time of incarceration or inpatient substance use disorder services. To combat this risk, ACDHS is now promoting distribution of overdose reversal drugs such as naloxone at discharge from the county jail and from county funded substance use disorder treatment. Further, drug treatment in jail is being increased along with transition services to help prisoners leaving jail access treatment after incarceration.

Better Allocation of Resources

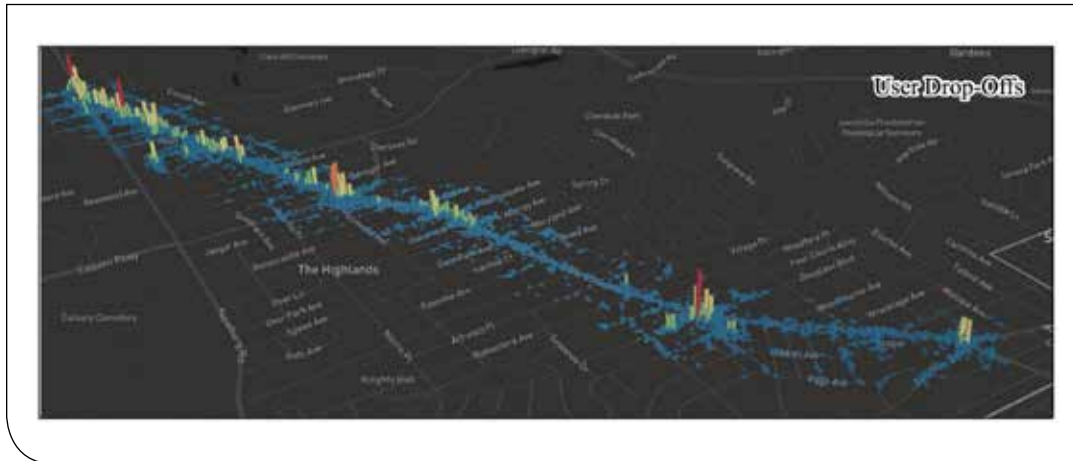
The Mobility Data Specification standard speeds deployment of data sharing for transit mobility. An intergovernmental data sharing model that goes beyond government boundaries to include private sector data, the Mobility Data Specification (MDS) enables over 80 cities to manage their micro-mobility vendors and to gather the data for local mobility policy decision making. This common data standard speeds analysis in cities, enabling multi-modal transportation and safety planning, and allowing real-time performance monitoring for micro-mobility vendors as well as transparency for regulatory compliance and pricing for permits. With MDS, Louisville, Kentucky, Mobility Manager James Graham can automatically generate a combined view of all electric scooter traffic by time of day and can look for patterns that would show or example if the nighttime safe riding curfew is being violated, as shown in Figure 1.

Figure 1: Electric scooter rides by time of day



In another example, Graham was able to generate heatmaps of where scooters are dropped off, analysis that was helpful when the city created designated parking spots for drop-off of dock-less scooters in response to concerns about scooter “clutter” on sidewalks in busy pedestrian areas as shown in Figure 2.

Figure 2: User drop-offs of electric scooters



Michael Schnuerle, former chief data officer for Louisville noted that the MDS data standard helped him quickly publish trip data after scooters appeared unexpectedly on local streets. The standard allows the city to require standard data from all e-scooter operators and enables the city to work with an API (application programming interface) feed to easily prepare mobility data for internal sharing or aggregation and preparation for publication on the open data portal. In 2018, Louisville became the first city to make aggregated MDS trip data available to the public, along with the methodology for doing so, and now four other cities are doing so as well.



MDS has also enabled the rapid deployment of equity and sustainability goals with easy access to data for analysis and pattern detection. For example, Kansas City, Missouri, has used data to designate “opportunity zones” characterized by low income and concentrations of minority residents. As vendors seek permits in the city to try out mobility innovations, the city uses these designated zones as required locations for a minimum threshold of services by those companies.

New York City has mandated that for the city’s carshare program,¹⁸ companies agree to put 20 percent¹⁹ of on-street spaces and vehicles in areas with low income residents. Using city income and equity data, the New York City Department of Transportation (DOT) can identify gaps in the new mobility networks and seek to address them. NYC also requires that bike-share companies provide the city with bike and dock availability data in real time. The information is published on the city’s open data portal and available to the public. For ride-hailing companies like Uber and Lyft, the city mandates that the companies share with the city data such as ride origin and destination, and whether the driver’s app was on or off. With this data, the city can monitor how long a driver has been on the road and can prevent overly-fatigued drivers from causing crashes.²⁰

A Seamless User Experience

Singapore shares data on the back end via data infrastructure investments. In the small city-state Republic of Singapore, integrating data on the back end enables a seamless user experience on the front end. Investment in shared platforms and a “single source of truth” for data governance significantly improves the efficiency of data and digital deployments.

For example, the Moments of Life app for families bundles services related to the birth of a child—a joyous time, but one that can also be a bit overwhelming. Now, registering the birth of a child has been reduced from 60 minutes to 15 and a parent need not go to 15 separate departments but can from one interface register the child’s birth, apply for the financial bonus paid to new parents, establish their unique national identity number, keep track of their immunizations (23 in the first 18 months of life!), and register them for a geographically convenient daycare and preschool. Two-thirds of eligible births are now registered through the app.²¹

18. <https://nycdotcarshare.info/>.

19. Gardner, Betsy. How NYC is using Transit Data for Equity Goals. February 2020. <https://datasmart.ash.harvard.edu/news/article/how-nyc-using-transit-data-equity-goals>.

20. For a more detailed description of how mobility standards have improved transit safety and efficiency, see the case study produced during the research process for this paper, “[Driving Toward Greater Safety and Efficiency: Urban Mobility Data Exchanges](#),” September 2020.

21. Smart Nation Singapore, “Building a Smart Nation with Tangible Benefits For Our Citizens and Businesses, Smart Nation Singapore,” February 28, 2020, <https://www.smartnation.gov.sg/whats-new/press-releases/building-a-smart-nation-with-tangible-benefits-for-our-citizens-and-businesses>.

Seven Challenges to Sharing Data Across Government



Data that an outsider to government might assume is routinely shared—for example, whether an individual in a local jail has been treated in that same community’s detox facility—is too often walled off and not shared outside the boundaries of the agency delivering the service. Despite the fact that sharing such data would provide safer, more humane service, it is too often not done simply because it is outside the status quo. With a fresh look, it makes sense to share that data, but the day-to-day reality of public service is that too often, workloads are high and resources are strapped so there is no bandwidth to think or work “outside the box” or the silo.

The largest challenge to data sharing in government is that it relies on human beings to do what might not be convenient, or that they may not view as important or in their own self-interest—and typically there is no forcing function to compel such data sharing. Here’s an example: Federal funding has enabled states to create integrated systems to connect outcome data on education by linking pre-school, elementary school, high school, and public college data to adult earnings. These Statewide Longitudinal Data Systems (SLDS) can add great value, as the ones in Georgia, Kentucky, Maryland, and Washington do. However, implementation has been inconsistent across states and in some cases the systems have stalled. One state SLDS leader noted that their system lacks college data because “the state university system just doesn’t want to share their data with us.”

And sadly, no one can make them do it, regardless of how incomplete the expensive system is without that essential data. This example is far from an anomaly—a survey of public sector IT officials found that 53 percent could not pinpoint problems in their data systems “because their systems were managed in silos.”²²



Challenge 1: People, Process, and Culture— Not Technology

The tools for matching data across sources and conducting analysis on large data sets are becoming both more widely available and less expensive. Yet, even with these advances in analytic tools, government data sharing remains a challenge. This is true in the private as well as the public sectors. In a survey by Government Technology, when asked their top data challenges, leaders in all sectors ranked “difficulty integrating data for multiple systems” as a top priority, with 63 percent in human services and 44 percent in transportation and infrastructure indicating this challenge.²³

The biggest obstacles are not related to the technology. Staff time for those with the talent to conduct analysis is limited, and turnover in public sector analytics roles further constrains available expertise. Additionally, data sharing goes against the prevalent organizational culture in government that avoids risk, focuses on narrowly defined tasks, and may resist sharing of information for fear of being wrong. Finally, in government, turf issues can significantly impede data sharing.

As one data leader noted, “The COVID-19 situation highlighted a lot of our typical data sharing challenges. We’ve had a very hard time even getting data from the county and state health departments, much less getting it in a timely and accurate fashion. I’ve also noticed inconsistencies in how case totals are reported from the county and the state, which is particularly unsettling.”

22. The State of IT Operations in the Public Sector, Splunk White Paper.

23. Government Technology, Interagency Data Sharing in the Time of COVID-19, Webinar Slide 6, May 29, 2020.

The problem is not limited to the public sector—the private sector lags in successfully integrating and using data. In their 2019 survey of Fortune 1000 companies, released in early 2020, NewVantage Partners found that while 99 percent of companies are investing in big data and artificial intelligence (AI), only 27 percent are achieving data culture.²⁴ The problem isn't technology, it's people—91 percent of survey respondents said that the biggest challenge is overcoming people, process and culture issues, not technology issues. And one in particular is the data literacy of senior leadership. In a survey of senior executives by analytics firm Splunk, while 83 percent of respondents agreed that data literacy is necessary to become a leader, 53 percent believed they are too old to learn such skills.²⁵

A 2020 podcast with leaders from the consulting firm McKinsey pointed out the challenge of sharing data: “Data—it's the lifeblood of the AI techniques used most often today. Most organizations have plenty of data within their veritable walls to fuel AI applications that improve areas from operations to product offerings. But it's the sharing of data across organizations that could unlock huge benefits for society. There's the potential to find cures to disease, to respond more effectively to crises, to combat climate change”.

McKinsey analysts David DeLallo and Jeni Tennison continued, noting: “Today, however, very little of the data sharing needed for such endeavors is happening. And this is due to a variety of reasons, from technical challenges to very legitimate privacy concerns and also because many organizations are simply hesitant to share their data because they see it as providing them with competitive advantage.”²⁶

While this reference to a reluctance to share data mentions competitive advantage, the hesitancy in government may stem from a variety of sources, including fear of being found to have data quality or completeness problems.



Challenge 2: Resistance to Sharing Data

A decade and a half ago, the author created a conceptual framework for creating a “whole customer” view of the families most at risk of being involved in the criminal justice system. A handful of likely scenarios were created that mapped to the experience of those most frequently in contact with the police agency leading the anti-violence working group. The cost of each of these scenarios were estimated using cost data for arrests, foster care, drug treatment, emergency medical services, and supervision provided a description of both the typical process pathways and the public cost of that hypothetical family of “customers” of the justice system.

After presenting this framework to the working group meeting of public and nonprofit stakeholders, the author was asked to stop participating in the group—the integrated snapshot of the “whole customer” connected to the multiple systems was viewed as a distraction. Focusing on the entire spectrum of the customer experience and looking at the combined government investment in these interactions was viewed as a waste of time. This experience took place at a highly regarded organization, which speaks to the status quo bias in even well-run government organizations.

24. NewVantage Partners, “Big Data and AI Executive Survey 2020: Executive Summary of Findings, 2020”.

25. Splunk, “The State of Dark Data,” Slide 18, 2019.

26. McKinsey. How to make the most of AI? Open up and share data. June 9, 2020, David DeLallo and Jeni Tennison. Podcast.



Challenge 3: Data Collected Without a Plan for Use, Quality

Most government administrative data is collected for the purpose of a particular service or transaction, not for research or for data analytics. This explains why for many data scientists, particularly at the state and local level, up to 90 percent of their time is spent cleaning data before they can use it.

When researchers, policy analysts, or data scientists in government want to compare data from one source to another, the task can be a tedious and manual one, often involving merging data from different sources into a single platform where it can be stored and analyzed. Integrating government data at the individual level (e.g., for someone being held in the local jail) involves merging records from all of the systems relevant to that individual (e.g., arrest record, jail booking, prosecutor and court records, mental health and substance use screening or treatment records as appropriate, medical services received in jail, educational, employment or rehabilitation programs, etc.). Data in these source systems is often of poor quality, with missing or inaccurate records.

The mere act of using data to explore policy questions can provide a virtuous cycle of improving data quality via the exploration and examination of the data. As noted by Janey Rountree, executive director of the California Policy Lab at UCLA, “You don’t understand the data quality issues until you work with the data and you don’t work with the data until you see a policy problem.”



Challenge 4: Incomplete Data Sets

The incompleteness of much data collection for bespoke purposes, and the challenge of poor data quality of the underlying data present significant problems. In commenting on the limitations of integrated data systems and administrative data in government, Henry Brady, dean of the Goldman School of Public Policy at the University of California, Berkeley noted:

Even with this linking, however, these data often lack useful ancillary information—unlike surveys, they do not automatically collect lists of socioeconomic characteristics such as education, income, age, and so forth on people or financial and historical information on firms or organizations. Moreover, even when this information is collected, it may be of low quality unless it is an essential part of the business purpose of the program (e.g., for welfare programs, income data are reliable because they are part of the application process, but education data are not). Intensive linking to other data sets can often expand their utility tremendously, but these matches are often precarious given the complexity of names, places, and other identifying information. Linkages using probabilistic matching techniques or geo-coding can help facilitate this process, but they still involve elements of uncertainty and incompleteness.²⁷



Challenge 5: Data Is Not Digitized

The success of data projects in Singapore owes a debt to the country’s significant investment in digital services. Each digital service generates data on the back end that can be used for analytics. Most governments are far from this level of digital maturity, and some

27. Brady, Henry E. “The Challenge of Big Data and Data Science,” *Annual Review of Political Science*, 2019. 22:297-323.

struggle to digitize and create value from paper records. The COVID-19 crisis has made clear how far behind the public health data infrastructure is—with 1,000 faxes a day creating piles of paper in Travis County, Texas, and with the State of Washington bringing in 25 National Guard troops to help do data entry on paper test results so that they can be included in the state system.²⁸

And, as noted by Sam Edelstein, former chief data officer for the City of Syracuse, New York, sometimes the data most valuable to an analytics project isn't in an electronic format. As he writes, in comparing the role of city data leader to a minor league baseball statistician, "Data that only exists on paper is a constant challenge. Sometimes digitizing the information is too costly and not valuable enough to justify the effort, so the data is never used."²⁹



Challenge 6: Lack of Data Standards

Currently, rather than a single integrated record for an individual that could facilitate a seamless transaction, there are many varied government records for every individual, sometimes even within the same agency, department, or bureau. Too often, even records for the same individual may not easily be matched.

