



DATA SHARING ACROSS STATE AGENCIES

Improving Constituent Services, Enhancing Policymaking & Reducing Costs

Executive Summary

State governments across the United States are beginning to recognize the value of interagency data sharing. Sharing data improves services to constituents, enables policymakers to make better-informed decisions, decreases data storage and maintenance costs, reduces fraud, and allows staff to spend more time serving constituents, and less time manually extracting and aggregating data.

Before interagency data-sharing initiatives can deliver on their promise, however, states first must overcome legislative, privacy, and technical barriers to data integration. It is equally critical that agencies ensure that the data they choose to integrate will deliver meaningful value.

This paper discusses the value of data-sharing initiatives, potential use cases, barriers to data integration, and recommended steps toward a solution. One key step involves establishing a data governance board and policies to help ensure that the initiative complies with applicable state and federal laws and regulations and delivers suitable return on investment through improved services for constituents, decreased costs, reduced fraud, and more-informed allocation of funding.

Another critical step is the implementation of technologies that support efficient data sharing. These will include a master data management solution as well as an integrated data platform with analytical and reporting capabilities.

We recommend that states start small, initially bringing together data from only a few agencies in order to demonstrate the value of data sharing and increase participation and confidence before engaging large numbers of agencies in a broader initiative.

Introduction

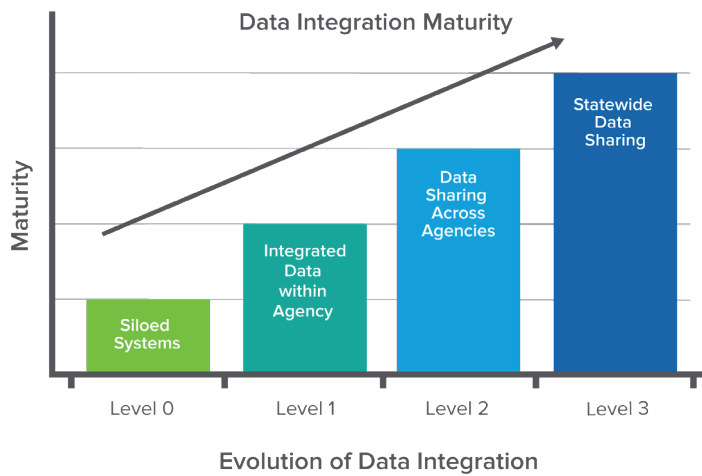
Data is one of the most valuable yet underutilized resources in the public sector. When policies and decisions are based on actionable data, they are more likely to align with constituent needs than policies and decisions based on fragmentary information, guesswork and supposition.

Yet, the data needed to improve governmental policymaking and decision-making is typically maintained in multiple, disparate systems and managed by multiple state government agencies. This makes it difficult if not impossible for analysts to access data that would provide a clearer view of constituents, their needs, and how programs can address these.

For decades, state agencies have operated in siloes, without an unequivocal means of identifying each unique constituent across systems. As a result, a Medicaid provider, for example, will not be informed that an enrollee recently phoned a state-run crisis line or was admitted to a state-run mental health facility. Even if the information were shared, it might be unhelpful, with the enrollee's name appearing as *J. Smith* in one system, *John Smith* in another and *John R. Smith* in a third.

To fully understand constituents and their needs, states must first integrate data from the various systems that contain pertinent information about constituents, de-duplicate data (such as the various renderings of a constituent’s name and other identifying information), and apply tools to analyze the data and create informative and actionable reports.

Although data integration and analytics have been strategic initiatives in the private sector since the late 1990s, they have only recently become high priorities in the public sector.



The CIOs’ Top 10 List

Each year, the National Association of State Chief Information Officers surveys state CIOs in order to determine their Top 10 strategies, management processes, and solutions for the coming year. In 2016, for the first time, business intelligence (BI) and data analytics made the list, with CIOs ranking them No. 4. The association defines BI and data analytics as “applying BI/Business Analytics within the enterprise, communicating the value, building expertise, delivering shared services, exploring big data, data analytics.”¹

The CIOs’ No. 3 priority for 2016 was consolidation/optimization, defined as “centralizing, consolidating services, operations, resources, infrastructure, data centers, communications and marketing.”

