

# UNDERSTANDING THE 2018 FEDERAL DATA LAW

MUCH MORE THAN OPEN DATA



# Table of Contents

<b>ABSTRACT</b> .....	<b>2</b>
<b>INTRODUCTION</b> .....	<b>3</b>
<b>TITLE I: EVIDENCE-BUILDING ACTIVITIES</b> .....	<b>4</b>
<b>TITLE II: OPEN GOVERNMENT DATA ACT</b> .....	<b>6</b>
OPEN, PUBLIC, ELECTRONIC, AND NECESSARY (OPEN) GOVERNMENT DATA ACT .....	6
DATA INVENTORY AND FEDERAL DATA CATALOG .....	7
CHIEF DATA OFFICER.....	8
CHIEF DATA OFFICER COUNCIL .....	8
EXPANDING THE CDO ROLE.....	10
<b>TITLE III: CONFIDENTIAL INFORMATION PROTECTION AND STATISTICAL EFFICIENCY ACT OF 2018</b> .....	<b>10</b>
CONFIDENTIAL INFORMATION PROCESSING PROTECTION.....	11
STATISTICAL EFFICIENCY .....	11
DESIGNATED STATISTICAL AGENCIES .....	12
ACCESS TO DATA FOR EVIDENCE .....	13
EXPANDING SECURE ACCESS TO CIPSEA DATA ASSETS .....	13
Applying to Access Data Assets for Developing Evidence .....	14
<b>ANALYSIS</b> .....	<b>14</b>
WHAT DOES THIS MEAN? .....	14
POTENTIAL CHALLENGES AND UNANSWERED QUESTIONS.....	16
<b>ACKNOWLEDGEMENTS</b> .....	<b>19</b>
<b>AUTHORS</b> .....	<b>19</b>
<b>REFERENCES</b> .....	<b>21</b>

## Abstract

While many Americans were shoveling their driveways and sidewalks from a series of winter storms, the Federal Government shut down, and the President signed the Foundations for Evidence-Based Policymaking Act of 2018 (FEPA). Much of what we know about the new law and its impact is still unclear, but several things are readily apparent. First, the role of the Chief Data Officer (CDO) is now incorporated into federal law and separated from the role of the Chief Information Officer (CIO). Second, government data is now open by default, and the Federal Government must maintain its data using open standards. Third, the law requires that the Federal Government carefully manage its data following industry best practices. Fourth, the Act fully recognizes the value and use of data in agency operations especially as part of evidence-based decision-making. Finally, the law anticipates that, collectively, these efforts will improve governmental decision-making and overall effectiveness.

FEPA's broad requirements present federal agencies with new opportunities and challenges. On the one hand, some enthusiastically support the arrival of data management at the national data. On the other hand, some are concerned that without new funding, the Federal Government couldn't reasonably meet these new expectations. There is little doubt, however, that these changes will impact data across federal systems regardless of its format, digital or paper. It is not unreasonable to think that compliance with FEPA requirements could require levels of effort reaching those expended for Y2K compliance in the late nineties.

This paper is a synopsis of the new legislation, the authors' views of the new law, a preliminary assessment of the law, and the impact it may have on government, industry, and the public. We summarize the legislation's significant features, examine how the new law differs from previous legislation, and review the requirement to expand the use of data across the federal landscape. Then, we identify outstanding questions we believe deserve further analysis and clarification.

This report exclusively reflects the authors' thoughts and opinions.

## Introduction

Open data is a type of U.S. data law, and FEPA's use of the acronym "OPEN" — open, public, electronic, and necessary — suggests that it is just that, just another open data law. However, using OPEN this way only serves to obscure the real distinction between FEPA and other open data laws. What we'll see in this paper is that FEPA goes well beyond open data. For example, FEPA includes critical decision-making, security, and has the potential to affect data usage across the entire country. Congress liked the idea — the Senate unanimously passed the bill, and the House approved FEPA by a vote of 356 - 17. Even the President approved of the bill without hesitation. Indeed, many may think that FEPA is just another open data law, but in fact, FEPA really represents the Federal Government's attempt to make data a strategic asset.