For example, the voter registration for an individual may say Jane Smith, and the census may list Jane M. Smith, but the tax records are for J.M. Smith—and a business ownership certification may use another variant of the name. In this example, each agency would have to devote time and money to standardizing name and address data, and likely other data fields as well. Yet there may be no direct benefit to the agency of doing so. If they really needed to use a standard format, wouldn't they already do so? The benefits are not directly seen by the agencies, but rather accrue to the city, agency, or department as a whole and to the individual recipient. The challenge of getting busy government managers to collaborate on a project that is not their own initiative and for which they are unlikely to see concrete results in the short term should not be underestimated.

The challenge of providing decision makers with accurate and timely counts of COVID-19 cases and deaths has brought this lack of data standards in public health into the spotlight. Some have proposed creating a working group to standardize public health data collection and facilitate interoperability among city, state, county, and federal government.

Another complicating factor in standardizing customer information is that for different agencies of government, the same individual can have vastly different customer characteristics. Consider a hypothetical woman—to the motor vehicle agency she's a driver, to the revenue department she's a taxpayer, to the small business department she's a business owner, and to the school department she's a mother. Each role may have vastly different attributes and permissions in the underlying systems.

The federal Commission on Evidence-Based Policymaking recognized this data standardization challenge in its 2017 final report and recommended the creation of a National Secure Data Service. If this resource were to be established, it would constitute an invaluable resource for intergovernmental data exchange, enabling a secure common platform for data sharing across federal programs and would allow both credentialed researchers and federal employees to conduct research and data analytics projects.

28. <https://www.nytimes.com/2020/07/13/upshot/coronavirus-response-fax-machines.html>.

29. Sam Edelstein, "Why Being a Chief Data Officer is Like Running a Minor League Baseball Team," June 3, 2019, <https://medium.com/@samedelstein/why-being-a-chief-data-officer-is-like-running-a-minor-league-baseball-team-745ccb478dc8>.

HOW INTERGOVERNMENTAL DATA STANDARDS LED TO IMPROVED URBAN MOBILITY

Few public policy areas could be more fundamentally intergovernmental than safe and efficient travel on the roads. The federal government funds a great deal of transportation infrastructure but the operations and safety enforcement functions of roads and transit systems are run locally. The challenge of making transportation safer for vehicles, pedestrians and cyclists requires a complex web of intergovernmental data collaboration. Add to this mix the new entrants such as ride hailing, bike sharing, and dockless scooters and the web becomes more challenging. Some mobility decisions are made by cities themselves, others by transportation agencies which may be regional and quasi-governmental.

When micro-mobility services unleashed shared-service bicycles, and then dockless electric scooters onto city streets, they disrupted more than just traffic. At that time, there were few if any public servants who knew exactly whose responsibility it was to regulate the safety of these new entrants to urban life. Should responsibility lie with those who regulate streets? Or sidewalks? Or parks? Or sharing economy services such as short-term housing rentals and ridesharing?

This disruptive innovation has unleashed a creative response from data-savvy individuals who have created a network for urban data innovators to create common standards and platforms and to share approaches to managing the services for the greatest public good. The Mobility Data Specification (MDS), rolled out in 2018, came about because of a need for data standards around dockless scooters, but the need for standards is wide-ranging in the mobility policy area. MDS demonstrates the value of building on existing efforts as it leveraged an existing standard for bike share data.

More than 80 cities are now using MDS to manage their micro-mobility vendors and to gather the data for local mobility policy decision making. The Open Mobility Foundation provides a forum for exchange of ideas, and hosts open source code and APIs to facilitate rapid replication and or transfer of success cases from one jurisdiction to another. This example demonstrates the power of disruptive innovation to inspire creative government solutions, and the value of networks in sharing success cases.

Federal leadership is also a factor in accelerating the sharing of data across mobility platforms. The U. S. Department of Transportation (DOT) is facilitating local adoption through a variety of research efforts and public private partnerships. For example, the public-private Work Zone Data Initiative for the first time sets a data standard that allows sharing of road closure data across platforms to improve safety and improve efficiency of travel. To advance open use and sharing of data and analytic insights, DOT created a Secure Data Commons (SDC). The SDC enables cloud-based collaborative and controlled integration and analysis of research data and provides privacy protections for personally identifiable information (PII) and confidential business information (CBI). Only approved users can access the data and they must be approved by DOT to do so. Nationwide Waze data is stored in the SDC and can be made available to researchers for work such as examining incident response times and use of incident frequency as a proxy for traffic volume. Analysis of the Waze data must be done within the Secure Data Commons and researchers may not export raw data from the SDC. Data analysis products, such as model results, figures, and tables can be exported from the SDC and shared publicly, subject to the approval of the DOT and Waze.



Challenge 7: Fears About Legal Authority to Share

There are both real and perceived legal challenges to data sharing. Both can be addressed with patience and persistence. For example, the federally-developed College Scorecard website combines student performance data, which is subject to Family Education Rights and Privacy Act (FERPA), and their wage and income data from protected IRS files. With thoughtful and patient work, the data sources were connected in a way that protects privacy of the underlying individual data but presents aggregate results in a way that provides insight.

Another example where a barrier was addressed is informing mental health or substance use disorder service providers when one of their patients has died from a drug overdose. To anyone not in government, this may seem like an obvious reason to share data, but doing this in Allegheny County, Pennsylvania, was not simple. According to Erin Dalton, deputy director for Allegheny County's Department of Human Services, "It seems straightforward, but there have been, historically and funding-wise, real walls between mental health and substance services." In some cases, statutory restrictions on data sharing had to be addressed, and in other instances the resistance to sharing data was rooted in fear and status quo bias rather than a true legal prohibition against data sharing.

Because the county's data warehouse includes human services and health data, there are legal requirements related to protected health information (PHI) and other data subject to privacy and confidentiality laws such as the Health Insurance Portability and Accountability Act (HIPAA). Data sharing agreements were not required for data sharing among county departments, because they all fall under the supervision of the county executive and thus are part of the same organization.

When the county sought to share data on the 20,000 children in the Pittsburgh Public School District (PPS), the data sharing agreement took 18 months to negotiate. The county was eager to connect data on child welfare and family homelessness with school attendance and to develop ways to increase school stability for children at risk of falling behind. School systems are often protective of their data, so the Memorandum of Understanding (MOU) between the county and PPS carefully outlines the roles, duties, and responsibilities of each party and specifies the governance, compliance, and privacy safeguards that are put in place to protect student data.

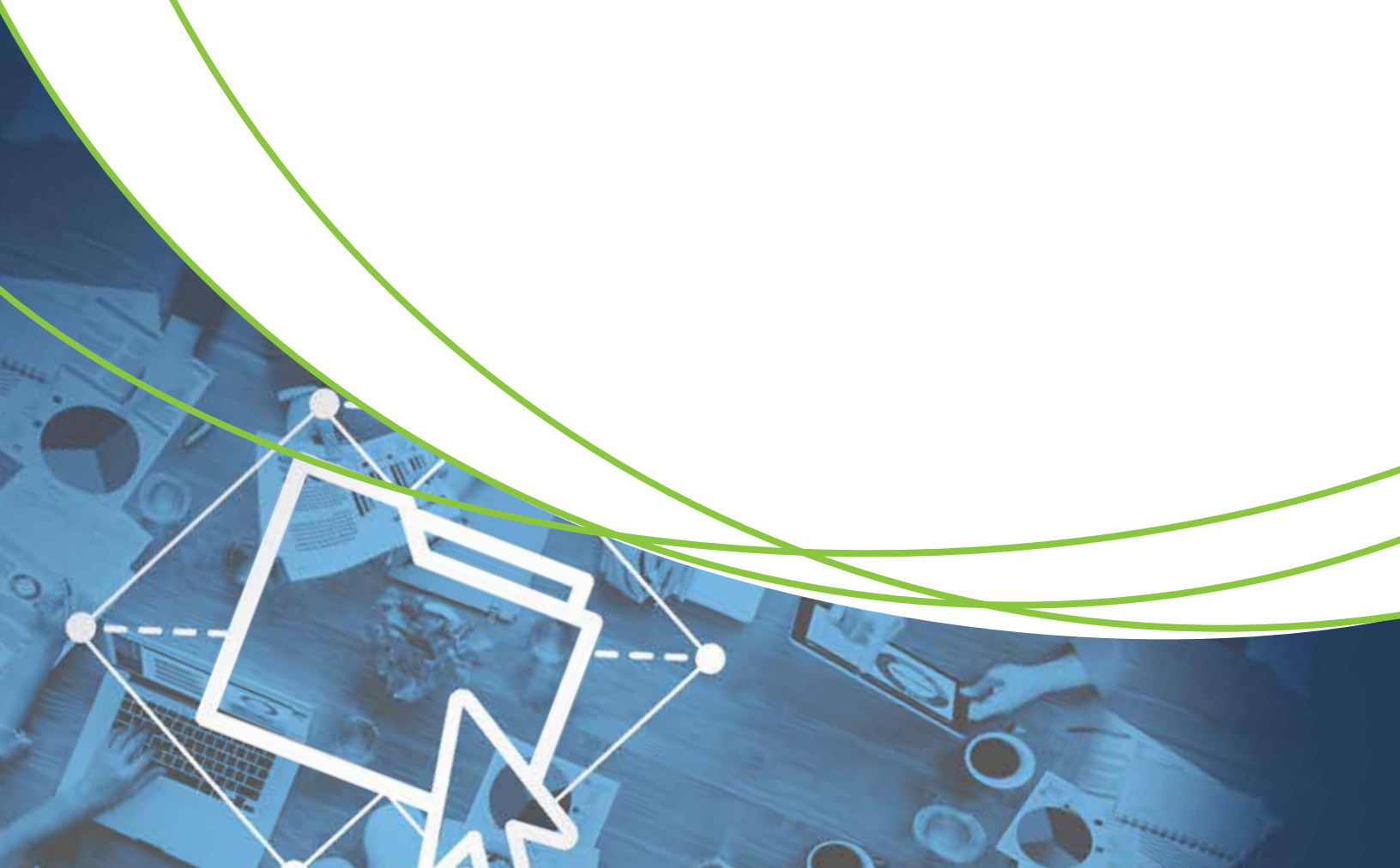
The Government Accountability Office (GAO) examined the challenges of data sharing in government with case studies and with a survey of stakeholders. Survey respondents were often confused about privacy rules, with 91 percent noting an extreme or great challenge with "confusion or misperceptions around what agencies are or are not allowed to share." Many survey respondents indicated that agencies may cite privacy requirements in general as a reason not to share data, without actually knowing the scope of the specific privacy laws at issue. With overlapping layers of state and federal law, some were legitimately confused, with one respondent noting, "It is challenging to understand what information may be shared with whom when four or five confidentiality laws may be applicable at the same time." Some survey respondents commented that their agencies may decide not to share data based on advice from their legal departments, "which may be based on a desire to minimize the risks associated with client data being shared improperly rather than due to a prohibition in law or policy."³⁰

While these challenges can be overcome—for example when a crisis creates urgency for collaboration, or when a disruptive technology compels government innovation—the barriers government data leaders face are significant. An exploration of factors that can accelerate data sharing and build momentum in the face of these challenges is discussed elsewhere.³¹

30. US Government Accountability Office, Human Services: Sustained and Coordinated Efforts Could Facilitate Data Sharing While Protecting Privacy, GAO-13-106. February 2013, page 13.

31. Jane Wiseman. "Eight Keys to Accelerate Government Data Innovation," Blog Post. IBM Center for The Business of Government, December 10, 2020, <https://www.dayoneproject.org/post/harnessing-data-analytics-to-improve-the-lives-of-individuals-and-families-a-national-data-strategy>.

Vision of an Ideal Data Sharing Ecosystem



Federal government data expert Kathy Stack noted, “For officials working inside government agencies, it’s no secret that government programs continually underperform because they can’t talk to one another.”³²

What would it be like if government data systems could easily link to one another, if data was seamlessly connected across all of government? Public sector employees could do their work without paging through paper documents and toggling back and forth among different sources of records in various electronic formats, such as a mainframe, locally resident data and data stored in an application at their agency or another agency. Error checks could be built in and mistakes could be nonexistent, saving time, money, and frustration. Customers would receive timely accurate information from government and speedy service, either via self-service tools or with the assistance of government employees. Transactions would be streamlined, efficient, and simplified. For example:

A couple walks into City Hall to get a marriage license. The City Clerk’s office scans the barcode on their driver’s licenses and instantly has accurate name and address data for each of them. Then the Clerk issues a marriage license and asks a few more questions. Would the couple like the city to automatically process a last name change for either party? If so, it can be automatically sent to all city, state, and federal government agencies, saving the couple time and complication with the name change. Address change? Same simple process. New driver’s license with the updated information? Click, click, it’s on its way. Change in tax filing status? Also done with one mouse click. Government just made this happy occasion a little better by taking out the paperwork.

The data sharing examples described in this report demonstrate a step toward such a customer-oriented government, personalized and responsive government. The hypothetical above represents an entirely new way of serving the public that leverages both the power of data sharing on the back-end, and the front-end capability to make transactions easier on the public with highly personalized, digital services. A few examples demonstrate that digital innovation can provide fuel for data sharing efforts as well.