Both priorities underscore state governments’ growing recognition of the value of data and analytics and of consolidating data and other resources across agencies.

State CIOs are not alone. State lawmakers also are embracing data integration. In 41 states and Washington, D.C., 318 bills on data sharing have been introduced since 2015.²

The federal government is aligned as well. In 2013 alone, \$42 million in federal grants were issued in support of state-level data sharing initiatives.³

In recent years, many governors and other state executives have recognized the criticality of integrating constituent data and have directed CIOs of state agencies to create integrated, statewide data-sharing policies and platforms.

State CIO Priorities for 2016

- 1 Security + Risk Management
- 2 Cloud Services
- 3 Consolidation/Optimization
- 4 Business Intelligence + Data Analytics
- 5 Legacy Modernization
- 6 Enterprise Vision + Roadmap for IT
- 7 Budget + Cost Control
- 8 Human Resources/Talent Management
- 9 Agile + Incremental Software Delivery
- 10 Disaster Recovery/Business Continuity

In Virginia, for example, Governor McAuliffe in May 2016 issued a directive encouraging state agency CIOs to “review all Commonwealth systems, practices, processes, policies, applicable laws and regulations governing the sharing of data across agencies and create an inventory of state agencies’ data analytics assets, capabilities, best practices, and data-sharing activities” in preparation for defining a strategy for leveraging shared data within the commonwealth.⁴

Other states already had been working toward that objective. New York, South Carolina, Indiana, Illinois, Michigan, Missouri and Arkansas have been among the early adopters of data-sharing initiatives and are currently working toward building data integration solutions.

Even within these states, however, many have integrated only subsets of data. For the most part, the focus has been on crimi-

nal justice and public health. Even there, considerable work remains to be done.

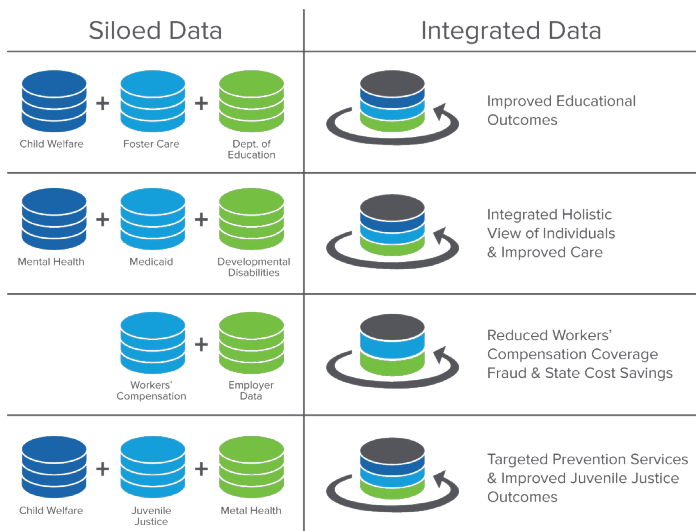
Each morning in Missouri, for instance, emergency room data is now made available to the state’s Medicaid healthcare providers so they can determine whether any of their patients require follow-up care or other services.⁵ While this is a step in the right direction, it does not reflect true data integration. Currently, the emergency room data is copied from a particular system and manually inserted in emails, which are then sent to providers.

An integrated approach would automate the transfer of data between systems and provide automated alerts to specific providers, saving time and money by eliminating the need for manually copying and emailing the data.

The Value of Inter-Agency Integration

Although the integration of state-government data delivers many benefits, four of the most important are:

1. Enhanced services and outcomes for constituents
2. More-informed policymaking
3. Decreased costs
4. Reduced fraud and abusive activity



Enhanced services and outcomes for constituents

Most Americans are served by more than one state agency. Yet in the absence of a statewide integrated data platform, the agencies that serve any given constituent have only a narrow view of that person and his or her needs.

