

Unleashing the Power of Financial Data in State and Local Governments

Written by C. Christina Ho and Dean Ritz

Introduction

Tracking the flow of money within the U.S. Government is difficult because the government's financial system is decentralized, with money flowing from the federal government to many state and local entities. However, if each level of government adopted *open data* principles, over time the government could track the flow of money between these disparate systems, thereby facilitating clear, transparent models of spending and performance.

This paper explains the circumstances for this present difficulty; introduces the principles of open data; provides a case study on the largest open data project completed by the Federal Government; and highlights key legislation (both present and pending) that mandates federal, state, and local governments adopt open data principles via *disclosure modernization* efforts. Lastly, it shows how existing software technology can enable this modernization effort to proceed in a decentralized manner.

This paper concludes by recommending the continued application of open data principles; that governmental agencies exercise their existing authority to speed their adoption; and that legislators continue advancing legislation that compels agencies to cooperate and apply their authority in order to, over time, unleash the power of financial data not just in federal government but also in state and local governments.

About the Authors



C. Christina Ho is the CEO & Founder of PolicyInsights, Inc., a public benefit company that helps governments use artificial

intelligence technology to make informed policy as well as spending decisions and to engage citizens better in those decisions. Formerly, she served as the Deputy Assistant Secretary for Accounting Policy and Financial Transparency at the U.S. Department of Treasury. There, she successfully led the implementation of the DATA Act with the goal to track nearly \$4 trillion dollars in annual spending and link data from multiple domains into one common standard.



Dean Ritz is the Senior Director of IP & Digital Reporting Strategy at Workiva, Inc., where he works in technical advocacy

roles for disclosure modernization and modern information architecture.

The authors wish to thank Christina Scheuer and Prof. Shannon N. Sohl for their assistance.

Organization of U.S. governments and their funding flow

In the United States, there are three levels of government: federal, state, and local. The U.S. Constitution establishes the authority and relationship of the federal and state governments. Under the Tenth Amendment to the U.S. Constitution, all powers not delegated to the federal government nor prohibited to the states are reserved to the states and the people. From a financial perspective, the federal government has significant impact on the finances of state and local governments. According to an analysis published by the Tax Foundation in 2014,¹ the federal government provides nearly 1/3 of the state revenue in the form of financial assistance. This amounts to more than \$600 billion each year. In addition, in 2016, the Pew Charitable Trusts reported that federal spending accounts for one fifth of the states' GDP.² This relationship introduces complexity and significant challenges in connecting spending to results because it requires tracking and integrating spending data at the federal, state, and local government levels.

History of financial reporting in U.S. governments

At the federal level, the Financial Report of the United States Government (FR) consolidates the financial statements of the government's component entities into a single government-wide report. Component entities include departments like the Department of the Treasury (Treasury), independent agencies such as the Securities and Exchange Commission (SEC), and government corporations or business-type entities such as the Pension Benefit Guarantee Corporation and the United States Postal Service. The FR has been published for over twenty years in compliance with the Chief Financial Officers Act of 1990 (CFOA).

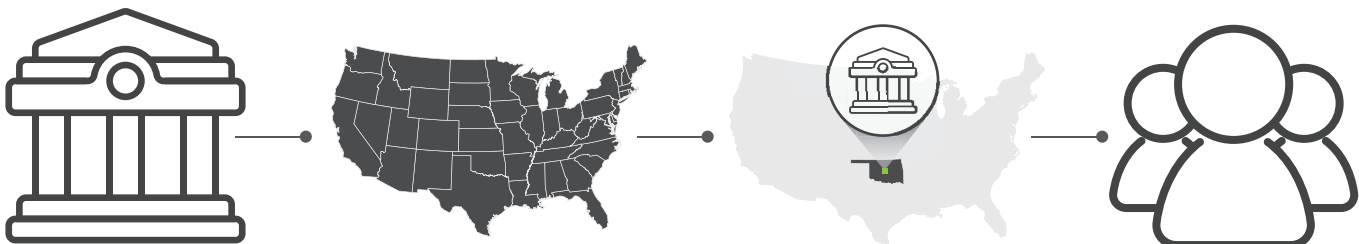
As a result of regulatory and legislative requirements contained in the CFOA and the Government Management Reform Act of 1994 (GMRA), the Secretary of the Treasury, the Director of the Office of Management and Budget (OMB), and the Comptroller General of the United States (who leads the General Accounting Office [GAO]) established the Federal Accounting Standards Advisory

Board (FASAB) in 1990 to develop the accounting standards that would guide preparation of the financial statements mandated by the CFOA.