- **Digital tools enable rapid response to the COVID-19 pandemic.** In early 2020, the City of Austin, Texas, quickly stood up a digital tool for residents to schedule COVID-19 tests and receive health advice, an effort that required speedy intergovernmental data collaboration on the back end, linking city and county systems in a seamless front end for the resident. The city’s coronavirus response website³³ provides information and access to the mobile tool, along with visualizations of which neighborhoods are being hit the hardest by the virus. The tool asks residents experiencing symptoms to fill out an online form to determine if they’re eligible to receive a test. For those eligible for testing, they can schedule an appointment from within the same app and receive a QR code that gives them a secure and private way to check in for an appointment, and also receive their results. When an Austin resident tests positive for the virus, the city gathers and anonymizes the data and includes it on a map that shows both where there are concentrations of cases, and where there is hospital capacity to meet that need. The city has also created a chatbot on the city’s website to answer resident questions about the virus.

32. Kathy Stack, *Harnessing Data Analytics to Improve the Lives of Individuals and Families: A National Strategy*, July 12, 2020, <https://www.dayoneproject.org/post/harnessing-data-analytics-to-improve-the-lives-of-individuals-and-families-a-national-data-strategy>.

33. <http://www.austintexas.gov/COVID19>.

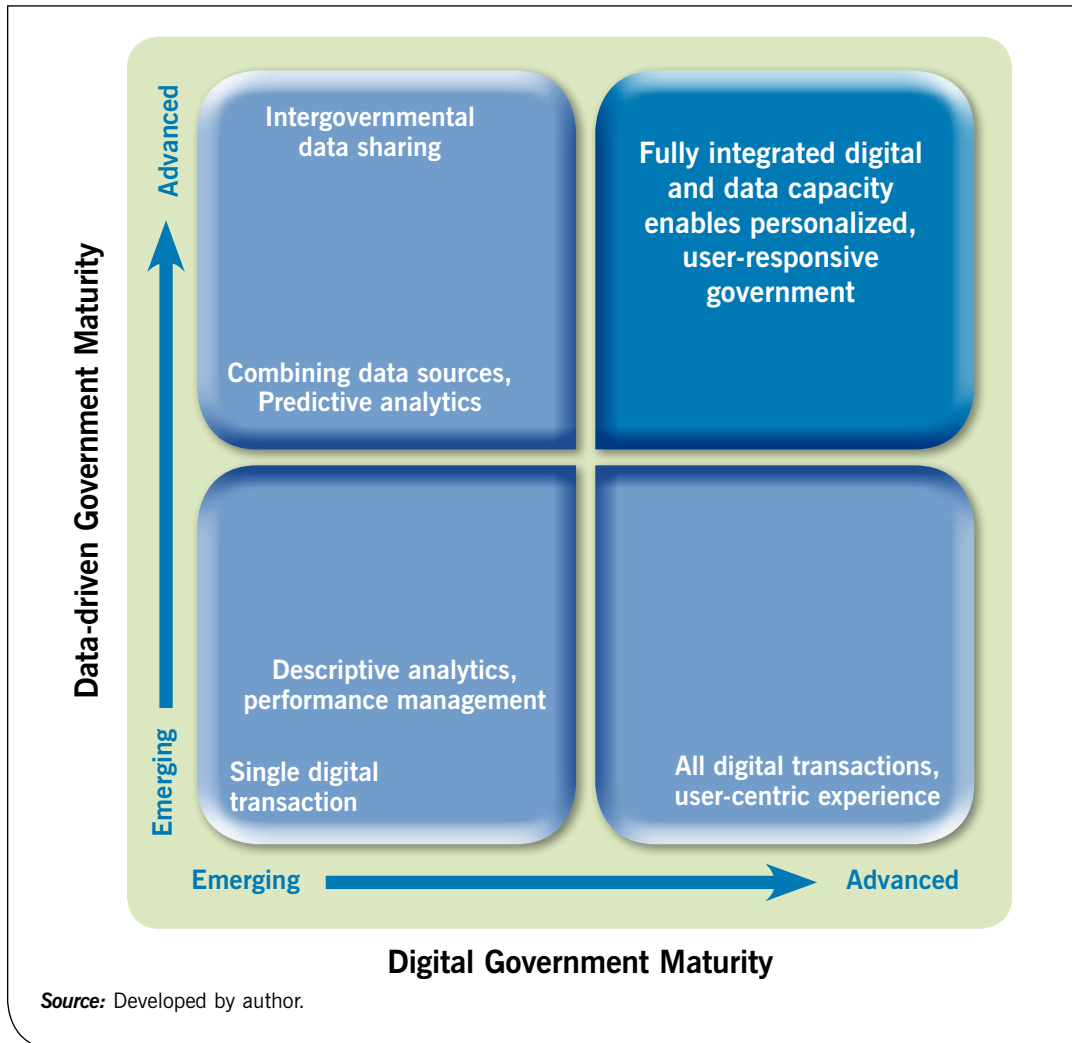
- **Single sign-on links government services on the back end for a smoother user experience.**
The federal Login.gov effort involved creating a single place that users could log on for personalized services on a variety of federal web sites. Cities and states have also adopted single sign-on. The State of Indiana developed a single sign-on portal called Access Indiana³⁴ which connects residents to 14 services that don't require a separate login, enabling the public to access services from a variety of agencies without knowing the organizational structure of the state, and simply being guided by their interests. Multiple, related services are bundled for the user, with back-end connections created that the user never has to see. According to chief administrative officer in Indiana's Office of Technology, Robert Paglia, "If you go in and renew a boating license, we can promote Indiana parks, or where you can book a cabin stay, or we can direct you to renew your fishing license."³⁵ Eventually, all state services that require login and password will be connected to the same single sign-on system. Not only will this make it easier on the public, but it will also save time for all of the individual department IT professionals who need to reset forgotten passwords for the many disparate systems that a user may only access infrequently. Eventually, voice-enabled digital assistant applications will be incorporated as well, with a seamless back end of integrated data sharing and processing.

Increasing maturity of digital government services, as more transactions are moved from paper to electronic increases the availability of data. If data can be fed back into analytic systems, it can then be mined for insight to drive more customer-responsive government services. This can create a virtuous upward cycle of better public services more aligned with customer needs—thus improving efficiency and customer satisfaction simultaneously.

There are ample examples of excellence in data-driven decision making in government, such as with the advance of the role of chief data officer and analytics teams in government across all levels. Digital government, too, has advanced significantly in the past decade, with the U.S. Digital Service and 18F (the federal technology and digital services provider within the General Services Administration) as well as a growing number of digital teams in city and state governments. Globally, Singapore, the United Kingdom, and Estonia stand out as exemplars of digital government service delivery. Yet, while there are outstanding examples of both data and digital, there are few government organizations that excel in both. As shown in Figure 3, pursuing initiatives for data-driven government and digital government in tandem can create the conditions for fully integrated and user-oriented government.

34. Access Indiana, 2020, <https://www.in.gov/inwp/2678.htm>.

35. Adam Stone, "Alexa and Siri are Resetting Public's Expectations for What a Good Experience with Government Could Be," *Governing*, September 20, 2019, <https://www.governing.com/news/headlines/Alexa-and-Siri-Are-Resetting-Publics-Expectations-for-What-a-Good-Experience-with-Government-Could-Be.html>.

Figure 3: While many governments excel at data use or digital services, few excel at both

Data and digital efforts pursued individually are powerful, but when pursued together they can be truly transformational for government. While the development of data and digital teams in government have both accelerated over the past decade, and in some governments the two roles are joined, it is important not to conflate these inter-related but separate functions. For leaders responsible for both data and digital efforts, it is helpful to remember that Clark Kent and Superman never did the same job at the same time. To be successful, integrated digital and data teams must have sufficient resources to perform both functions, and must have clear delineation of roles with strong leaders at the helm of each team.

SINGAPORE: AN EXAMPLE OF A GOVERNMENT THAT IS AT THE FOREFRONT OF BOTH DIGITAL AND DATA MATURITY

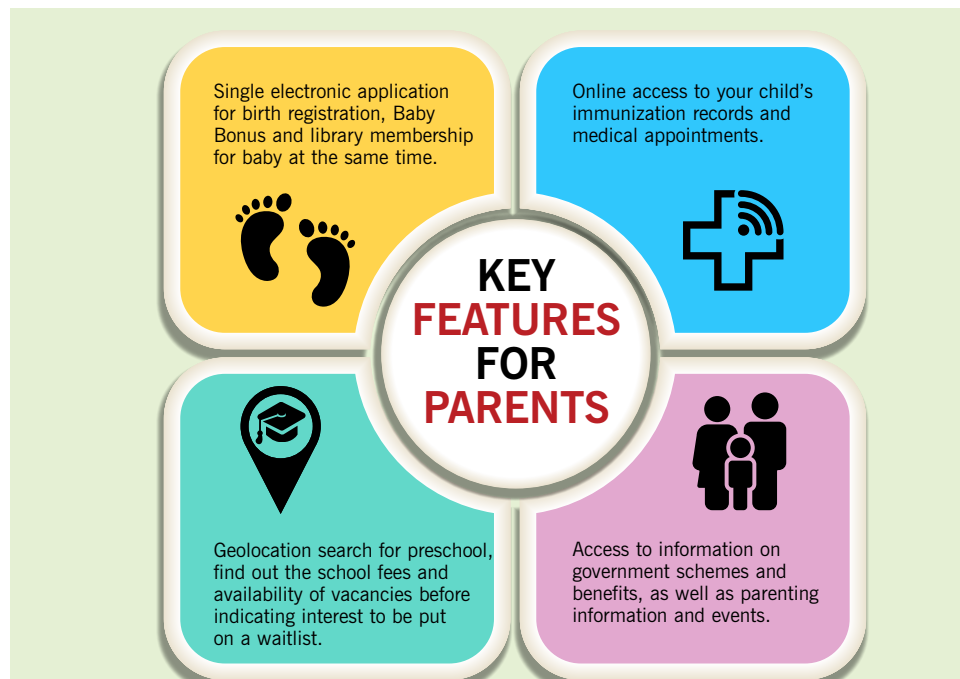
The Republic of Singapore invests in the long term for both public and private sector prosperity. It is a beacon for those interested in government and in particular in technology and data policy in the United States and around the world. Singapore is in many ways different from the United States, which has 50 times the population of this small city-state island, and has four times as long a history. Yet, Singapore while a very different society than the U.S., still offers a helpful point of reference.

Since its founding in 1965, the Republic of Singapore has been at the forefront of government innovation. This small city-state is bringing top talent and ideas like agile development and design thinking into everyday government operations. This relatively new and digital-savvy government puts the needs of the public at the center of public services, using data to connect across the “silos” of government to create a seamless digital government experience for the customer.

Singapore has made strategic investments in data and digital infrastructure. For example, since 2003, the public has been able to use secure single sign-on for government services. SingPass, or Singapore Personal Access has since 2003 enabled users to securely access over 300 digital government services from 110 government agencies. SingPass is now enabled for mobile devices, too. Extending the single sign-on to further simplify government transactions, MyInfo is a “Tell Us Once” service that pre-fills a digital form with authoritative personal data each time the user performs a transaction. Another example is digital payments—Singapore has long been moving toward cashless payments, and now has enabled digital payment for many government transactions.

The overarching leadership and vision for technology and innovation in data and digital services in Singapore is its Smart Nation and Digital Government Office (SNDGO). The office was created in 2017 to be a centralized hub of digital government activity as well as to build up digital infrastructure to serve the needs of citizens and businesses.

The Smart Nation team is charged with re-engineering government operations for efficiency, and with enhancing digital services to citizens and businesses. Throughout the work, they seek to bring in the voice of the public and to foster co-creation and engagement opportunities. This team has been working with various government agencies to create dozens of apps for public use, ranging from reporting municipal issues, crowdsourcing first responders for emergency situations, to parking. One app for seniors bundles services for them such as signing up for events that encourage active ageing and access information on government benefits. Another app bundles all services related to the birth of a new child, an exciting but also overwhelming experience. Now, this one app enables a new parent to complete birth registration, application for the government’s baby bonus payment, signup for day care and school registration, along with library membership application for young parents. As can be seen below, the resulting user interface is easy to navigate, and reflects the feedback of young parents during the development phase, including the use of eye tracker software so that app developers could literally “see” how people were interacting with the app.



Source: Singapore Government Technology Agency.³⁶

Perhaps nowhere is the government of Singapore's emphasis on technology and connecting across the silos of government to provide the citizen with a seamless experience more visible and valuable as in the response to the COVID-19 virus. Within days of the first news reports of the virus, public servants in Singapore began working on web sites and apps that would help provide the public with information to protect themselves against the virus. They weren't starting from scratch because there were already data and digital services staff in place, along with agile development processes. The government of Singapore made the decision to hire engineers and data scientists as government employees and to develop their own technology, rather than to contract out to private sector firms for this service. This has enabled more rapid response and has created a startup culture of experimentation within government. Referring to the decision a few years ago to bring more capacity for building apps inside government, GovTech Director Der Yao Leong noted, "With the COVID response, the app development team had to push out many new apps rapidly, and this has shown how valuable it is to have the engineering talent in house."

Government Data Office Director Quek Su Lynn described her organization's role as helping government "better harness the power of data to improve policies and services to citizens."³⁷ In addition to setting policy direction, her office oversees development of the infrastructure that enables efficient data flow across government with a goal of enabling data to flow easily between and among government agencies. "Single Sources of Truth" (SSOTs) are designated as authoritative sources for data elements that may be used by multiple government entities instead of collecting such data themselves—this ensures quality, consistency, and inter-operability of data across the government.³⁸

36. See its LifeSG app: <https://www.life.gov.sg>.

37. Medha Basu, "Meet the Women in GovTech 2019," *GovInsider*, December 17, 2019, <https://govinsider.asia/data/women-in-gov-tech-special-report-2019/>.

38. For a more detailed description of the data and digital innovation achievements of the Singapore government, refer to the case study produced during the course of research on this project, "July 2020, available at https://scholar.harvard.edu/files/janewiseman/files/engines_of_innovation_singapore_case_study.pdf which was completed as part of the research for this report.