In many states, for example, data regarding behavioral health history, medical history, and involvement with the criminal justice system is maintained in separate systems. If a law enforcement officer encountered a person with a mental illness, information about the person’s behavioral health history, state-provided treatment, and medical history would not be available to the officer. Making that information available could impact the officer’s interactions with the person as well as the outcome for the person.

Interagency data integration offers a broader and more cohesive picture of constituents and can result in better care and outcomes for constituents. It can answer questions regarding how certain services and programs impact lives, enabling the state to coordinate and enhance services, improve the efficiency of service delivery and reduce the information collection burden on the public.

More-informed policymaking

Many policies, regulations, and laws cut across multiple agencies’ data. An integrated data platform can lead to better-informed, data-driven policymaking.

In New York, for example, integration of welfare and criminal justice data enabled policymakers to determine that individuals who first entered foster care after age nine were more likely to be involved in the juvenile justice system. Based on this analysis, the state allocated more than \$20 million targeting prevention services for children in this age group.⁶

Decreased costs

Data sharing also reduces data duplication, processing and storage, ultimately leading to lower costs over time. Additional cost savings arise as agencies spend less time entering data and managing duplicative data. Through the implementation of self-service tools and automated reporting, less time can be spent aggregating data and more time can be spent serving constituents.

Decreases in costs and labor ensure that a greater portion of taxpayer dollars is spent on policies that support constituents as opposed to the management and maintenance of data.

Reduced fraud and abusive activity

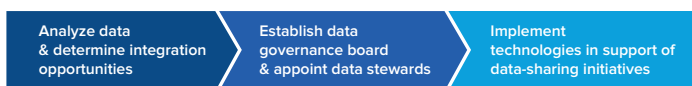
By addressing the problem of siloed data, states are also bet-

ter able to prevent and reduce fraud which provides additional cost savings.

In North Carolina, for example, integrating data from the Division of Employment Security, Rate Bureau, and Department of Insurance, has allowed the state to improve identification of workers' compensation coverage fraud. The data integration has provided visibility to the number of employees by employer as compared to worker insurance coverage and provide alerts for potential noncompliance. Since integrating the data, the state has collected nearly \$1 million in civil penalties and 600 employers who lacked coverage previously now have coverage.⁷

Solving the Problem: Recommended Steps

An effective strategy for establishing an integrated data platform requires analysis, governance and technical expertise.



Analyze data and determine integration opportunities

Before states move to a path of integration, it is important to first define desired outcomes. These outcomes should inform the strategy and the determination of what data should be integrated. In order to accomplish this, it requires analyzing existing systems and identifying data that would be of value in an integrated environment. It is important to recognize that not all data should be integrated.

Integrating data from a state education department with data from state-run hospitals, for example, probably wouldn't deliver sufficient return on investment (ROI) to justify the cost and effort required to bring the data together.

In contrast, integrating select data from state-run hospitals with select data from in-home behavioral health providers could deliver significant value, as it would enable staff to provide better follow-up care after determining, for example, that a person had been hospitalized for a suicide attempt or drug overdose.

In determining what data merits integration, it is important to consider not only potential ROI, but also public perceptions and state and federal regulations. Assembling data regarding

constituents' criminal background, medical history and prior treatment for substance use might provide the basis for valuable insights for policymakers, but it could lead to a public outcry about privacy rights, including those afforded by the federal Health Insurance Portability and Accountability Act (HIPAA) and Code of Federal Regulations (CFR).

In Illinois, where multiple state agencies are participating in a data integration project, the agencies "first had to establish the rules of the road," according to GovTech.com. "First and foremost was understanding privacy rules. It's understood, for example, that (HIPAA) restricts the use of personal medical information. Less well known is the law's ban on the use of certain depersonalized data, even for statistical purposes."⁸

Establish a data governance board and appoint data stewards

The most effective way to establish and enforce data sharing policies and regulations is through data governance and data stewardship.

Data governance involves setting standards, policies, and procedures that apply to collecting, integrating, and sharing data. Policies should stipulate what data can and cannot be integrated; who can access the integrated data; how the integrated data can be used; how permissions will be granted; how data will be stored; and how metadata (or information about data) will be created and managed. Data governance also sets standards for data quality.