There are arguably only a handful of federal data laws requiring the government to share data with its constituents. The first significant law was the Freedom of Information Act (FOIA) in 1967. FOIA requires federal agencies to disclose requested information to the requestor unless it falls under one of nine exemption categories (5 U.S. Code § 552). Years later, the Federal Government passed the Federal Funding Accountability and Transparency Act of 2006 (FFATA) empowering every American with the ability to hold the government accountable for each spending decision. This law resulted in the data sharing site, USASpending.gov, where the public can use government-provided tools to analyze more than 80% of annual government spending. Other laws that affect data include the Government Performance Reduction Act of 1993 (GPRA) and Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA). For the purposes of this paper, though, we plan to focus on FEPA and the affect it has on data management.

A more recent data law was the Digital Accountability and Transparency Act of 2014 (The DATA Act). The DATA Act required the Federal Government to transform and report its spending information into open data. The DATA Act also amended the Federal Funding Accountability and Transparency Act of 2006, addressing two fundamental issues.

Standardized Agency Reporting. First, the DATA Act required the Treasury Department (TRE) and the White House Office of Management and Budget (OMB) to establish government-wide data standards for reporting expenditures to TRE, OMB, and the General Services Administration (GSA). When the government collects data from the public, TRE and OMB publish that information at no cost to the public via the Internet.

A future in which rigorous evidence is created efficiently, as a routine part of government operations, and used to construct effective public policy. Advances in technology and statistical methodology, coupled with a modern legal framework and a commitment to transparency, make it possible to do this while simultaneously providing stronger protections for the privacy and confidentiality of the people, businesses, and organizations from which the government collects information (American Statistical Association, 2017).

The commission focused on improving how data are used to generate evidence in support of federal policies and programs. Additional support for FEPA arrived in the 2018 President's Management Agenda, which called for the development of a long-term national data strategy and for the government to leverage data as a strategic asset (Hart, N., & Shaw, T., 2018). Though this legislation seemed slightly "out of the blue," according to a representative from the Data Coalition (The Data Coalition, n.d.), the group worked for more than three years with help from both political parties to shape FEPA and position it for success in Congress. The Data Coalition is a group that advocates for the increased use of open standards by the Federal Government.

## TITLE I: Evidence-Building Activities

The Federal Government engages in a wide variety of activities and where possible, tries to measure its performance in quantitative terms. For example, the U.S. Department of Commerce (DOC) conducts the Population and Housing Census every ten years to present a count of every person living in the United States. Similarly, the Bureau of Justice Statistics (BJS) submits a report describing justice systems, crime, criminal offenders, and victims of crime, and the Bureau of Transportation Statistics (BTS) prepares a report describing airline on-time performance, pirates at sea, transportation safety and availability, and more. The Federal Government uses these reports — and many others — to measure the health and welfare of the United States and as critical inputs to legislative and executive decision-making, specifically when it comes to building federal budgets and administering services across the country.

Collectively, the analysis behind these reports is known as evidence-building activities. These functions include the collection, compilation, processing, analysis, and dissemination of data to create general purpose, policy- and program-specific statistics. Evidence-building activities also include program evaluation, research, policy- and program-related analysis, performance measurement, and public health surveillance. Federal evidence-building is highly decentralized, requiring each agency to carry

out all the functions. While many departments and agencies have capacity to do some of this work, not every organization can consistently meet the new legislative requirements.

Similarly, how agencies perform this work varies by department and agency. For example, some organizations have established centralized offices for executing one or more of these evidence-building duties, and other organizations have assigned functions across program areas. Generally speaking, departments conduct most statistical activities across mission areas (e.g., health or energy), and the statistical methods that agencies use vary by agency as well. Organizations performing evidence-based work are collectively known as Federal Statistical Agencies (FSA). FSAs are responsible for collecting, compiling, processing, and analyzing data for statistical purposes. Within these agencies, more than 120 analytical components are likely to be impacted by the new legislation.

The table below indicates agencies expected to have the most disruption from the FEPA.

Topic	Bill Section #	CEP No.
<b>Strengthens Privacy Protections</b>		
Establish an Agency Official for Data Policy	101(a) (§314)	3-3
Designate Chief Data Officers	202(e) (§3520)	3-3
Codify Statistical Policy Directive #1	302(a) (§3563 and §3572(b))	3-4
Conduct of Comprehensive Risk Assessment and Analysis of Data Sensitivity	303(a) (§3582)	3-1
<b>Improves Secure Access to Data</b>		
Establish the Advisory Committee on Data for Evidence Building	101(a) (§315)	4-2
Create and update Data Inventories with Metadata	202(d) (§3511)	4-5
Make Data Available for Statistical Activities	303(a) (§3581)	2-3
Develop a Single Process for Researcher Access to Data	303(a) (§3583)	2-8
Improve Transparency About Projects Using Confidential Data	303(a) (§3583(a)(6))	4-3
<b>Enhances Evidence-Building Capacity</b>		
Requires Agencies to Produce Evidence-Building Plans (learning agendas)	101(a) (§312)	5-2
Recognizes the Evaluation Function with Designation of Evaluation Officers and Requirements for Written Evaluation Policies	101(a) (§313)	5-1
Generate an Inventory of Evidence-Building Units	101(c)	5-1
Establish a Chief Data Officer Council	202(f) (§3 520A)	5-3