Examples of Successful Data Sharing Initiatives

The government data sharing examples that follow demonstrate that significant public value can be created. However, creating the necessary supporting data infrastructure efforts take time, and that creating linkages across sources sometimes requires executive pressure, along with staff level persistence and creativity. Data platforms combine and compile information, and can either grant secure, role-based access to government staff such as a data warehouse does, or can provide access to the general public such as open data and performance dashboards do. Intergovernmental data sharing drives insight by comparing across more than one level of government or department, or expands beyond government. Federal, state, and local examples include:

- One-stop dashboard: U.S. Department of Education's College Scorecard
- Linked global data: U.S. State Department's repatriation of Americans' overseas during the COVID-19 crisis
- Linked education and wage data: state-level longitudinal data sets
- Linked cross-sector data: fighting the opioid crisis in Massachusetts
- Data modeling partnership: COVID-19 data in Texas
- Data integration across silos: improving justice system in Cook County, Illinois
- Data warehouse: intergovernmental human services in Allegheny County, Pennsylvania
- Integrated human services data systems: service delivery, policy research

Examples of Federal Data Sharing Initiatives

One-Stop Dashboard: U.S. Department of Education's College Scorecard.

The U.S. Department of Education created College Scorecard to help families learn out about and be able to compare similar data across schools when choosing a college. The tool brings together in a standardized format information about college course offerings, costs, admission rates, average test scores for students accepted, graduation rates, and average earnings upon graduation for different fields of study. Earnings data come from IRS tax records of students six and 10 years after college enrollment, and student loan repayment data come from the National Student Loan Data System, and shows for each college the median cumulative loan debt for graduates along with their loan repayment rates.

The mobile-friendly tool allows a user to search for colleges based on any of the criteria—for example, engineering programs in a specific region, or architectural programs below a certain size. The site also shares racial and ethnic information about students as well as the percentage of students receiving financial aid. Creating the site required compiling data across a large number of sources and bringing it into a common platform. From the site, a user can begin to fill out their college financial aid application to apply for federal grants or loans—integrating across yet another major government data system.

The tool aims to equip young people and their families with data to aid in what may well be the single most important financial investment, and one of the most significant personal and professional investments in the young person's career. The project involved collaboration from the Department of Education, the Department of the Treasury, the White House Domestic Policy Council, the Council of Economic Advisors, and Office of Management and Budget, the General Services Administration's 18F, and the U.S. Digital Service.

Linked Global Data: U.S. State Department's Repatriation of Americans Overseas During the COVID-19 Crisis. In January 2020, before the rest of country had yet changed our behavior due to the novel coronavirus, the data team at the State Department was bringing together

data from disparate public and private sources to create real-time information updates for department leadership on how to bring Americans home safely, first from Wuhan, China, and then from outposts around the globe. Under the leadership of Janice deGarmo, acting chief data officer, the data team quickly brought together all the data they could to help understand, monitor, and respond to the crisis, both from across the department and from external sources. Applying lessons from the Ebola outbreak and using both publicly available information, and their own on the ground intelligence, the State Department quickly created a series of data products to help their senior leaders take action based on timely, accurate data to protect U.S. citizens abroad—both State Department employees and others needing help in repatriation, or coming safely home during this crisis.

In January 2020, the work of the data team went from focusing broadly on the operational management of the department's workforce to laser-like focus on repatriation. This unprecedented effort led to the safe repatriation of over 100,000 Americans from more than 135 countries on over 1,000 flights working with embassies and consulates in every corner of the globe. "This success rests on the strong foundation of the data team's high-quality, trusted, and timely data and the expertise of our data scientists and leaders,"³⁹ deGarmo said.

This and other pandemic response data efforts at the State Department have required data not only from State Department and open-source data, but also included data coordination from across government, including the Centers for Disease Control and Prevention (CDC), the Department of Homeland Security, Customs and Border Protection, the White House Coronavirus Task Force, as well as private industry.

Examples of State Data Sharing Initiatives

Linked Education and Wage Data: State-Level Longitudinal Data Sets. Most states have a Statewide Longitudinal Data System (SLDS)⁴⁰ that links person-level data from preschool through elementary and secondary education, and to post-secondary education and work. The linkage from the start of a child's education to their adult wages can provide a detailed view of the individuals, programs and policies most likely to succeed, and can help identify groups most in need of support in order to become self-sufficient. Interesting and important policy questions could be answered by linking this data, such as:

- Which preschools best prepare students for kindergarten?
- Which schools, and which teachers best prepare children to read?
- Which students (and from which schools and teachers) most need remediation classes when they enroll in higher education?
- Which credentials or vocational programs provide the greatest return on investment for the student? For the taxpayer?
- How successful are college graduates in securing jobs and in staying in the workforce by type of degree or area of study?

Some states operate these systems out of their education department as is done in Massachusetts, other states host the system from a centralized statistical agency as in Kentucky, some states have a performance or budget office host the system as is done in Michigan, while other states have a university partner host the system, such as in Rhode Island.

39. Email correspondence, August 5, 2010.

40. <https://nces.ed.gov/programs/slds/>.

Federal funding from the Department of Education starting in 2005 gave states the latitude to design their own systems tailored to local needs. Since that time, \$826 million has been awarded to states over the course of six rounds of grants.⁴¹

Some have public facing web sites showing the research insights discovered from these linked data sets, as is done in Washington. Others allow individual access so that for example a parent can track a child's progress far beyond the latest report card, as is done in Georgia. This parent engagement tool has proved valuable. Robert Swiggum, chief information officer at the Georgia Department of Education, noted, "Imagine a 10th grader's parent coming in and seeing 11 years of history on that child. That kind of mutual access has really helped bring together conversations about the individual student's education."⁴²

Unfortunately, many of these state systems remain underleveraged, or unsuccessful. Implementation has been uneven. Some states are using their systems to innovate, fine-tune policy and empower local decisions right down to the classroom level. In other states, turnover from one administration to the next has meant that state leaders are no longer focused on using the data for insight. Some states received grant funds but were never able to successfully implement their systems due to typical technology challenges such as vendor management, scope creep, and unrealistic expectations. Some states faced more tactical challenges such as when attempts to use student performance data to assess teacher performance were scuttled by unions who advocated for decoupling student and teacher data.

One of the key issues with these systems is that they grew organically, as states requested federal grant dollars for specific projects in their state. There was limited effort at the federal level to mine best practice and then impose or incentivize adoption of those practices. There was no requirement for a long-term data use strategy at the start of the program, which could have given states a roadmap for how to turn data into insight and into action.

Another challenge with the systems is the disconnect between the state as the host of the system and the local schools and school districts as the creators of the data. Incentives and accountability could improve data quality at the input. Finally, an ongoing battle for states has been to secure sustainable funding for system operations.

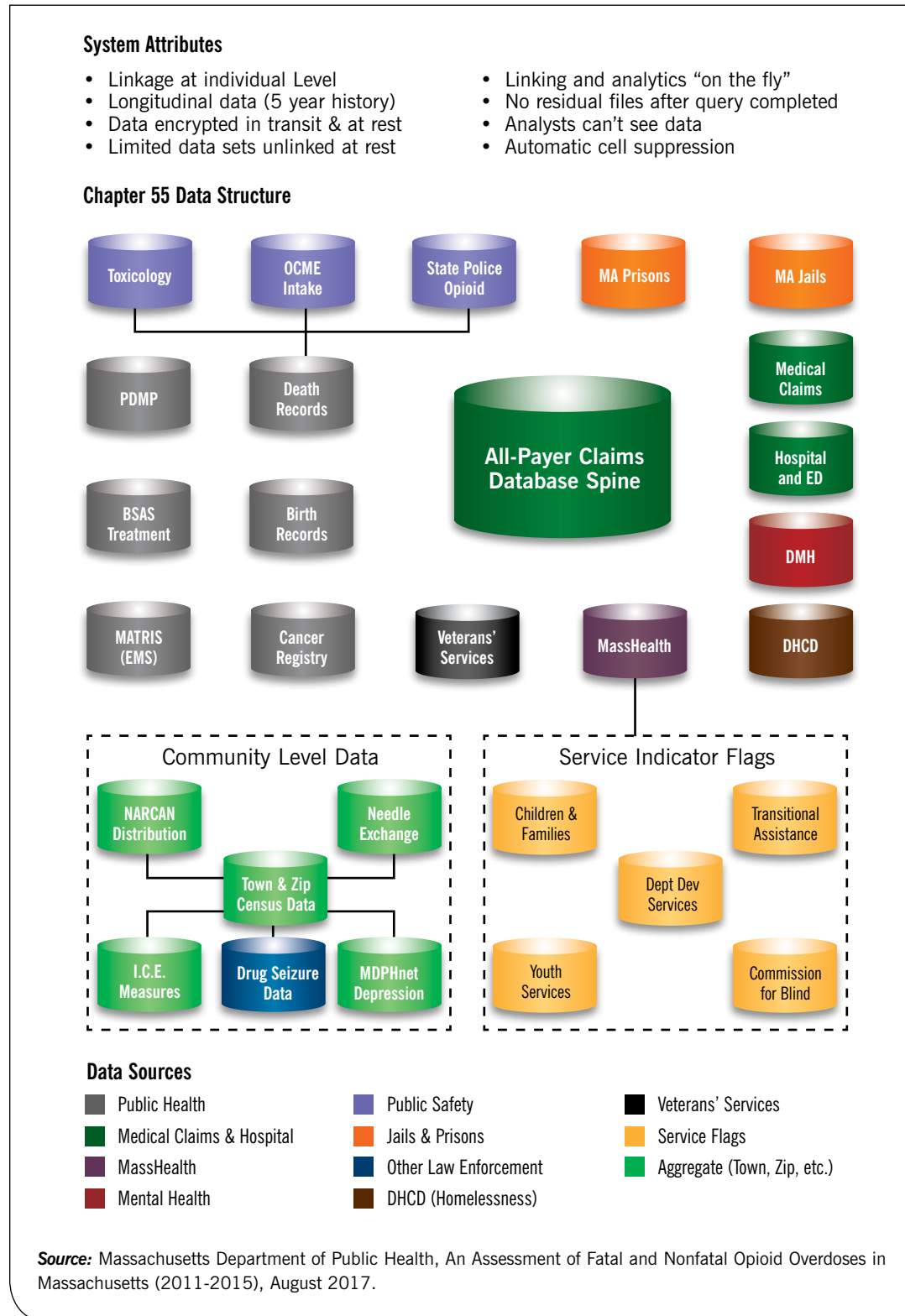
Linked Cross-sector Data: Fighting the Opioid Crisis in Massachusetts. The need to use data to address the opioid challenge in Massachusetts was well described by that state's Department of Public Health Commissioner Monica Bharel, MD, MPH, who said, "We collect a lot of data in government but what we end up with is a lot of data points, not actionable information." What she and her team did was ground-breaking work to find insight by mining many sources of data from over 20 public sector entities.

41. Institute of Education Sciences, "History of the SLDS Grant Program," April 2020, https://nces.ed.gov/programs/slids/pdf/History_of_the_SLDS_Grant_Program_Apr2020.pdf.

42. Jessica Leigh Brown, "4 Best Practices for Implementing State Longitudinal Data Systems," Education Dive, May 2, 2019, <https://www.educationdive.com/news/4-best-practices-for-implementing-state-longitudinal-data-systems/552355/>. Request to reuse quote granted by Swiggum.

The opioid data analysis drew insight from a database that linked mental health data, jail and prison data, vital records, substance addiction treatment data, ambulance encounter information, the state’s all-payer claims database, and others. Figure 4 shows how the many data sources came together:

Figure 4: Data Architecture for Massachusetts’s Opioid Integrated Data Sharing Initiative



The ability to conduct analysis on individual-level data from disparate data sets while protecting data privacy required innovation. Individual level data were temporarily linked and never stored in a common database, going back to their original files once the analysis was complete. Data was only temporarily linked and calculations were completed without the source data being available for download. High-level encryption assured that data transfers were secure.

The result satisfies federal HIPPA privacy requirements as well as the federal regulatory (42CFR part 2) requirements that provide additional protections to data about individuals receiving substance abuse services. As noted by Dr. Thomas Land, one of the key data analytics leaders on the project, “The protections we put on data went far beyond federal and state law.” This method took a great deal of planning and cooperation to get everyone comfortable and required bringing together lawyers, technologists, and data people from a very early stage of the process. Forging an innovative way to protect privacy while doing individual-level data matching took patience in working across disciplines with countless multi-hour meetings to hammer out all the details. As Dr. Land said, “There’s a way around every issue, and people had the patience to sit down and work it out. We didn’t get to zero risk, but we minimized risk and created better protections than anything else we could find.” At the best moments, the privacy protection and data access issues were dealt with in tandem, while the moments of challenge and difficulty were when the privacy protection or data access teams were working independently and sometimes at cross purposes.

No one agency had all of the necessary skills or resources, nor sufficient manpower for inter-agency data sharing. The most effective teams leverage diverse talents in new ways. For example, in Massachusetts the opioid data initiative required legal, data, technology, and analytic skills that were found in the Department of Public Health, the Center for Health Information and Analysis, the Executive Office of Technology Services and Security, along with data, legal and privacy experts from across the many agencies that shared data.

Data Modeling Partnership: COVID-19 Data in Texas. A partnership forged in early 2020 between Texas A&M University and the Texas Department of State Health Services facilitates access by university researchers to deidentified state data for use in developing models to forecast the spread of COVID-19. The hope is to map the spread in advance, identify anticipate hotspots, and plan for the allocation of necessary supplies and equipment in the locations that have the greatest need.

The modeling team includes faculty members with expertise in infectious disease, epidemiology, emergency management, public service, health policy, biology, economics and mathematics. The agreement with the state is instrumental in providing rapid data to the modeling team, facilitating near real-time review of and adjustments to the forecasts as the epidemic unfolded.

Olga Rodriguez—special advisor to the chief operating officer and senior vice president of the Health Science Center at Texas A&M University, and a key player in the effort to forge the memorandum of understanding (MOU) between the university and the state—had herself recently joined the university after serving for two decades in state government. This understanding of how state government works, and importantly how procurement and data sharing negotiations work from the state’s perspective, in addition to existing relationships of trust with key owners of the necessary data, significantly facilitated the development and approval of the MOU.