Data governance should be established at both the state level and the agency level, with awareness of and alignment with applicable state and federal laws and regulations.

A major reason that defining standards, policies, and procedures is necessary in an interagency environment is that each agency, historically, has followed its own path, collecting and managing data in its own way. When multiple agencies share data but do not share data governance principles, confusion inevitably will arise regarding access, privacy, and security and regarding the data itself. One agency, for example, might refer to a given person's race as "Caucasian" while another agency might refer to it as "white." Establishing data standards will promote data consistency and higher data quality, leading to more reliable information.

Within each agency, a data steward should be appointed to enforce data governance principles, ensuring that data is collected, stored, and managed in compliance with the agency's standards as well as the state's standards.

At the state level, a data governance board, with representation from the agencies participating in the data-sharing initiative, should lead the effort to define data governance policies and procedures and keep these current and consistent with evolving state and federal laws.

One example of a state that has successfully implemented a state data governance board is New Jersey. To coordinate data sharing among agencies, the state established the New Jersey Data Governance Office, responsible for executing data governance policies and processes set by the state's Data Governance Executive Committee.⁹

Implement technologies in support of data-sharing initiatives

With data governance and data stewardship established, the next step in an interagency data-integration project involves selecting and implementing technologies to support the initiative. These will include a master data management solution as well as an integrated data platform with analytical and reporting capabilities.

It is worth noting that these technologies won't necessarily replace existing technologies used by the agencies involved in a data-sharing initiative. In most cases, each agency's employees will continue to enter and manage data within existing systems. What will change is that select data from participating agencies will be sent from existing systems to the master data management solution and, in turn, to the integrated data platform.

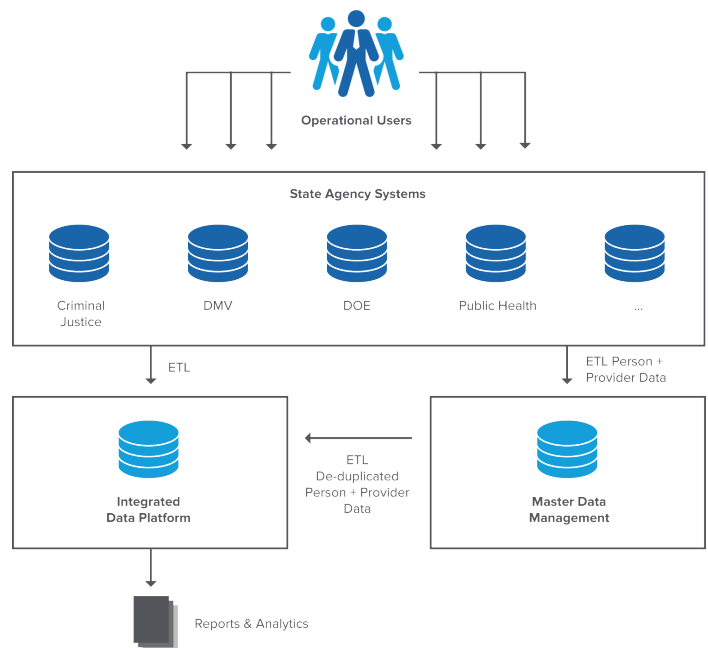
That said, we recommend that states leverage existing technologies and resources when possible in order to save time, labor and money; however, they may need assistance from outside consulting or technology firms in developing an effective data strategy and model and implementing needed tools. Typically, the master data management solution will be an extract, transform and load (ETL) tool that will de-duplicate data regarding people and establish a master record for each unique person. If different agencies collect data regarding the same person but have captured some details differently (for

example, using the name John Doe in one system and Johnathan Doe in another), the ETL tool will determine whether these differences suggest that the data refers to two different people or to one person to whom different designators, or values, have been applied. Along with this, a single value will be selected for each data field in the individual's master record. This will enforce reporting consistency going forward. De-duplication and data standardization are essential to the effective analysis and reporting of data.

To bring together all available information about constituents whose data appears in multiple systems, an integrated data platform is needed. This usually takes the form of a data warehouse.