---

Under Title I of FEPA, all federal agencies are required to manage their data using industry best practices, to regularly analyze the data, and to use the results to inform policymaking. These practices encourage the government to use data for measuring and understanding outcomes, rather than merely counting outputs and activities. As part of managing their data, federal agencies must develop plans around statistical questions and appoint officials who represent agency equities as they relate to analytical inquiries. Agencies must identify data that they propose using to conduct qualified analyses. Researchers must also describe the methods they expect to use, legal obstacles that could impede their work, the plan they intend to follow, and any supporting information the Director of the OMB may request.

Additionally, Title I requires a non-partisan chief evaluation officer who serves as a plan and implementation evaluator with demonstrated domain expertise. The selected individual is expected to correlate program evaluation activities to data production and data quality improvements over time. This role is similar to the European Union's (EU) requirement for a neutral data quality official for all organizations. Title I also requires OMB to establish an advisory council that reviews, analyzes, and makes recommendations for promoting the use of data for evidence-based decisions. The group is called the Advisory Committee on Data for Evidence Building (the Committee). Membership consists of thirteen participants with at least three federal CDOs. The Committee also includes a single CIO to ensure continuity with information technology.

## TITLE II: Open Government Data Act

### Open, Public, Electronic, and Necessary (OPEN) Government Data Act

Title II is the OPEN Government Data Act. The title's overall legislative intent includes increased transparency, self-empowerment, new or improved products, and services, potential economic innovation, and new knowledge and growth areas. The OPEN Government Data Act is the portion of FEPA that details how open data will work for federal data. For instance, Title II requires the government to provide three new foundational capabilities. 1) a government-wide data inventory, 2) a new role called the Chief Data Officer (CDO), and 3) a data governance body that is known as the Chief Data Officer Council (CDO Council).

Open data is content that people can use without restriction. People can modify and share open data with anyone for any purpose. The Federal Government collects an enormous variety of data types to deliver services to its citizens, and many argue that the government has not leveraged its data nearly enough. With the passage of the OPEN Government Data Act, federal data must be open and available for others to use, including other government agencies. Open data also includes access to analog data, which is data not in digital form.

Another critical difference is how data is released to the public. Where FOIA enables access to government data, by request in its original form, the OPEN Government Data Act makes data electronically available by default. No specific application needed, regardless of its original type. Instead of using exempted data categories like those found in FOIA, the Federal Government must now develop a formal release process that ensures that the government does not release sensitive information as part of open data.

## Data Inventory and Federal Data Catalog

All organizations keep track of things that are valuable. Organizations typically record information about their assets from the day they acquire something until the day they dispose of it. Organizations manage inventories using catalogs, which regularly describe assets like automobiles, buildings, and raw materials. The Federal Government now formally recognizes data as an asset. Simply put, federal data has monetary value, and federal agencies need to keep track of it. Inventories help organizations know what data they have. Inventories also allow agencies to see what data others need. They can provide access to save money or charge them to use it. In either case, data can be as valuable — if not more — than the physical products and processes that produce them.

To repeat, the OPEN Government Data Act requires agencies to maintain an inventory of their data assets. Agencies must inventory and submit their data holdings to a centralized federal data catalog that OMB administers. From there, other federal agencies, researchers, and the public can discover data that might be valuable. In this way, federal agencies can give everyone better access to their data, including other agencies, while decreasing the costs required to create the same data across multiple federal agencies.