While the federal financial reporting requirement was mandated by law, financial reporting for state and local governments is voluntary, motivated by their financing needs in the municipal bond market. The Comprehensive Annual Financial Report (CAFR) is governed by standards promulgated by the Government Accounting Standard Board (GASB), established in 1984. The CAFR's origin can be traced back to the 1940s, but it did not become nationally adopted until the 1970s. In 1945, the Government Finance Officers Association (GFOA) began awarding certificates of achievement for excellence to state and local governments that comply with the CAFR requirements. These certificates recognize the individual governments in their achievement of transparency and sufficient disclosure especially for municipal bond investors needing to assess the financial health of the issuers. As of the latest reporting period (2016), the GFOA has awarded the certificates to over 4,000 state and local government entities, including municipalities, counties, special purpose districts, and school districts. It should be noted that this recognition does not address the latency in the delivery of these reports. Whereas public companies must issue their annual financial report within 60 to 90 days of their year end (depending on the size of the company), entities providing a CAFR have 180 days to file.³

The flow of funds and outcomes

Generally speaking, and as illustrated below, the federal government provides funding to states and local governments. Local governments (and sometimes community-based organizations) are primarily responsible for delivering services to the people and for achieving outcomes. Given the distinct nature of the financial reporting in the public sector, in which all state and local financial reports are prepared separately with different accounting standards, it is currently impossible to use the FR or CAFRs to evaluate the performance of the governments in a systematic and meaningful manner.



Considering the number of states/territories (56) and local governments (>89,000),⁴ the diversity of the systems, and magnitude of the data captured within the public finance ecosystem, anyone attempting to extract data and translate it into meaningful information is presented with a significant challenge. In addition, since the financial reports for federal, state and local governments are not integrated, intergovernmental transactions are not consolidated or eliminated to avoid double-counting. For example, federal assistance to state and local governments is reported in the FR as costs and in the CAFR as revenue as well as expenditures. Additionally, the delivery of the outcomes are not reported as part of the financial reports.

In the private sector, the performance of companies are measured quantitatively in terms of profit. In the public sector, however, the performance of the organization cannot be easily measured quantitatively or holistically. Given the decentralized nature of the service delivery and the lack of standard performance measures, it is difficult to both collect the data and systematically interpret them. With the open data movement and the complementary advancement of software technology, data in disparate systems and various physical locations can be linked and analyzed. This creates opportunities for improving program and financial performance and providing broader transparency to the public.

Disclosure modernization

The first step to unleash the power of financial data is having open data. Open data is data standardized in structure and freely available. With open data, organizations can increase the speed and quality of expenditure analysis and draw more trustworthy conclusions on program performance.

The adoption of open data standards would be a great boon for many state and local governments because their current methods for producing financial reports simply produce static documents; that is, they produce 8.5" x 11" sheets of paper that limit information portability. Valuable data is locked up in PDF files, word processing and spreadsheet documents, and sometimes as scans—pictures of spreadsheets. It is rarely formatted even as data.

Many governments and regulators worldwide have already implemented disclosure modernization. Disclosure

modernization describes the move from document-based reporting to open data reporting. In the U.S., this move most notably includes successful investments by the SEC for quarterly and annual reports by public companies; the FDIC for Bank Call Reports; and the U.S. Federal Government with the implementation of the DATA Act, the first open data federal law in the U.S.

Further disclosure modernization mandates will affect the Bureau of Consumer Financial Protection, Commodity Futures Trading Commission, Federal Housing Finance Agency, and—most notably for government entities issuing municipal bonds—the Municipal Securities Rulemaking Board (MSRB).