Once the integrated data platform has been designed, built, and tested (to ensure that the information it contains is usable and actionable), the next step is to create automated reports and provide tools for self-service reporting and analytics.

These capabilities will enable users to inform decision-makers through real-time access to information in the form of reports, dashboards, and analytics, leading to data-driven policymaking and decision-making.



Approach: Starting Small

We recommend that states start small, integrating only two or three agencies' data as a proof of concept, and then build on

that through an iterative approach. This will make the project manageable while providing an opportunity for project sponsors to build trust and gain interest across agencies, helping boost participation in the initiative.

North Carolina began by piloting state data integration through two projects¹⁰:

- Criminal Justice Law Enforcement Automated Data Services, which consolidates data from multiple law enforcement agencies into a single portal
- The North Carolina Financial Accountability and Compliance Technology System, which shares data from multiple sources in order to identify fraud and waste in areas such as unemployment insurance

Based on the success of these projects, North Carolina since has established the Government Data Analytics Center, which serves as the analytics hub for all state agencies. North Carolina has also begun developing data-sharing initiatives involving data regarding child welfare, juvenile justice, and Medicaid. Additionally, the state is developing a statewide health information exchange to enable sharing of medical information with Medicaid providers.

“The strategy has been to start small and think enterprise,” said John Correllus, director of the North Carolina Government Data Analytics Center, speaking in an interview with Government Technology magazine.¹¹ “This has allowed the organization to grow with each success to what it has become today. Our strategies have always been focused on business enablement and benefits.”

Challenges

States that initiate data-sharing projects will face challenges, including legislative, privacy, and technical issues. With the right tools, support, and expertise, however, these challenges can be overcome.

Legislative

Creation of data-sharing policies may be time-consuming and complex, calling for the enactment of new legislation. To clear the way for data sharing in Illinois, for example, the state Legislature passed a bill creating an open data platform as well as regulatory architecture. “The law requires agencies to make

architectural choices with open data in mind,” notes an article on the Govtech.com website.¹²

Privacy

Protecting the privacy of data and the individuals it is about requires protocols that all agencies sharing the data agree to follow. Establishing these protocols requires research and an understanding of applicable federal and state laws.

A variety of laws may apply, depending on the data shared and the state in which the data-sharing initiative takes place.

Here are a few examples:

- Federal HIPAA provisions protect the privacy and security of personal health information (PHI) and allow regulators to impose criminal and civil penalties for violations. If PHI is compromised in a data breach, covered entities must provide notice according to federal guidelines.¹³
- CFR-42 restricts the disclosure and use of patients’ alcohol and drug abuse records, including prognosis, diagnosis, and treatment information.¹⁴
- The Federal Trade Commission (FTC) Act prohibits unfair and deceptive practices. It applies to offline and online privacy and data security policies. The FTC has brought enforcement actions against companies for failing to comply with posted privacy policies and for the unauthorized disclosure of personal data. The FTC also enforces the Children’s Online Privacy Protection Act, which covers online collection of information from children.¹⁵
- Many states also have enacted laws that protect personal data. Organizations that collect personal information online from California residents, for example, must comply with the California Online Privacy Protection Act.¹⁶
- Montana lawmakers have enacted protections that go beyond HIPAA. In Montana, healthcare information can be shared with another state or public health agency only when it is necessary to provide health services to a patient.¹⁷
- A total of 46 states as well as the District of Columbia, US Virgin Islands, Guam, and Puerto Rico have enacted breach notification laws, requiring data owners to notify affected individuals in the event of unauthorized access to or acquisition of personal information. Fifteen states further require that a state regulator, usually the state attorney general’s office, be notified in the event of a breach.¹⁸

To limit violations involving the sharing of PHI, agencies can remove data elements that would identify individuals. These might include name, age, race, address, date of birth, and Social Security number. However, removing this information may not always be an option as some personally identifiable data fields are necessary for certain types of reporting and data sharing.

“The nuances of state law – and HIPAA rules – arguably make it easier not to share data out of fear of reprisal,” notes a white paper by The Governing Institute. “However, it’s critical that states understand HIPAA exceptions and develop a safe and effective framework for data sharing.”¹⁹

Technical

State agencies today use diverse database platforms. They also have diverse data models, which determine how data is captured and structured and how data elements relate to one another. The data is also likely to be varying quality.