Rodriguez noted that the shared sense of urgency around the coronavirus, along with having the right people at the table from the beginning saved considerable time and effort and noted they were able to complete the MOU in about 10 days. As Rodriguez noted, “Everyone stepped up to the plate. Given the urgency, we all knew what elements of the MOU may necessitate compromise to get it done.”

Examples of Local Data Sharing Initiatives

Data Integration Across Silos: Improving Justice System in Cook County, Illinois. Cook County (which includes Chicago), is the nation's second largest county by population with over five million residents. It has achieved the ambitious goal of linking real-time data across multiple justice agencies spanning some of the most difficult "silos" of government to penetrate—the judiciary, corrections, prosecution and defense. The integration of data into an "enterprise service bus" enables decisions about release of defendants to be made based on a full review of their historical record so that the most dangerous can be detained while those who pose the least risk are released. An additional data integration enables sending automated text messages to remind defendants of court dates, which reduces inefficiency with lower no-show rates at court, and has been demonstrated in other jurisdictions to reduce no-shows by over a quarter.⁴³

Begun with a county resolution in 2002 to integrate its justice data systems, a 2005 strategic plan, and a \$2.3 million investment in 2015, this effort took significant effort not just on the technology side, but also to build relationships of trust across the agencies.

Strong executive leadership and persistent attention were keys to the eventual success of this initiative. County President Toni Preckwinkle, whose attention to this priority never wavered, was key to the success of the effort.

Data Warehouse: Intergovernmental Human Services in Allegheny County, Pennsylvania. Allegheny County (which include Pittsburgh) hosts one of the nation's leading integrated data systems. It links human services data, and many related data sets that form a seamless flow of data, in order to improve public service to vulnerable individuals, spanning city, county, and state sources. This data warehouse includes data from over 20 different sources across the full range of its services provided from childhood to aging, along with other sources such as county jail data, school data, and overdose data. All of the data can be matched across the different sources on an individual basis, allowing a level of research and insight rare in the public sector.

The data warehouse was created to improve the quality of service delivery to clients by leveraging data to enable frontline workers to perform their jobs better, and to support executive and management decision making. There is a wide range of data in the data warehouse, including:

- Client demographic data such as name, social security number, date of birth, and address
- Service information such as the client's past and present services (available to the client or parent/guardian if the client is under 14 years old) and the cost of that service
- Service provider information such as the, location, type of provider, and the services they have delivered to clients.

Unique to this tool, from the start, it was also intended to be the foundation of data and an information resource for both internal research and also external transparency and accountability—providing public value both to direct users and to a range of other interested stakeholders.

Keys to success for this project were:

- Outside funding to incubate the project
- Strong leadership vision from Human Services Director Marc Cherna

43. Ideas42, "New Text Message Reminders for Summons Recipients Improves Attendance in Court and Dramatically Cuts Warrants," 2018, <https://www.ideas42.org/new-text-message-reminders-summons-recipients-improves-attendance-court-dramatically-cuts-warrants/>.

- The authority to consolidate data for a unified department
- A long-term vision and the ability to generate momentum from early wins
- A cooperative working relationship between operations leaders and legal staff that helped forge data sharing agreements

None of this multiyear success would have been possible without persistence and patience. For example, the data sharing agreement forged between the county and the Pittsburgh Public Schools took eighteen months to negotiate—a team with less focus on outcomes and public benefit might have given up rather than pursued the goal relentlessly.⁴⁴

Integrated Human Services Data Systems: Service Delivery, Policy Research. Over the past decade, a movement to combine social service data sets to provide a clearer picture of client needs has created several exemplary integrated data systems (IDS). The description of the Allegheny County Department of Human Services data warehouse above is an outstanding example of an integrated data system.

To give greater visibility into the interconnected and related complex needs of clients, IDS link individual-level data from multiple administrative record systems, creating a more detailed understanding of client service needs. Integrating data across systems allows a case worker to make an individualized service delivery plan for a client, and then monitor progress toward goals across programs. Some IDS allow the client to see their progress too, and some offer longitudinal and comparative data for policy makers. For some, a partnership with a researcher unlocks more insight on the data over time and across programs and outcome goals. Many IDS are operated by government agencies, but some are operated by universities in partnership with the government, under agreements that protect the privacy and security of client data.

The advancement of the notion of IDS and the success of specific case studies owes due credit to the University of Pennsylvania and the network created there to support these efforts, called Actionable Intelligence for Social Policy (AISP). AISP has played a significant role in advancing the cause of integrated data systems through its technical assistance and network building activities, and by publishing case studies of success. The power of this network demonstrates the value of peer sharing of data challenges and progress.

One of the exemplary AISP sites is the state of Washington, where the state's Department of Social and Health Services (DSHS), Research and Data Analysis Division is responsible for its integrated data system that spans health and human services programs. The data system enables both state government researchers and external researchers (with approval of an Institutional Review Board) to conduct research on integrated client data for the purposes of informing program managers, legislators, and the public. All research done using this system is made public, and its web site includes data visualizations, over 100 reports⁴⁵ and several topical dashboards.⁴⁶

One research project from the DSHS team showed a financial return on investment. Using data from the IDS, the state team created a predictive model to identify high-risk Medicare and Medicaid clients who could most benefit from additional support and integration of their care. A pilot project that assigned high-risk patients to a coordinated care approach saved the state six percent on Medicare costs, or \$21 million, when compared to the cost of the study's comparison group.⁴⁷

44. For a more detailed case study of this data sharing effort see: "Better Government Through Data: Using the Allegheny County Human Services Data Warehouse to Design More Effective Results," available at <https://scholar.harvard.edu/janewiseman/publications/better-government-through-data-using-allegheny-county-human-services-data>, which was completed as part of the research for this report.

45. <https://www.dshs.wa.gov/ffa/rda/research-reports>.

46. <https://www.dshs.wa.gov/ffa/research-and-data-analysis/dashboards>.

47. The Annie E. Casey Foundation, Using Integrated Data Systems to Improve Case Management and Develop Predictive Modeling Tools, 2017.



Four Success Factors for Sharing Intergovernmental Data



Regardless of how a data sharing effort is structured, and whether a government-led or collaborative model is chosen, or whether a hybrid of the two models is engaged, there are four key success factors that are common across the two models. The key elements relate to the people leading the effort, the teams managing the work, the process for completing the work, and the data itself. These four success factors apply to both of the models that are described above.



Success Factors Related to Leadership

- **A leader's long-term *vision* is key to success in the short term and in the long run.** The long-term value of an intergovernmental data sharing effort should be part of the conceptual design as that will enable it to be valuable both in the short and long terms.
- **Consistent executive *engagement* matters.** Without the leadership imperative in taking the risk to engage in data sharing, none of the exceptional work profiled here would have happened. For example, in Allegheny County, Human Services Director Marc Cherna personally oversaw the initial development of the data warehouse. In Massachusetts, the opioid data sharing project that spans over 20 entities was a high priority project for that state's governor and received his frequent oversight.
- **Clear purpose for the use of data can mitigate confusion and reduce obfuscation.** The purpose for using the data and the intended outcome has to be very, very clear. This can be articulated by the senior executive sponsor, or by the project leader, but it should be clear to all involved. It can't be a broad statement like "to improve health"—it has to be very specific, like "to develop an infectious disease spread model and forecast anticipated utilization of health services in Texas." Being that specific alleviates uncertainty and risk for all parties to a data sharing agreement. As noted by Olga Rodriguez, special advisor to the chief operating officer and senior vice president of the Texas A&M University Health Science Center: "Setting clear boundaries around the use of data can lower data sharing risks and speed the approval process."
- **Senior executives' *conspicuous* use of data can set a great example.** As seen in Allegheny County, the value of the human services data warehouse became visible when policy changes began to be enacted based on data insights. One of the keys to success in states that have had effective Statewide Longitudinal Data Systems is visible leadership at every level, from the state department of education to the school in creating a culture of data use. "Leaders themselves should use information from these systems to make clear to people how these data are relevant, that they're not abstract but actually affect people's lives," noted Claus von Zastrow, principal, Education Commission of the States.



Success Factors Related to the Team

- **Experienced staff often have credibility as well as valuable knowledge of both processes and workarounds.** While those new to government bring fresh ideas, often those who have been in government for a career know how to get things done. In Santa Monica, California, Chief Wellbeing Officer Julie Rusk had two decades of credibility in the human services agencies of her city when she took on the creation of a well-being index, so her ongoing relationships of trust helped open doors that might otherwise have been closed. Likewise, when data scientists at Texas A&M needed an MOU with the state public health agency for their COVID-19 data analytics project, it helped that their university had on the team someone with two decades of prior experience in state human services who had existing relationships of trust with some of the very people who would be involved in the data sharing MOU discussions.
- **Existing working relationships facilitate and ease communication during innovation.** Data sharing requires trust and that can often be accelerated when there are existing relationships forged by other projects. This is seen in the case studies profiled here as well as in the research literature. In describing the many ways the government of Bangladesh mobilized to provide digital services in the wake of the COVID-19 crisis, Anir Chowdhury, United Nations Development Programme (UNDP) policy advisor, described the work as “spontaneous,” but noted that they had been working together for 13 years before that.⁴⁸ He noted that most of the recent COVID work was based on prior relationships and that the trust built up over a decade of digital transformation efforts afforded some level of trust in many agencies.
- **When building new relationships, sometimes working together informally can build trust in advance of developing a formal data sharing agreement.** Sam Edelstein, former chief data officer for the City of Syracuse, New York, recommended starting with an informal data sharing project and building trust through small wins before moving to develop a formal data sharing agreement. He noted that the city and Syracuse University began a data sharing project that enabled students to use city data with university oversight of their work. After two years of informal data sharing, it was much easier to create a formal agreement as the roles and requirements became clear over time. As Edelstein said, “Start small, show wins and then formalize the agreement.” He also recommended seeking out and then working with people who are willing to say “yes” instead of “it’s impossible because of xyz regulation.”

48. Regional Innovation Centre UNDP Asia-Pacific, “The Innovation Dividend Podcast: How the Bangladesh government is repurposing assets for COVID-19”, medium.com, April 20, 2020, <https://medium.com/@undp.ri/how-the-bangladesh-government-is-repurposing-assets-for-covid-19-650c26019259>.



Success Factors Related to Processes

The process of establishing data sharing arrangements can be challenging, but in success cases challenges were mitigated. Data sharing efforts can disrupt the status quo and may require significant attention to change management and people issues throughout the implementation process. For example, the Allegheny County data warehouse team is proud of their data sharing agreement with the Pittsburgh Public Schools, but noted that it took 18 months to negotiate. The enablers of success relate to the process and data challenges that need to be addressed and the key strategies for tackling them. Regarding the process for establishing and maintaining a data sharing effort, there are several success factors related to the process of getting the work done:

- **The process can be long and may require *patience*.** Data sharing is complex and often not well understood. The Cook County, Illinois, integrated justice project took over a decade to be launched, as well as did the realization of true intergovernmental reach of the Allegheny County Human Services data warehouse. One expert noted the importance of “patience and persistence” in getting a data sharing agreement completed. Carlos Rivero, the chief data officer for the Commonwealth of Virginia, recognized the importance of patience: “Sharing data is a vague and ill-defined process which takes a lot of time and effort to be successful (even for non-sensitive data).” In the survey of state and local data sharing professionals conducted by the GAO, 60 percent of respondents to their survey noted establishing data sharing agreements was an extreme or great challenge. According to one survey respondent, “The process of getting agencies and individual entities (such as schools) on board and getting to data sharing agreements is so exhausting that the effort just dies.”⁴⁹
- **The mechanics of getting an MOU done can be complex and intimidating and without *persistence* may cause a project to stall or fail.** One helpful insight was shared by Olga Rodriguez, special advisor to the chief operating officer and senior vice president of the Texas A&M University Health Science Center. Again, Rodriguez helped her university forge a data sharing agreement with her state’s public health agency in about 10 days. Reflecting on the keys to their speedy negotiation and approval process, she noted, “You need to know about the approval process for an MOU in the organization and figure out who needs to be at the table and be sure you’ve got the right players from the start. You don’t want 23 people at the table, just the ones who need to be there, probably two or three from each organization. That way you can really streamline the negotiations. For our MOU we needed the contract administration, legal, and program experts from the state, and from the university, we needed the privacy and IT staff to be on board with the data elements and data transfer needed by our modeling team. If you work with the right people from the start, you save a lot of time in the approvals at the back end.”
- **Plan for ongoing *sustainability*.** Many data sharing efforts fail if there is leadership turnover or if operations resources are not secured. Bob Swiggum, chief information officer at the Georgia Department of Education, and leader of that state’s longitudinal data system, noted: “Whenever you’re designing the system, make sure you can afford to manage it and keep it running. If you can’t provide the training and continue to support the system with the needed infrastructure, don’t bother in the first place.”

49. U.S. Government Accountability Office, Human Services: Sustained and Coordinated Efforts Could Facilitate Data Sharing While Protecting Privacy, GAO-13-106. February 2013. P. 23.

- **Providing confidence in the privacy protections in place and that data will be used as agreed is a key to success for data sharing in government.** Generally, public servants want to protect the public's interest, and often that means protecting data too. Unfortunately, there is a history of risk-aversion in government that can make data sharing more difficult. Government employees are often conditioned to be very protective of their data, not because of ill will, but because of legal rules and attempting to stay out of trouble. The Government Accountability Office (GAO) examined the challenges of data sharing in government with case studies and with a survey of stakeholders.⁵⁰ Their survey demonstrated this tendency toward data protection. Survey respondents were often confused about privacy rules, with 91 percent noting an extreme or great challenge with "confusion or misperceptions around what agencies are or are not allowed to share." Many survey respondents indicated that agencies may cite privacy requirements in general as a reason not to share data, without actually knowing the scope of the specific privacy laws at issue.