## Chief Data Officer

Data already supports strategic and operational efforts every day, from the largest and most prominent organizations like the Department of Defense (DoD) to the most obscure organizations like the U.S. Board on Geographic Names (BGN). Data is also a key input for the creation and management of strategy and direction setting for organizations. Effectively using data requires decision-makers that have access to high quality, fit-for-purpose data. Well-formed data can include data from other agencies, which have different charters, authority, and compliance requirements. To that end, the OPEN Government Data Act requires every federal agency to identify an individual who serves as the CDO, the primarily role responsible for managing the Nation's data. The Act directs CDO to be responsible for the following:

- Set standards for data formats, Negotiate terms for data sharing, and develop processes for data publishing;
- Coordinate with any agency officials responsible for using, protecting, disseminating, and generating data;
- Review the impact of agency IT infrastructure on data accessibility and coordinate with agency Chief Information Officer to reduce barriers that inhibit data accessibility;
- Maximize the use of data in the agency to support evidence-based analysis, cybersecurity, and operational improvement;
- Acquire and maintain training and certification related to confidential information protection and statistical efficiency; and
- Be responsible for overall data lifecycle management.

## Chief Data Officer Council

How an organization uses its data is not always clear, and to remedy that confusion, some organizations rely on data governance to set data policy. Still, some think that data governance is no more than clarifying who stewards data. Others recognize that data governance is a human-based business process designed to optimize the value and use of data.

In practical terms, data governance is a deliberative body that an organization uses to make decisions about its data, how people use it, and what standards are necessary to conduct operations. And while more organizations talk about data governance, some realize that data governance is technically complex, organizationally challenging, and politically sensitive. Compounding matters, getting executive support for data governance is critical, and unfortunately, many executives don't understand how crucial leadership commitment is. Most often, data governance lands in information technology (IT) hands because people believe that data is a consequence of IT. When organizations define data as a byproduct of information technology, data typically becomes an object of computing and treated like other IT assets.

However, the OPEN Government Data Act separates business problems and technical problems using the Council as moderator. The Council requires that the government recognize data as the raw material that enables and improves business processes, decisions, and interactions. Additionally, the OPEN Government Data Act recognizes that data MUST be an asset that the organization controls from acquisition through final disposition in the pursuit of generating the highest possible return on taxpayer investments.

The Council is composed of a chair — appointed by the Director OMB — CDOs from each agency, two representatives for all federal chief information officer and evaluation officers, and the Administrator of the Office of Electronic Government (OEG). The Council convenes regularly and sets policy regarding the following:

- Establishing government-wide best practices for the use, protection, dissemination, and generation of data;
- Promoting and encouraging data sharing agreements between agencies;
- Identifying ways agencies can improve the production of data for analysis and policymaking;
- Consulting with the public and engage with private users of government data;
- Consulting other stakeholders on how to improve access to federal data; and
- Identifying and evaluating new technological solutions for improving the collection and use of data.

## Expanding the CDO Role

Of particular interest is a small and seemingly insignificant reference tucked deep inside FEPA. Most people might glance over the citation, but it's worth a second look. Specifically, the law says the following:

FUNCTIONS.—The Chief Data Officer of an agency shall (...) (5) carry out the requirements of the agency under subsections (b) through (d), (f), and (i) of section 3506, section 3507, and section 3511;

These references describe the requirements for government efficiency, collection, and planning as they relate to federal data management. What this means is that FEPA formally requires CDOs to set the policies and procedures across the entire data lifecycle from planning a data collection through final records disposition. In some cases, this means CDOs take a more active and direct role. For example, CDOs must communicate regularly with the public to ensure that an agency's data collection is necessary and not overly burdensome. CDOs, in cooperation with the agency Chief Financial Officer (CFO), are also required to develop a full and accurate accounting of information technology expenditures and goals for improving information resource management's contribution to the productivity, efficiency, and effectiveness of government operation.

## TITLE III: Confidential Information Protection and Statistical Efficiency Act of 2018

Title III's scope is considerable and affects nearly every part of the Federal Government; therefore, oversight and coordination are essential to its success. OMB oversees and coordinates the development of policies to help lead oversight efforts. OMB expects to develop, document, and make transparent business processes to ensure that all federal agencies support the same rules and implement consistent programs. Title III also reinforces the notion that all agencies must centralize their data. Additionally, agency heads provide OMB reports as needed, while designated statistical agencies (DSA) submit annual reports to OMB, the Committee on Oversight and Government Reform of the House of Representatives, and the Committee on Homeland Security and Governmental Affairs of the Senate. Annual reports include information on the Federal Government's progress relative to the law.