All data integration projects, including those taking place in the private sector, face similar challenges. They are particularly complex when integration takes place across government

agencies, because each agency historically has followed its own standards regarding data collection and design. Before an integration project gets under way, these inconsistencies will need to be remediated.

Conclusion

Data sharing within and among state agencies is critical to improving services to constituents, enabling policymakers to make data-driven decisions and reducing unnecessary data storage and maintenance costs.

For too long, information has been an underutilized resource in the public sector, and data integration and analytics have been de-prioritized because of legislative, privacy, and technical barriers. However, recently, the value and benefits of data sharing are being recognized across the United States, and governors are mandating agencies to take action in identifying data sharing opportunities, defining policies for governance, and implementing a solution. These directives, while challenging and time-consuming to implement, will ultimately lead to improved services for constituents and more-informed allocation of funding.

How CapTech Can Help

CapTech has worked extensively with state governments, including performing a wide range of work at state agencies in Virginia, Georgia, Missouri, Nevada, and Kentucky. CapTech consultants are experienced and certified in big data, data strategy and architecture, data analytics, data governance, and data engineering.

In Virginia, a CapTech team integrated Medicaid, state hospital, behavioral health, developmental disability and criminal justice data with the implementation of a data warehouse and master data management solution. The integrated system, which houses information about all aspects of patient care, has enabled the state to ensure a higher quality of life for the individuals supported.

The project received Virginia's 2015 Governor's Technology Award for Innovative Use of Big Data and Analytics.²⁰

The master data management solution matches individuals across systems, enabling consistent reporting of individual information and improved traceability and management of individuals' records. The data warehouse, which serves as the integrated data platform, is complemented by self-service tools that give the staff quick and easy access to information. Previously, employees spent hours each day searching for information and trying to aggregate data for reporting. The self-service tools enable them to get the information and run reports in a matter of seconds. This frees up time to focus on other priorities, including constituents and how to provide better care for them.

Using the tools that the CapTech team implemented and trained staff on, employees with varying technical backgrounds can answer a wide range of questions in real time, without the assistance of data analysts or business analysts.

In connection with the Virginia project, CapTech helped establish a data governance process and, along with it, a data governance board. The project streamlined, automated, and standardized reporting for internal and external needs, saving an average of 80 man hours per week and \$100,000 per year.



AUTHOR BIO

Gabriella Lively | glively@captechconsulting.com | 804.355.8059

Gabriella Lively is an IT professional with expertise in state government and the healthcare industry. She has experience in data warehousing, data engineering, data analytics, and business intelligence. Ms. Lively has helped a variety of clients utilize their existing data more efficiently to drive business planning and make more informed decisions through the integration of disparate systems and the implementation of analytic solutions to report, monitor, and predict outcomes.