With overlapping layers of state and federal law, some were legitimately confused, with one respondent to GAO's survey noting: "It is challenging to understand what information may be shared with whom when four or five confidentiality laws may be applicable at the same time." Some survey respondents commented that their agencies may decide not to share data based on advice from their legal departments, "which may be based on a desire to minimize the risks associated with client data being shared improperly rather than due to a prohibition in law or policy." One stakeholder said that "a great deal of practice is based on avoiding risk and taking what is assumed to be a safe course without any necessary connection to the original policies' requirements or intentions."



Success Factors Related to Data

The data gathered by government is fraught with incompleteness, inaccuracy, and other challenges, but these can be overcome with care. Data sharing across sources amplifies underlying data quality challenges in source systems and requires solid data governance. For example, most states don't actively audit their criminal history records, but when one state did, it found 18 percent of records were misclassified, and that up to 30 percent were misclassified at some law enforcement agencies. Success factors related to data are described below.

- **Data quality issues should be addressed at the point of data entry.** Data in source systems can have errors and those data quality issues can make it challenging to successfully link data across systems if the source data is incorrect. As one data expert said, "If it's not being collected accurately it doesn't matter what you do with it." For example, if a typo exists in the client name in one database, or the name is incomplete, that means "fuzzy matching" and other data management techniques are needed. The National Center for Health Statistics noted that even before the surge in deaths due to the novel coronavirus in 2020, one in three death certificates noted the wrong cause of death, citing the differences in local regulations about who can or does complete the certificate.⁵¹ Most states don't actively audit their criminal history records, but when one state did, they found 18 percent

50. U.S. Government Accountability Office, Human Services: Sustained and Coordinated Efforts Could Facilitate Data Sharing While Protecting Privacy, GAO-13-106. February 2013.

51. Jessica Priest, Coronavirus: 1 in 3 death certificates before virus were wrong, and it's getting worse, SouthCoast Today, April 28, 2020, <https://www.southcoasttoday.com/news/20200428/strongcoronavirus-1-in-3-death-certificates-before-virus-were-wrong-and-its-getting-worsestrong>.

of records were misclassified, and that up to 30 percent were misclassified at some law enforcement agencies. Further, in a national survey of state criminal history records, only 68 percent of them included disposition data, with some states having as few as 10 percent of their records including this essential information. If data are missing or wrong, any analysis of the data will be misleading and could result in research findings and or policy decisions that are misguided. And it's not only government that has data quality issues. Aaron Klein of the Brookings Institution pointed out: "More than one in five consumers have a 'potentially material error' in their credit file that makes them look riskier than they are" to lenders."⁵²

- **Data quality issues are amplified when the data is shared.** Data collected in one operational system for one purpose may be later used for another purpose either by a practitioner or a researcher. Even where the accuracy is improved by the secondary users, the source system record is unlikely to be updated or improved, meaning that quality remains low often even when the data is used. Where data quality is better, it is typically uneven across system components limiting cross-agency sharing, or even within the same organization. For example, data collection standards and metadata may vary even within a police department between their records management and dispatch systems, limiting the ability to mine across systems for patterns and insight.
- **Lack of data standards makes data sharing more complicated.** Local differences undermine national data consistency. Much of the data collected for national reporting purposes is collected in local schools, courthouses, and city clerk's offices. The lack of consistent methods and tools means that in some cases data is collected electronically, and in far too many cases it is still collected on paper. Further, with over 19,500 cities and towns and nearly another 69,500 other units of local government, the myriad of ways of interpreting terms and definitions makes the rolled-up data potentially wildly inconsistent. As an example of this challenge, researchers trying to assess the feasibility of using linked administrative data to measure the national incidence of youth who had contact with both the foster care and juvenile justice systems, found that they were ultimately not able to attempt a national measure because they could not "apply the same definitions of system involvement because juvenile justice and child welfare practices vary widely across each jurisdiction."⁵³
- **Data governance is often underappreciated but is critical to success.** Data governance is critical to success of data sharing efforts but is seldom given sufficient attention. As noted by Natalie Evans Harris, "While we've seen many success stories in deploying data sharing collaboratives, many government agencies are also experiencing the pains of data sharing governance practices that are limited and under-resourced, resulting in practices that can be unsustainable, not effective, and not forward-thinking."⁵⁴ Evan White succinctly described the keys to effective data governance: "Successful governance processes empower subject-matter experts to review project proposals based on well-defined technical criteria, and include provisions—such as defaulting to approval if a proposal is not affirmatively rejected within a designated timeframe—to ensure that the approval pipeline does not become a project graveyard. Another aspect of successful processes is that they separate concerns over privacy and data security from those about the substantive content."⁵⁵

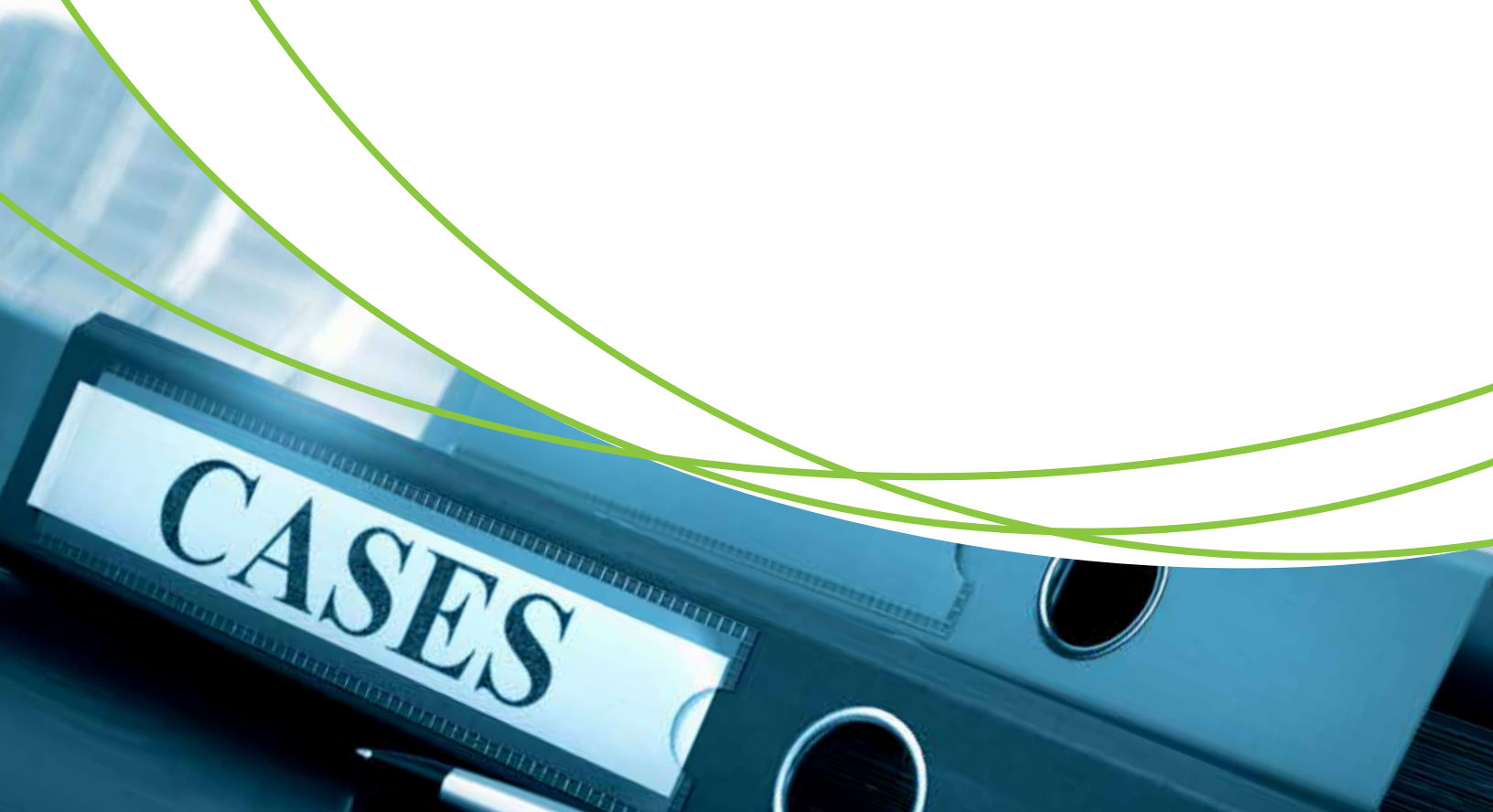
52. Charlie Warzel, All This Dystopia, and for What?, New York Times, February 18, 2020, <https://www.nytimes.com/2020/02/18/opinion/facial-recognition-surveillance-privacy.html>.

53. Denise C. Herz, Ph.D. and Carly B. Dierkhising, OJJDP Dual System Youth Design Study, p134, National Criminal Justice Reference Service, March 2019, <https://www.ncjrs.gov/pdffiles1/ojjdp/grants/252717.pdf>.

54. Natalie Evans Harris, "Sharing Data for Social Impact: Guidebook to Establishing Responsible Governance Practices," Beeck Center for Social Impact + Innovation at Georgetown University, January 2020, <https://beeckcenter.georgetown.edu/wp-content/uploads/2020/01/Data-Sharing-Report.pdf>.

55. Evan White. "A Roadmap for Linking Administrative Data in California," UC Berkeley, California Policy Lab, May 7, 2019.

Recommendations



The following recommendations to advance intergovernmental data sharing are based on findings from the relevant literature, expert interviews, and the case studies and examples described here. While these recommendations are targeted to the federal government, they apply equally to legislators, executives, and data leaders at the state and local government levels as well.

Recommendation One

Create a policy and governance framework. Congress and the president should define a broad data and digital excellence vision, with incentives to act and a strong data governance infrastructure. This would include the following actions:

- Action 1: Establish an “ask once” goal for government data collection, and require agencies to create a strategy to gradually achieve it, subject to privacy constraints.** With state and local open data policies, and with the federal Digital Accountability and Transparency Act, when government data policies have gone from “closed by default” to “open by default”, the volume of open data published by public agencies increased dramatically. Similarly, if legislatures at all levels compel government agencies to default to not asking the same individual for the same information multiple times, a good deal of redundant data collection could be avoided and time saved. In Estonia, where an “ask once” policy has been in place for many years, and with the government’s first digital service⁵⁶ going online in 2000, it now takes only three minutes⁵⁷ to file taxes online, and 47 percent of the population uses digital voting. The estimated benefit is significant with 800 years⁵⁸ of working time saved each year, and an estimated two percent of GDP⁵⁹ saved due to the use of digital signatures. While the U.S. is far larger and more diverse and complex than Estonia, with its 1.3 million people, this model of integrating government data for the convenience, security, and privacy of the public is a model worth examining. Legislators should mandate U.S. government agencies to set the goal of not asking the public for information it already has on hand, or can get from a peer government agency.
- Action 2: Mandate that government calculate the burden on the public of redundant data collection.** The Paperwork Reduction Act requires that federal agencies complete a structured process to calculate the level of “burden hours” imposed on the public and businesses before asking them to provide information in surveys and other data collection tools. Yet, there is no parallel requirement to assess the burden of the failure of government to use or share data that it already possesses. When an agency asks for the same information twice, or when one agency asks for information that another agency has, the public has no choice but to repeatedly submit the information. A new resident in a town may need to go to the library, school district, pet registration, and voter registration offices to establish residency, and depending on the services offered in the town, may also have to separately register for utilities and for trash disposal. If they want to renovate their home, there will be an entirely new set of forms and procedures from the planning department, the building permits, and electrical and plumbing work approvals. If federal, state, and local government each individually commit to viewing the customer’s time as valuable and places the burden on government not the individual to duplicate data, significant efficiencies could be gained. A longer-term solution would be to interconnect across the levels of government, but a good starting point would be to require each level of government to take responsibility for acknowledging and accounting for the burden placed on the public of

56. E-Estonia.com, Frequently Asked Questions: Story of e-Estonia, March 2020, <https://e-estonia.com/wp-content/uploads/2020mar-faq-story-of-e-estonia-1.pdf>.

57. E-Estonia.com, e-Estonia facts, April 15, 2020, <https://e-estonia.com/wp-content/uploads/e-estonia-facts-20-04-15.pdf>.

58. E-Estonia.com, e-Estonia Guide, April 4, 2018, <https://e-estonia.com/wp-content/uploads/eas-eestonia-vihik-a5-180404-view.pdf>.

59. Ibid.

providing duplicate data. The next step would be to create a strategy for sharing data internally to the greatest degree possible without compromising privacy. Public input to the design of such data sharing would be key to success.

- **Action 3: Require or reward linking administrative datasets.** As an incremental step toward achieving an “ask once” policy, government agencies could begin linking their existing data. Many different government programs often touch the very same people’s lives without knowing it; the unemployment office, the child welfare, public housing, and the mental health agencies may all be serving the same clients without knowing and without sharing data. Data sharing could improve the personalization and quality of service delivery. The federal government could require that data be shared across programs serving vulnerable populations, perhaps as part of the response to the COVID-19 economic crisis. Either by requiring linkage of administrative data sets, or by rewarding those agencies that do provide linkages, the barriers of inertia and lack of local leadership can be overcome in ways that are not confined to bureaucratic silos.
- **Action 4: Establish intergovernmental data councils.** As noted by government data expert Kathy Stack, “It is currently no one’s job in the federal government to understand the challenges that state and local governments face in harnessing data, analytics, and evaluation to improve the impact of funding they receive from hundreds of federal grant programs.”⁶⁰ A major source of insight for federal agencies would be to listen to the data needs and challenges of state and local government, by convening intergovernmental data councils. There is a precedent—the Department of Transportation reaches out to state and local governments in an ongoing data exchange discussion. Formalized data councils could provide a mechanism for feedback on data standards which would make data more interoperable across the levels of government.
- **Action 5: Establish rules or deadlines for standard data sharing and use agreements.** When legislation requires it, data sharing is accelerated significantly. For example, Massachusetts broke down information-sharing silos after legislation called for data sharing to solve the opioid epidemic. At the start of the state’s opioid data sharing initiative, there were 78 different types of data sharing agreements in use in departments across the state and the average time to reach agreement to share data between two agencies was 133 days. Inspired by the opioid project, a working group of executives across departments together developed a standard Data Use Licensing Agreement (DULA) that can be used and adapted for any agency in the state for sharing, protecting, and securing data. Establishing this mechanism for sharing data will likely have benefit for years to come in ways not even yet imagined as state agencies and departments open up new avenues of data collaboration. Having the legislative deadline provided a significant incentive to finish the work on time, and provided urgency and focus.