## Confidential Information Processing Protection

Perhaps the most crucial part of Title III is the part describing the requirements for protecting sensitive information. Congress found that the Federal Government had no single method for collecting data and guarding public information, and as a result, the public's trust in government declined. Congress recognized that protecting an individual's confidential information serves public and societal interests, and if the public does not trust the Federal Government with its data, the public's distrust negatively affects the accuracy and completeness of any resulting statistical analysis. Consequently, FEPA distinguishes two types of data uses: data collected for statistical and non-statistical reasons. Statistical use refers to data that the Federal Government uses as part of an analytic inquiry. Non-statistical use relates to information that the government uses for purposes other than statistical activities. Statistical data collection can be shared among agencies but may not be shareable outside the agency; nonstatistical data can be shared more broadly and has fewer restrictions.

Title III also places restrictions on how the government releases confidential personal information. Agencies may disclose confidential information only with the consent of the respondent and just for statistical purposes. Title III requires agency-head approval for each release, and the releasing agency is responsible for confirming there are no other restrictions to prohibit the agency from releasing the information. For example, agencies need to ensure that there are no other legal restrictions on a release, like a private contractor statutory limitation. To help drive home the seriousness of protecting confidential data, Title III assigns strict fines for nonconformance. Anyone who willfully discloses data to someone who is not entitled to receive will be guilty of a Class E felony and imprisoned for not more than five years or fined not more than \$250,000, or both.

## Statistical Efficiency

In today's fast-paced world, business, consumers, investors, and others are using data to inform their decisions better all the time. As part of their research, the Committee discovered that federal agencies encountered legal constraints that prevented federal agencies from sharing data and increasing the efficiency and efficacy of official statistical inquiries. At the same time, survey respondents began questioning whether the government could protect their confidential personal information.

The Committees realized that the quality of governmental research directly depended on the willingness of respondents to answer statistical surveys honestly. The Committee also recognized that

lowering the reporting burdens on federal agencies would lead to more accurate depictions of the economy and its health. For instance, if the Bureau of the Census (DOC), the Bureau of Economic Analysis, and the Bureau of Labor Statistics (DOL) share data more effectively, the Federal Government would have a better understanding of American businesses and how the business landscape was changing. Also, if federal agencies used uniform data standards when categorizing industries, the government could drill down to identify and resolve specific industry problems related to their agencies, be able to adjust for new businesses entering and exiting the U.S. economy, and be able to detect any irregularities that arise.

Last, Title III expands the amount of data each agency can collect and analyze. Title III also extends the number of industry classifications from 135 to over 800, providing the government with more resolution when reviewing the state of the economy.

## Designated Statistical Agencies

Part of Title III's goal is to increase the government's overall understanding of the United States economy, especially in areas reflecting the Nation's most important economic indicators, like income and product accounts. Accordingly, FEPA improves the semantic and syntactic similarity and accuracy of federal economic statistics that these agencies produce. Title III explicitly authorizes the sharing of business data among the Bureau of the Census (USCB), the Bureau of Economic Analysis (BEA), and the Bureau of Labor Statistics (BLS) for statistical purposes to reduce paperwork burdens on businesses. These three organizations are known as designated statistical agencies.

DSAs are doing a lot of the heavy lifting under Title III. These agencies (or divisions of agencies) produce and disseminate relevant and timely statistical information. Title III charges DSAs with conducting credible and accurate statistical inquiries and operating objectively. Following OMB's guidance, each DSA must establish policies, best practices, and other procedures to perform their work and produce evidence for policy-makers. Because DSAs and FSAs are responsible for holding the public's data, these agencies represent the last line of defense for protecting confidential information and making sure that data is used for statistical purposes, only.

While DSAs are being held accountable for what you would expect — the elimination of duplicate data, improvement of data quality, and cost control — they are also responsible for protecting the confidentiality of collected sensitive personal information. Specifically, DSAs must ensure that leadership, staff, and agents (e.g., academia and contractors) are aware of the importance of keeping

confidential information under tight control. In this case, awareness applies to any agency that utilizes confidential information. DSAs are responsible for ensuring that everyone fully understands their legal obligation to protect confidential information. This means that people having access to sensitive information must follow prescribed physical and electronic security procedures. These security procedures include keeping an audit log of everyone who accesses sensitive data. Finally, when DSAs share data, they must also conform to new sharing requirements, and written agreements reflecting binding terms. Only officers, staff, and agents who are parties to the data sharing agreements are allowed access to the data.