End Notes

- ¹ “State CIO Top Priorities for 2016.” Nov. 10, 2015. National Association of State Chief Information Officers. Available at http://www.nascio.org/Portals/0/Publications/Documents/2015/State_CIO_Top%20Ten_Policy_and_Technology_Priorities_for_2016.pdf.
- ² “N.J. bill would move state toward automated criminal justice data system.” July 19, 2016. StateScoop. Available at <http://statescoop.com/n-j-bill-would-move-state-toward-automated-criminal-justice-data-sharing-system>.
- ³ “Cross-state data sharing.” Workforce Data Quality Campaign. Publication date unknown. Accessed July 26, 2016. Available at <http://www.workforcedqc.org/federal-solutions/cross-state-data-sharing>.
- ⁴ “Executive directive 7 (2016). Leveraging the use of shared data and analytics.” 2016. Commonwealth of Virginia. Available at <https://governor.virginia.gov/media/6007/ed-07-leveraging-the-use-of-shared-data-and-analytics.pdf>.
- ⁵ “Data sharing and analytics: Changing HHS for the better.” Jan. 9, 2015. Government Technology. Available at <http://www.govtech.com/health/Data-Sharing-and-Analytics-Changing-HHS-For-the-Better.html>.
- ⁶ “Innovations in NYC Health and Human Services Policy. Data Integration and Cross-Agency Collaboration. January 2014. Vera Institute of Justice. Available at http://www.nyc.gov/html/ceo/downloads/pdf/policybriefs/DataIntegrationBrief_011314.pdf.
- ⁷ “Data sharing helps North Carolina detect workers’ comp fraud.” August 6, 2015. GNC. Available at https://gcn.com/articles/2015/08/06/workers-comp-fraud-detection.aspx?s=gcn_tech_070815
- ⁸ “Data sharing and analytics: Changing HHS for the better.” Jan. 9, 2015. Government Technology. Available at <http://www.govtech.com/health/Data-Sharing-and-Analytics-Changing-HHS-For-the-Better.html>.
- ⁹ “2014 NASCIO recognition award submission.” Nov. 6, 2013. National Association of State Chief Information Officers. Available at <http://www.nascio.org/portals/0/awards/nominations2014/2014/2014NJ11-NASCIO%202014%20Nomination-Data%20Governance%20Office.pdf>.
- ¹⁰ “North Carolina IT takes control of state data.” Oct. 14, 2014. Government Technology. Available at <http://www.govtech.com/data/North-Carolina-IT-Takes-Control-of-State-Data.html>.
- ¹¹ “Taking a Peek Inside North Carolina’s Government Data Analytics Center.” March 14, 2016. Government Technology. Available at <http://www.govtech.com/data/Taking-a-Peek-Inside-North-Carolinas-Government-Data-Analytics-Center.html>.
- ¹² “Data sharing and analytics: Changing HHS for the better,” Govtech.com. Jan. 9, 2015. Available at <http://www.govtech.com/health/Data-Sharing-and-Analytics-Changing-HHS-For-the-Better.html>.
- ¹³ “Data protection in the United States: an overview.” July 1, 2016. Practical Law. Available at <http://us.practicallaw.com/6-502-0467>.
- ¹⁴ “42 CFR Part 2, Subpart A – Introduction.” 2015. Cornell University Law School Legal Information Institute. Available at Added this one-sourced from here: <https://www.law.cornell.edu/cfr/text/42/part-2/subpart-A>.
- ¹⁵ “Data protection in the United States: an overview.” July 1, 2016. Practical Law. Available at <http://us.practicallaw.com/6-502-0467>.
- ¹⁶ “Data protection in the United States: overview.” July 1, 2016. Practical Law, op cit. Available at <http://us.practicallaw.com/6-502-0467>.
- ¹⁷ “How data sharing can prevent fraud in state public assistance programs.” 2016. The Governing Institute. Available at <http://www.governing.com/papers/How-Data-Sharing-Can-Prevent-Fraud-in-State-Public-Assistance-Programs-44948.html>.
- ¹⁸ “Data protection and privacy in 26 jurisdictions worldwide.” 2014. Law Business Research Ltd. Available at https://www.hunton.com/files/Publication/1f767bed-fe08-42bf-94e0-0bd03bf8b74b/Presentation/PublicationAttachment/b167028d-1065-4899-87a9-125700da0133/United_States_GTDT_Data_Protection_and_Privacy_2014.pdf.
- ¹⁹ “How data sharing can prevent fraud in state public assistance programs.” 2016. The Governing Institute. Available at <http://www.governing.com/papers/How-Data-Sharing-Can-Prevent-Fraud-in-State-Public-Assistance-Programs-44948.html>.
- ²⁰ “Big Data Saves \$100,000 of Tax Payer Money While Improving Quality of Care for Patients.” Sept. 27, 2015. Press release, CapTech. Available at <https://www.captechconsulting.com/news/big-data-saves-tax-payer-money>.

Related CapTech Blogs

<https://www.captechconsulting.com/blogs/five-ways-modern-data-governance-will-make-your-organization-more-productive>
<https://www.captechconsulting.com/blogs/a-better-way-for-government-to-modernize>
<https://www.captechconsulting.com/blogs/quality-is-a-habit>