Recommendation Two

Congress and the president should establish funding and capacity building mechanisms to support implementation of increased data sharing across government. This would include the following actions:

- **Action 6: Adequately fund data and digital infrastructure.** Effective data sharing requires investments in hardware, software, and network infrastructure to support the movement of data across departments and its analysis and storage. Funding common platforms reduces one barrier to sharing. At all levels of government, policy makers should not only fund the infrastructure (a data warehouse for example) but should also fund the necessary data

60. Kathy Stack, “Harnessing Data Analytics to Improve the Lives of Individuals and Families: A National Strategy,” July 12, 2020, <https://www.dayoneproject.org/post/harnessing-data-analytics-to-improve-the-lives-of-individuals-and-families-a-national-data-strategy>.

governance and management, along with automation of common processes.

- **Action 7: Provide ongoing support funding for data sharing projects.** Simply mandating data linkage may initiate the connection but will not sustain it. Without ongoing support funding, some programs may founder, as was seen in some states' Statewide Longitudinal Data Systems created by the U.S. Department of Education. Given that leadership turnover can happen at the state and local level, a federal funding mechanism for ongoing support for integrated data systems would be optimal.
- **Action 8: Support data literacy in federal agencies and among federal leaders.** Recent research from Fortune 1000 companies shows that while 99 percent of firms are investing in big data and artificial intelligence, only 27 percent are achieving culture change, with one of the key gaps being executive leadership data literacy.⁶¹ One study found that 53 percent of IT and data leaders think they are "too old" to learn about data.⁶² For government to turn the tide and create a culture of data-driven decision making, data literacy will need to become a priority, and the federal government can build tools and resources that can be used by state and local governments as well. While a handful of leading cities have already developed data academies to increase the skill levels of staff, the majority of government agencies lack a comprehensive professional development program to create a generation of managers who understand and embrace the power of data. This is particularly true for senior and mid-level executives who are largely without available skill development options as most current training is tool-based learning for analysts, not decision-making training for executives. One excellent example of addressing this challenge is in Pittsburgh where the city's director of Innovation and Performance, Santiago Garces, is building momentum on this cultural shift toward greater adoption of data use by providing executive level training on how to use data.
- **Action 9: Provide resources to improve data quality.** The lack of incentive at the front end to collect and record data accurately is often tied to lack of resources, lack of training, and a disconnect between those who enter data into operational systems and those who use the data to make decisions. Data quality and availability in rural areas is less studied and may be prone to greater gaps and inconsistencies. As one data expert noted, "Data collection will be better when people see a reason to get it right." Funding dedicated to improving data quality could include automation tools, training resources, auditing, and feedback loops as well as personnel.



Recommendation Three

The nonprofit and philanthropic sectors need to proactively support intergovernmental data sharing efforts. Philanthropy has played a significant role in many of the successes to date and has inspired a wide range of data sharing innovations, including several of those described in this report. Continued leadership would include the following actions:

- **Action 10: Provide incentives to innovate and link different sources or types of data at the state and local levels.** Philanthropy can spur innovation by providing incentives to link administrative and survey data with other sources inside and outside of government. Many examples exist already including the groundbreaking work of the City of Santa Monica in creating a well-being index, and the use of multiple levels of government data along with external sources in mobility data sharing efforts. More collaboration across entities could include the U.S. statistical agencies as well as universities, civic, and media journalism sources.

61. NewVantage Partners, "Big Data and AI Executive Survey 2020: Executive Summary of Findings," 2020.

62. Splunk, "The State of Dark Data," Slide 18, 2019.

- **Action 11: Fund the documentation and sharing of data sharing best practices.** Data leaders have relied on networks of peers for idea sharing and for the spread of innovation. The Harvard-based Civic Analytics Network is a leading example, where two innovations spread widely among network members. The Data Academy launched in San Francisco has now been adapted in many other network cities, and the data science use case process pioneered in New Orleans is also widely shared among members. With the opportunity for in-person peer sharing diminished in the near future due to COVID travel restrictions, and to promote replication beyond the confines of peer networks, data leaders should document their successes in extensible and sharable repositories including blog posts, via presentations and webinars and where possible by developing repeatable tools and templates for others to adopt.
- **Action 12: Fund replication tools and information exchange networks.** Success cases exist, but for the typical busy government executive, a limiting factor is the time it takes to learn about and determine how to replicate a success found elsewhere. Funding could be provided for academics and other outsiders to support government staff in the documentation of best practices along with multimedia tools, templates, and technical assistance to create replication strategies and implement them would help spread success cases.



Recommendation Four

Agency managers and data leaders at all levels of government need to champion data sharing efforts. The recommended actions apply to all who are working with data at all levels, from chief data officers and other data leaders to the data scientists, data analysts, and data engineers in the public sector, as well as data entrepreneurs in the civic tech community. This recommendation would include the following actions:

- **Action 13: Articulate and share a clear vision for data sharing and its value.** Intergovernmental data sharing is hard, and every participant will at times be challenged to stay committed. Having a leader who articulates and shares a clear and inspiring vision can provide common cause. When the goal is agreed and understood by all it helps provide focus and clarity. In addition to executive level vision, a deadline also helps but is not required. In the case of the Massachusetts opioid data sharing effort, a deadline imposed by the legislature to answer seven key policy questions with data gave structure to the data sharing work.
- **Action 14: Establish and widely share data standards and protocols.** Data standards can make interoperability easier, particularly when data standards transcend one organization. Within an organization, a chief data officer (CDO) can help advance standardization of data. Former CDO of the General Services Administration, Kris Rowley, noted, “When it comes to data definitions, data standards, and business rules, you cannot define those enough. And when there is a conflicting data standard, the best thing to do is to articulate the conflict and why it’s a problem. Then, the CDO should facilitate the process of getting to consensus on the data standard rather than trying to impose a standard on others, because that can backfire.” Regarding data standards that extend across organizations, there are several examples of collaboratively-developed data standards, such as the Mobility Data Specification (MDS), the General Bikeshare Feed Specification (GBFS), Work Zone Data Exchange (WZDx), the National Open Court Data Standards (NODS), and the National Information Exchange Model (NIEM). While these examples demonstrate what is possible, they each remain far from ubiquitous in their adoption.
- **Action 15: Be patient as it takes time.** Janey Rountree, executive director of the California Policy Lab at UCLA pointed out: “Data sharing is very personal, it’s about building relationships of trust and that takes time. It took us two and a half years to create the data sharing agreement between the Policy Lab and Los Angeles County. But now that it’s in place it’s a powerful foundation for analytics and policy insight on a range of topics, so it was well

worth the effort.” Often, success comes from building incrementally and leveraging what already exists. Small wins build not only momentum for larger ones, they also inspire loyalty. For example, the State of Georgia’s Statewide Longitudinal Data System (SLDS), started as a basic platform for standardizing data received from school districts to the state, and was incrementally built over time. Unlike some SLDS efforts that attempted to launch ambitious systems from scratch, Georgia built on what it already had, such as an existing unique identifier for students and a district-level student information system, and then built a data sharing exchange, and gradually created increased functionality, to the point where now it is among the most sophisticated of all state SLDS systems.

- **Action 16: Create communities of practice for data enthusiasts.** The power of networks to support and enhance innovation and its spread is widely documented.⁶³ In every city, county, state, or federal agency, data leaders should create communities of practice for the exchange of ideas and to support mutual ideation. The General Services Administration, under the tenure of its first chief data officer, Kris Rowley, was an early leader in the federal government in creating this capacity and advancing the government’s data culture with both structured and unstructured community of practice events. In Los Angeles, the Citywide Data Collaborative was created to bring together people across departments who don’t otherwise interact but who have a common interest in data, from the airports to pensions or parks departments. Once the collaborative was under way, with both formal and informal ways of engaging, the group was expanded to include both the school district and the county. When the school district hired a data officer in 2018, he found this a valuable resource for data and ideas. Sari Ladin-Sienne, former chief data officer for the City of Los Angeles commented on the value of idea exchange and building a community: “Too often, data people are disconnected from each other, sitting at their desks, without peers to exchange ideas with. This provided them a forum and a way to share common challenges, jointly problem-solve, and deepen their knowledge.”

Conclusion

The COVID-19 crisis has put data at the center of decision making for many chief executives in government. An emergent situation such as this can inspire new ways of sharing data. The clarity and urgency of the policy problem can provide both focus and an incentive to find solutions efficiently, and external pressure can inspire speedy action.

While intergovernmental data sharing has the potential to dramatically improve government operations and provide more personalized and efficient service to the public, it remains an underexploited area. As governments respond to the crisis and opportunity of workforce and economic changes caused by COVID-19, an area that should be examined is how sharing data across silos can make government better and more responsive to customer needs, both now and into the future. Future research should examine not just how data is shared but the policy implications of such sharing, including how the privacy requirements of linked data differ from the requirements of the source data, how data bias in source systems is impacted by the linking of data across systems, and how the automation of data ingestion and processing will change the value proposition for such systems.

Data and the sharing of data for insight is increasingly important to everyday operations in government. Additional focus on intergovernmental data sharing can benefit the health, safety, prosperity, and well-being of the public.

63. Stephen Goldsmith and Jane Wiseman, “What Networked CDOs Can Bring to Their Cities,” *Governing*, May 1, 2018, <https://www.governing.com/blogs/bfc/col-cities-chief-data-officers-networking.html>.

SOURCES

Access Indiana. 2020, www.in.gov.

Allegheny County Department of Human Services. Allegheny County Data Warehouse, July 2019.

The Annie E. Casey Foundation. Using Integrated Data Systems to Improve Case Management and Develop Predictive Modeling Tools, 2017.

ATS Corporation. AST Provides CJIS Solution for Nation's 2nd Largest County, 2020.

Auspos, Patricia. Using Integrated Data Systems to Improve Case Management and Develop Predictive Modeling Tools. The Annie E. Casey Foundation, March 17, 2020.

Brady, Henry E. The Challenge of Big Data and Data Science, *Annual Review of Political Science*, 2019. 22:297-323.

Brown, Jessica Leigh. 4 Best Practices for Implementing State Longitudinal Data Systems. Education Dive, May 2, 2019.

Chowdhury, Anir, Policy Advisor, UNDP and the Government of Bangladesh, Member of the Prime Minister's National Digital Task Force. Regional Innovation Centre UNDP Asia-Pacific, The Innovation Dividend Podcast: How the Bangladesh government is repurposing assets for COVID-19, medium.com, April 20, 2020, <https://medium.com/@undp.ric/how-the-bangladesh-government-is-repurposing-assets-for-covid-19-650c26019259>.

Cook County Bureau of Technology. Modernization and Innovation Report. 2018.

Cook County Government. "Cook County Board Approves New Technology Modernizing Data Systems Used by Justice Agencies." December 16, 2015.

Coral, Lilian. "The Power of Data Visualization in Cities: Los Angeles' GeoHub." *Data-Smart City Solutions*, Harvard University, Ash Center for Democratic Governance and Innovation, September 6, 2017.

DeHart, Dana and Shapiro, Cheri. "Integrated Administrative Data & Criminal Justice Research," *American Journal of Criminal Justice*. Vol. 42 Issue 2, p255-274, June 2017.

Doar, Robert and Linda Gibbs. "Unleashing the Power of Administrative Data: A Guide for Federal, State, and Local Policymakers," Results for America and American Enterprise Institute. October 19, 2017.

Edelstein, Sam. "Why Being a Chief Data Officer is Like Running a Minor League Baseball Team." June 3, 2019.

E-Estonia.com. e-Estonia Facts. April 15, 2020.

E-Estonia.com. e-Estonia Guide. April 4, 2018.

E-Estonia.com. Frequently Asked Questions: Story of e-Estonia. March 2020.

Fry, Hannah. "Your Number's Up." *The New Yorker*, September 9, 2019.

Gelobter, Lisa. "Under the Hood: Building a New College Scorecard with Students." Home Room, U.S. Department of Education, September 12, 2015.

Gardner, Betsy. "How NYC is Using Transit Data for Equity Goals," *Data-Smart City Solutions*, Harvard University, Ash Center for Democratic Governance and Innovation, February 18, 2020.

Gardner, Betsy. "Moving Beyond Mobility as a Service: Interview with Seleta Reynolds," *Data-Smart City Solutions*. Harvard University, Ash Center for Democratic Governance and Innovation December 18, 2019.

Goldsmith, Stephen, and Leger, Matthew. "Mobility and the Connected City: Effectively Managing Connected Mobility Marketplaces," *Data-Smart City Solutions*. Harvard University, Ash Center for Democratic Governance and Innovation February 20, 2020.

Goldsmith, Stephen and Leger, Matt. "An Ambitious Scheme to Bring All of a State's Data Together." *Governing*, February 19, 2020.

Goldsmith, Stephen and Wiseman, Jane. "What Networked CDOs Can Bring to Their Cities." *Governing*, May 1, 2018.

Governing Institute. Making "Good Government" Great: The Role of Integrated Data Systems. 2018.

Government Technology. "Interagency Data Sharing in the Time of COVID-19." Webinar Slide 6, May 29, 2020.

Groves, Patrick. "Building a Data-Driven Future: Digital Cities 2019 Revealed." *Governing*, November 16, 2019.

Harris, Natalie Evans. "Sharing Data for Social Impact: Guidebook to Establishing Responsible Governance Practices." Beeck Center for Social Impact + Innovation at Georgetown University, January 2020.