## Access to Data for Evidence

Title III presumes, DSAs have access to data as part of their approved analytical inquiries. For that reason, DSAs must help other federal agencies make copies of data upon request. Title III also requires that DSAs promptly respond to data requests and provide requestors with specific statutes that prohibit DSAs from sharing data with requestors. Also, Title III requires agencies to establish and follow regular and transparent processes to conduct interagency data exchanges.

## Expanding Secure Access to CIPSEA Data Assets

Not all federally collected data will be appropriate for the public, however. Some data contains sensitive governmental information that allows agencies to perform their missions. The Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA) was passed to handle these situations. This law provides strong confidentiality protections for statistical information collections that are sponsored or conducted by federal agencies.

FEPA also defines confidential information more specifically than previous laws. Title III requires federal agencies to produce a universal system to categorize the relative sensitivity and the corresponding level of accessibility for each data asset. The categorization system includes shared sensitivity levels, criteria for assigning sensitivity, rules for producing a lesser sensitive version of a data asset, and standards to improve access by redacting specific information. Title III also requires FSAs to perform a risk assessment for any data asset that the government intends to release to the public. FSAs need to ensure that the evaluation is easy to understand and available on the national data catalog. Additionally, FSAs must publish their standards and procedures on as well.

## Applying to Access Data Assets for Developing Evidence

Title III also directs OMB to develop a process through which the Congressional Budget Office (CBO), State, Local, and Tribal government researchers, and other individual users to apply for access to data assets covered under FEPA. The law explicitly requires OMB to ensure that FSAs produce identical processes, using a standard application form, criteria, timeframes for a response, appeals, and technical standards. Title III also requires OMB to develop guidelines so that FSAs develop the same processes in coordination with the public and private stakeholders.

## Analysis

New legislation can be tricky to understand and parse through, because it is uncontested. FEPA contains a number of new requirements and seems to create a number of interesting opportunities for federal agencies. As these new areas undergo rollout and implementation, some of the hazier details will become more solid. However, we believe key areas already exist that require additional insight and analysis, and we welcome more discussion.

### What Does This Mean?

Better data sharing among designated groups has a variety of benefits, not the least of which is less spending by government agencies to produce the same data — a big win for taxpayers. FEPA's long-term impacts are difficult to measure and will be heavily dependent on the law's practical implementation. However, what is clear is that we now have a substantive law and an opportunity to track any resulting issues and measure their relative impact going forward. For example, federal agencies must produce formal plans describing what open data they use to produce named outputs. To do this requires the Federal Government to manage its data and use those assets as part of the government's decision-making process, which we believe will affect the entire U.S. economy. Below are some potential effects the authors have already identified:

**Overall Impact Potential.** The Federal Government accounts for 33% of our national economy and has a considerable proportion of data assets. Because many organizations interact with federal data, FEPA will affect many organizations outside of government. What sort of consequences the new legislation creates and how those impacts affect different groups is not readily apparent.

**Greater Data Sharing Examples.** FEPA provides specific authority to DSAs to share data, allowing those agencies to access each other's data. The Federal Government can create more interoperability only through detailed and carefully controlled interoperation. The success of this effort can be used as a model to apply throughout other parts of the Federal Government.

**Knowledge Worker Value.** Presently, the Federal Government has trouble paying market rates for data expertise. As the government shifts to using data for more decision-making, there are opportunities for data personnel to receive higher compensation at rates comparable to industry rates. However, with the commensurate increase in demand for these skills, the government may be falling farther behind and unable to compete for scarce resources in the market.

**Data as a Gateway to Better Practices.** To be useful across agencies, the Federal Government must adopt a unifying measurement framework. If the government can better manage its data, many other things become possible including decision-making, process design, and standards. According to The Center for Data Innovation (CDI), open data has enormous value for businesses, journalists, academics, civil society groups, and even other government agencies. These organizations use this same data to develop innovative products and services, establish critical business decisions, conduct research, and ensure accountability and oversight in government. (Riley, D., 2019). All will be required to improve just to keep up.

**Uniform Confidentiality Approach.** When all agencies are sharing data in the same way, individuals and organizations are more likely to trust that the government is protecting their information, which leads to an increase in data sharing and cost reduction (Bean, R., 2018). It remains unclear whether or not agencies will fully conform to the same confidentiality rules.