Harsha, Daniel. "From Liability to Asset: Stephen Goldsmith on How Ridesharing has Transformed the Humble Sidewalk into a Valuable Piece of Urban Real Estate," *Data-Smart City Solutions*. February 27, 2020.

Herz, Denise C., Dierkhising, Carly B. "OJJDP Dual System Youth Design Study." *National Criminal Justice Reference Service*, March 2019. P.134.

Ideas42. New Text Message Reminders for Summons Recipients Improves Attendance in Court and Dramatically Cuts Warrants. 2018.

Institute of Education Sciences. *History of the SLDS Grant Program*. April 2020.

Institute for Social Capital. Status of ISC's Data Sharing Agreements. UNC Charlotte.

Jankowicz, Nina. "Estonia Already Lives Online—Why Can't the United States?." *The Atlantic*, May 27, 2020.

Jenkins, Adelia and Culhane, Dennis. "Integrated Data Systems in the U.S.: A National Survey of State and Local Governments and their University Partners." *International Journal of Population Data Science*, November 19, 2019.

Johnston, Ryan. "Austin launches coronavirus testing enrollment service." *State Scoop*, April 17, 2020.

Kitzmiller, Erika M. "IDS Case Study: Washington State's Integrated Client Data Base and Analytic Capacity. Actionable Intelligence for Social Policy (AISP)," University of Pennsylvania, 2013.

Konkel, Frank. "How Customer Experience is Disrupting Government Technology," *Nextgov*, April 5, 2019.

Lane, Julia. "After Covid-19, the U.S. Statistical System Needs to Change," *Significance Magazine*, Royal Statistical Society, August 2020.

McKinsey. How to make the most of AI? Open up and share data. June 9, 2020, David DeLallo and Jeni Tennison. Podcast.

MetroLab. Digital Government Transformation at Full Speed, Webinar, May 8, 2020.

NewVantage Partners. Big Data and AI Executive Survey 2020: Executive Summary of Findings. 2020.

Office for National Statistics (United Kingdom), Census-taking in the ancient world, January 18, 2016.

Palfrey, Quentin. "5 Strategies for Evidence-Based Policymaking." *Governing*, February 27, 2017.

Parrish, Rick. "The U.S. Federal Government Still Ranks Near the Bottom of Forrester's Customer Experience Index." *Forrester*, May 31, 2018.

Parsons, Jim and Sandwick, Talia. "Closing the gap: Using criminal justice and public health data to improve the identification of mental illness." Vera Institute of Justice, July 2012.

Perez Jr., Zeke. 50-State Comparison: Statewide Longitudinal Data Systems. Education Commission of the States, August 19, 2019.

Pew Research Center. Little Public Support for Reductions in Federal Spending. April 11, 2019.

Portillo, Ely. "Do youth end up in the justice system because they disengage from school?," UNC Charlotte Urban Institute, June 10, 2019.

Population Reference Bureau, *Milestones and Moments in Global Census History*, September 4, 2019.

Priest, Jessica. "Coronavirus: 1 in 3 death certificates before virus were wrong, and it's getting worse." *SouthCoast Today*, April 28, 2020.

Regional Innovation Centre UNDP Asia-Pacific. "The Innovation Dividend Podcast: How the Bangladesh government is repurposing assets for COVID-19." *medium.com*, April 20, 2020.

Rivero, Carlos. "Data Governance Roadshow," presentation. 2019. Commonwealth of Virginia.

SanGIS.org. About SanGIS. 2012.

Sorrells, Analisa. "UNCC's Institute for Social Capital links data across education, housing, social services, and more." *EdNC*, May 28, 2019.

Splunk. The State of Dark Data. Slide 18, 2019.

Stack, Kathy. "Harnessing Data Analytics to Improve the Lives of Individuals and Families: A National Strategy." July 12, 2020, <https://www.dayoneproject.org/post/harnessing-data-analytics-to-improve-the-lives-of-individuals-and-families-a-national-data-strategy>.

U.S. Department of State, "Coronavirus Global Response Coordination Unit: Taking care of our people during the COVID-19 pandemic." *State Magazine*. May 2020.

U.S. Government Accountability Office, "Human Services: Sustained and Coordinated Efforts Could Facilitate Data Sharing While Protecting Privacy," GAO-13-106, February 2013.

Stone, Adam. "Alexa and Siri are Resetting Public's Expectations for What a Good Experience with Government Could Be." *Governing*, September 20, 2019.

The Citywide Analytics Team. Our 2020 Vision. Boston.gov, February 2020.

Tran, Billy. "9 Situations in Your HDB Estate That Your Town Council Can Actually Help You With." *The Smart Local Singapore*, March 31, 2020.

Vincent, Brandi. "State Department Harnesses Data to Help Bring Americans Home Safely During the Pandemic." *NextGov*, June 2, 2020

Warzel, Charlie. "All This Dystopia, and for What?." *New York Times*, February 18, 2020.

White, Evan. A Roadmap for Linking Administrative Data in California, UC Berkeley, California Policy Lab, May 7, 2019.

Williams, Jake. "Universities modeling coronavirus impact using state data." *Ed Scoop*, April 14, 2020.

Wiseman, Jane. "Lessons from Leading CDOs: A Framework for Better Civic Analytics." *Data-Smart City Solutions*. January 25, 2017.

Wiseman, Jane. "Engines of Innovation: How investments in data and digital infrastructure and human capital paved the way for customer-responsive and data-informed government in Singapore," July 2020. <https://scholar.harvard.edu/janewiseman/publications/engines-innovationhow-investments-data-and-digital-infrastructure-and-human>.

Wiseman, Jane. "Driving Toward Greater Safety and Efficiency: Urban Mobility Data Exchanges," September 2020. <https://scholar.harvard.edu/janewiseman/publications/driving-toward-greater-safety-and-efficiency-urban-mobility-data-exchanges>.

Wiseman, Jane. "Better Government Through Data: Using the Allegheny County Human Services Data Warehouse to Design More Effective Results," October 2020. <https://scholar.harvard.edu/janewiseman/publications/better-government-through-data-using-allegheny-county-human-services-data>.

Wiseman, Jane. "Two Models for Successful Data Sharing," Blog Post. IBM Center for The Business of Government, December 8, 2020. Retrieved at: <http://www.businessofgovernment.org/blog/two-models-successful-intergovernmental-data-sharing>.

Wiseman, Jane. "Eight Keys to Accelerate Government Data Innovation," Blog Post. IBM Center for The Business of Government, December 10, 2020. Retrieved at: <https://www.dayoneproject.org/post/harnessing-data-analytics-to-improve-the-lives-of-individuals-and-families-a-national-data-strategy>.

INTERVIEWS WITH AUTHOR

Costa Leabo, Stefanie, Chief Data Officer, City of Boston, June 2, 2020.

deGarmo, Janice, Acting Chief Data Officer, U.S. Department of State, July 24, 2020.

Edelstein, Sam, Chief Data Officer, City of Syracuse, June 4, 2020.

Goodman, Elizabeth, Director of Design, A1M Solutions, April 17, 2020.

Gypalo, Dessa, Chief Data Officer, Cook County, IL, January 24, 2020.

Hefner, Rebecca, Data and Analytics Officer, City of Charlotte, May 8, 2020.

Ladin-Sienne, Sari, Program Manager, Bloomberg Harvard City Leadership Initiative, May 29, 2020.

Lane, Julia, Professor, NYU, August 6, 2020.

McWhorter, Robin, U.S. Citizenship and Immigration Services, May 15, 2020.

Nicklin, Andrew, Futurist-At-Large, Centers for Civic Impact at Johns Hopkins University, June 4, 2020.

Perez, Zeke, Jr., Policy Analyst, Education Commission of the States, April 14, 2020.

Ritualo, Amy, Division Chief, U.S. Citizenship and Immigration Services, May 15, 2020.

Rodriguez, Olga, Special Advisor, Office of the Senior Vice President/Health Science Center, Texas A&M University, April 28, 2020.

Rountree, Janey, Executive Director, California Policy Lab at UCLA, June 3, 2020.

Stack, Kathy, CEO, KB Stack Consulting, August 3, 2020.

von Zastrow, Claus, Principal, Education Commission of the States, April 14, 2020.

White, Evan, Executive Director of the California Policy Lab at UC Berkeley, June 24, 2020.

ABOUT THE AUTHOR

Jane Wiseman leads the Institute for Excellence in Government, a nonprofit consulting firm dedicated to improving government performance. She is also an Innovations in American Government Fellow at Harvard University's Ash Center for Democratic Governance and Innovation. She has served as an appointed official in government and as a financial advisor and consultant to government.

Her current consulting, research, and writing focus is on government innovation, data-driven decision making, and operational efficiency in government. With the Harvard Ash Center and the United Negro College Fund, she is developing a data-informed approach for urban responses to the upskilling and reskilling of unemployed and underemployed individuals. With the Harvard Ash Center, she supports a national network of urban chief data officers to accelerate the use of analytics in local government. With Harvard Ash Center, she created an open platform for government access to the best examples of government operational efficiency approaches, with case studies showing successful implementation. She has written on customer-centric government, data-driven decision-making in government, pretrial justice, and 311 for a variety of audiences.

Her prior consulting work has included organizational strategy, performance management, and eGovernment strategy work for Accenture and Price Waterhouse. Selected clients include the National Governor's Association, the United States Department of Veterans Affairs, the National Criminal Justice Association, the Commonwealth of Massachusetts, the United States Postal Service, the State of Michigan, the United States Department of Housing and Urban Development, and the United States Department of Commerce.

Ms. Wiseman has served as Assistant Secretary, Massachusetts Executive Office of Public Safety and as Assistant to the Director for Strategic Planning, National Institute of Justice, United States Department of Justice. Ms. Wiseman represented the Justice Department on detail as a Staff Assistant for the U.S. House of Representatives Appropriations Committee. Ms. Wiseman holds a Bachelor of Arts in Government from Smith College and a Master of Public Policy from the Harvard Kennedy School.



JANE WISEMAN

KEY CONTACT INFORMATION

To contact the author:



Jane Wiseman
CEO, Institute for Excellence in Government, and
Senior Fellow, Ash Center for Democratic Governance and Innovation
Harvard University
124 Mt. Auburn Street
Cambridge, Mass. 02138
Phone: (617) 448-4250
Jane.Wiseman@gmail.com

REPORTS FROM THE IBM CENTER FOR THE BUSINESS OF GOVERNMENT



For a full listing of our reports, visit www.businessofgovernment.org/reports

Recent reports available on the website include:

Agility:

How Localities Continually Adapt Enterprise Strategies to Manage Natural Disasters by Katherine Willoughby, Komla D. Dzigbede and Sarah Beth Gehl

The Road to Agile Government: Driving Change to Achieve Success by G. Edward DeSeve

Transforming How Government Operates: Four Methods of Change by Andrew B. Whitford

Agile Problem Solving in Government: A Case Study of The Opportunity Project by Joel Gurin and Katarina Rebello

Applying Design Thinking To Public Service Delivery by Jeanne Liedtka and Randall Salzman

Digital:

The Rise of the Sustainable Enterprise by Wayne S. Balta, Jacob Dencik, Daniel C. Esty and Scott Fulton

Innovation and Emerging Technologies in Government: Keys to Success by Dr. Alan R. Shark

Risk Management in the AI Era: Navigating the Opportunities and Challenges of AI Tools in the Public Sector by Justin B. Bullock and Matthew M. Young

More Than Meets AI: Part II by the Partnership for Public Service and The IBM Center for The Business of Government

The Impact of Blockchain for Government: Insights on Identity, Payments, and Supply Chain by Thomas Hardjono

A Roadmap for IT Modernization in Government by Dr. Gregory S. Dawson

Delivering Artificial Intelligence in Government: Challenges and Opportunities by Kevin C. Desouza

Effectiveness:

Measuring the Quality of Management in Federal Agencies by James Thompson and Alejandra Medina

Mobilizing Capital Investment to Modernize Government by Steve Redburn, Kenneth J. Buck and G. Edward DeSeve

Scaling Evidence-Based Programs in Child Welfare by Patrick Lester

Responding to Global Health Crises: Lessons from the U.S. Response to the 2014-2016 West Africa Ebola Outbreak by Jennifer Widner

Insight:

Integrating Big Data and Thick Data to Transform Public Services Delivery by Yuen Yuen Ang

A Practitioner's Framework for Measuring Results: Using "C-Stat" at the Colorado Department of Human Services by Melissa Wavelet

Data-Driven Government: The Role of Chief Data Officers by Jane Wiseman

Integrating and Analyzing Data Across Governments—the Key to 21st Century Security by Douglas Lute and Frank Taylor

People:

Distance Work Arrangements: The Workplace of the Future Is Now Edited by John M. Kamensky

Preparing the Next Generation of Federal Leaders: Agency-Based Leadership Development Programs by Gordon Abner, Jenny Knowles Morrison, James Perry and Bill Valdez

Off to a Running State Capital Start: A Transition Guide for New Governors and Their Teams by Katherine Barrett and Richard Greene

Risk:

Managing Cybersecurity Risk in Government by Anupam Kumar, James Haddow and Rajni Goel

About the IBM Center for The Business of Government

Through research stipends and events, the IBM Center for The Business of Government stimulates research and facilitates discussion of new approaches to improving the effectiveness of government at the federal, state, local, and international levels.

About IBM Global Business Services

With consultants and professional staff in more than 160 countries globally, IBM Global Business Services is the world's largest consulting services organization. IBM Global Business Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues, and the ability to design, build, and run those solutions in a way that delivers bottom-line value. To learn more visit ibm.com.

For more information:

Daniel J. Chenok

Executive Director

IBM Center for The Business of Government

600 14th Street NW
Second Floor
Washington, DC 20005
202-551-9342

website: www.businessofgovernment.org
e-mail: businessofgovernment@us.ibm.com

Stay connected with the IBM Center on:



or, send us your name and e-mail to receive our newsletters.



IBM Center for
The Business of Government