**State Government Impacts.** State and local governments utilize federal data whenever they must through federal grants and other data-sharing mechanisms. As the Federal Government changes its approach to data and usage, there is an opportunity for state and local governments to change their approach as well.

**Remainder Impacts.** Though not explicitly established by the statute, we have seen instances with previous federal data mandates where other "knock on" effects can occur, including:

- Continued focus in the financial and other historically numbers-driven industries.

- Corresponding increases in training and other immersion opportunities for federal employees should result.
- The potential for research communities to focus more on the use of data instead of just new results.
- Additional experiences, benchmarks, and best practices to compare against the private sector.
- Governments uncover additional public-private partnership opportunities.
- Improved efficiencies and effectiveness provide better leverage the power of government in all forms.

## Potential Challenges and Unanswered Questions

As data advocates, we believe that leveraging data assets is a good thing. The fact that FEPA was passed so quickly compared to other federal data statutes is exciting. However, it's important to understand and track the impacts so that the government can make timely measured adjustments, and the Federal Government can continue to lead the way for state and local governments to drive more data usage in their constituencies.

Greater data availability benefits all citizens and American businesses, but it is not without consequences. Responsible rollout and management are critical at this exciting juncture. While the statutory changes show promise in many areas, our team remains concerned and sees some unanswered questions:

**Separation of Duties.** For the first time, the new legislation shows a clear separation of duties between the CDO and CIO. No longer is data a byproduct of technology and the responsibility of the CIO. Instead, data is the business, and how organizations use and share data are decisions that the business owns. More work needs to be done to ensure that technology does not eclipse these ideals.

**The Potential for Creation of Information Imbalances.** Information is power, or, to put it more finely, disproportionate access to information is power. While the new law gives the public greater access to massive amounts of government data, the ability to process and make sense of all that data will most likely not be equally available to all citizens.

**Taxpayers are the Product.** Based on some initial research, large companies, academia, and watchdog organizations are the most likely beneficiaries of FEPA. These organizations represent a tiny part of the

"public." On the other hand, companies like Experian, Equifax, Transunion, Facebook, and a host of other large companies have strong abilities to process and monetize government data to their advantage. Right now, those companies already collect a massive amount of data from citizens — making taxpayer data the product they sell to other organizations. With the passage of FEPA, all companies will have access to government data, which may either enrich the company products currently available or render them less valuable since the government will make the data available for free.

International Trade Imbalance. Foreign corporations will also have equal access to U.S. data under FEPA. Foreign access may give overseas enterprises a market advantage when understanding and selling to the American market. Few countries have open data programs that are as sweeping as the United States, and it is not clear whether the greater data availability will inspire other countries to do the same.

Implementing Standards. FEPA requires OMB to establish rules regarding the marking and handling of federal data. President Obama signed Executive Order 13556 in 2010 that requires that federal agencies classify data using standardized marks and tags to distinguish data that the Federal Government deems sensitive and worthy of control. It remains unclear how FEPA affects existing directives like EO 13556.

Unreasonable Performance Standards. Industry consumes federal data and incorporates that data into products and services. As demand for commercial services grows, industry makes various claims about performance, quality, relevance, and more to customers. Because the Federal Government is the originator of open federal data, the government could inherit costs associated with improved data timeliness without its knowledge or consent. So, while the law has not funded this sort of performance, industry's needs may influence the government's hand in the future.

No Additional Spending, But No Payoff Either. Governments charge taxes and tariffs on products and services, which have inherent value. Today, businesses sell data for a profit. If data has value, why can't the government receive additional revenue for it? By freely releasing data under FEPA, the government may be losing revenue to the detriment of the American taxpayer.

Lack of Objective Criteria for Personnel. What constitutes a qualified CDO is very unclear. Depending on the perspective of any agency, CDOs would likely come in all shapes and forms. The legislation specifies: "(b) QUALIFICATIONS.—The Chief Data Officer of an agency shall be designated on the

basis of demonstrated training and experience in data management, governance (including creation, application, and maintenance of data standards), collection, analysis, protection, use, and dissemination, including with respect to any statistical and related techniques to protect and de-identify confidential data.”

How these issues translate into a formalized credentialing system remains to be seen. Were we discussing a traditional and formalized profession like dentistry, we would have shared expectations. However, because CDOs are new and academia has not designed an underlying educational curriculum, those shared expectations do not exist.

In conclusion, as with any legislation, FEPA has positives and negatives, and the data community is poised to experience some exciting changes as a result. The possibility of this new legislation energizes us, and we look forward to an ongoing discussion. The International Society of Chief Data Officers sponsored this research. Readers can continue the dialog on the organization’s site at <http://iscdo.org>.

## Acknowledgements

We greatly benefitted from many reviewer comments. We especially thank the Bipartisan Policy Center's Dr. Nick Hart (one of the legislations strongest proponents) and Natalie Evans Harris, former Senior Policy Advisor with the President's Office of Science and Technology Policy and now BrightHive's Co-founder and Head of Strategic Initiatives for keeping our analysis on course.

## Authors

Peter Aiken is an acknowledged Data Management (DM) authority. He is an Associate Professor of Information Systems at Virginia Commonwealth University (VCU), past President of the International Data Management Association (DAMA-I) and Associate Director of the MIT International Society of Chief Data Officers. As a practicing data consultant, professor, author, and researcher, he has studied DM for more than 30 years. International recognition has come from assisting more than 150 organizations in 30 countries including some of the world's most important. Peter also hosts the longest running and most successful webinar series dedicated to DM (hosted by dataversity.net). In 1999 (before Google, before data was big, and before data science), he founded Data Blueprint, a consulting firm that helps organizations leverage data for profit, improvement, competitive advantage and operational efficiencies.

Todd Harbour Todd was the Chief Data Officer (CDO) for New York State and data leader for the U.S. Federal Government. In these roles, Todd led work to develop descriptive and diagnostic big data analytic solutions. He also worked with mission partners to produce data strategies, frameworks, roadmaps, data standards, and information architectures. Todd also led to interagency data sharing work and was instrumental in establishing a community-wide data analytics community of practice. Todd was also co-author of Data Strategy and the Enterprise Data Executive with Dr. Peter Aiken.

Kathy Walter is CEO of Nsoma, a company that solves the challenges of intersection points—technology, programming, and information and a consumer protection attorney with New York State. Kathy has a decade of experience managing IT projects, first as a COBOL programmer in the utility industry, then as an IT project manager in a variety of website consulting firms and the financial services industry. She also worked for nearly a decade in product management and

development as a technical product manager at Instinet Corporation, associate brand manager at Gillette and Proctor & Gamble, and as a senior product manager for Iron Mountain, before branching into education with the NYC Department of Education, information processing, and the law.

## References

1. Harris, N. (2019, January). Phone Interview with Dr. Peter Aiken.
2. U.S. Department of Justice. (n.d.). What is the Freedom of Information Act? Retrieved February 21, 2019, from <https://www.foia.gov/about.html>.
3. Ryan, P. (2019, January 14). H.R.4174 - Foundations for Evidence-Based Policymaking Act of 2017. Retrieved February 21, 2019, from <https://www.congress.gov/bill/115th-congress/house-bill/4174>.
4. U.S. General Accounting Service (GSA). (n.d.). About the Federal Funding Accountability and Transparency Act (FFATA) Subaward Reporting System (FSRS). Retrieved February 21, 2019, from <https://www.fsrs.gov/>.
5. American Statistical Association (ASA) (September 7, 2017). Commission on Evidence-Based Policymaking Issues Final Report. Retrieved February 21, 2019, from <https://community.amstat.org/blogs/steve-pierson/2017/09/07/commission-on-evidence-based-policymaking-issues-final-report>.
6. The Data Coalition. (n.d.). The Data Coalition. Retrieved February 21, 2019, from <https://www.datacoalition.org/>.
7. Hart, N., & Shaw, T. (2018, December 22). Congress Provides New Foundation for Evidence-Based Policymaking. Retrieved February 21, 2019, from <https://bipartisanpolicy.org/blog/congress-provides-new-foundation-for-evidence-based-policymaking/>.
8. Foundations for Evidence-Based Policymaking Act of 2018', §§ 101. Federal evidence-building activities-303. Increasing access to data for evidence (U.S. Government Printing Office 2019).
9. Riley, D. (2019, January 17). Trump signs law to open government documents to public scrutiny. Retrieved February 21, 2019, from <https://siliconangle.com/2019/01/16/president-trump-signs-new-law-opens-government-documents-public-scrutiny/>.
10. Bean, R. (2018, January 30). The Chief Data Officer Dilemma. Retrieved February 21, 2019, from <https://www.forbes.com/sites/ciocentral/2018/01/29/the-chief-data-officer-dilemma/#49344f53896a>.