Different types of organizations answer to different masters. Some reporting requirements fall upon private sector entities to ensure transparency to investors, health and fairness of the market, etc. Grant issuers may require federal aid recipients to submit reports showing spending or progress. Some public entities, such as municipalities, have reporting requirements contractually placed upon them by private market participants, e.g., by underwriters, in order to issue bonds and access that market. Lastly, some masters have influence rather than authority. Though it does not regulate municipalities, the MSRB, for example, can support disclosure modernization by leveraging its responsibility for facilitating a functioning municipal bond market.

Reporting requirements are meant to inform constituents and potentially influence the reporting organizations as well. The late Supreme Court Justice Louis Brandeis famously wrote that “sunlight is said to be the best of disinfectants.” For performance-based policy-making and good governance, that sunlight needs to fall upon open data. Good governance needs open data in order to go beyond “good intentions” and “educated guesses” into a new era of informed and carefully measured decision-making.

Case study: The DATA Act

The passage of the Digital Accountability & Transparency Act (the “DATA Act”) on May 9, 2014,⁵ marked the beginning of the Federal Government’s most significant step toward standardizing, linking, and publishing one of its most high value data sets.

The G8 open data charter published on June 18, 2013, defined high value datasets and noted that spending data is one of them.⁶

Although the Federal Government has been publishing spending information since 1789, it was the first time annual spending could be tracked and linked across multiple functions, including budget, accounting, procurement, and financial assistance (totaling nearly \$4 trillion in 2017). As a result, the public could have better insights into the who, where, what, why, and how much of federal spending. Previously, accessing such information would be costly at best and impossible at worst because all of the data resided in non-interoperable systems at over 100 federal agencies.

The DATA Act implementation took three years, but other large scale and complex government reforms have taken decades to implement. The implementation faced significant technical challenges, but the cultural challenges were even greater.

Technical problems faced in implementation

The law required over 100 federal agencies to report over 400 interconnected spending data elements—a *data vocabulary*—from hundreds of disparate systems to Treasury; the information was then published on the USA Spending.gov website launched in May 2017. The biggest technical challenge was therefore data integration. The challenge was met by a focus on integrating open data rather than integrating computer systems.

Historically, the Federal Government would take a system-centric approach by building a big system and mandating agencies report their data to a central agency based on some proprietary data standard. Once all data was incorporated into one system, it could then be published. This approach would cost more money and would take much longer because a new complex system would have to be built, and the agencies' systems would have to be enhanced to extract the newly required data based on the proprietary data standards. Many federal agencies still have legacy systems, and making changes to them (as assumed by a system-centric approach) is costly and time consuming.

As the leader of this large scale implementation project, Treasury, in consultation with a coalition of forward-thinking technologists within the Federal Government, wanted to instead apply innovative and proven methods and technology. Adopting a pure data-centric approach that requires agencies to embed data standards with the data itself at the source appeared to be the right technical solution. With this approach, the data can be validated at the source and then published and integrated

into a common data set representing the trillions in annual spending.

Unfortunately, sometimes the best technical solution is challenged to generate enough momentum to achieve the goal due to cultural resistance and gaps in understanding.

Cultural problems faced in implementation

The DATA Act was not a popular law when it was enacted. It was viewed by federal agencies as yet another reporting mandate that would add no value. In addition, the law provided no additional funding for the agency implementation. In that context, the initial resistance was understandable.

When driving large-scale changes in an organization like the Federal Government, navigating the people, policy, and politics is key. The Federal Government is a large and complex organization. Over many decades, layers of bureaucracies and processes were established with the goal of mitigating risks and chaos. As a result, the federal workforce's time spent on managing processes limited its opportunities to serve as proactive problem solvers. Business functions, policies, and even legal authority are often in silos. Integrating the data from various business functions, as required by the DATA Act, would therefore blur the lines of some institutional domains as well as highlight the gaps in skills and responsibilities. Imposing a technologically advanced solution, such as the purely data-centric approach, would generate even more resistance due to cultural and human capital barriers.

Applying open data principles

To accommodate the reality of the technical and cultural challenges, Treasury designed an implementation approach that focused on minimizing agency resistance by making it as easy as possible for agencies to provide the data. Agencies could provide the data in CSV format (a technical compromise appropriate at this stage for this project given the technical and cultural challenges; more appropriate purpose-built data encoding languages like XBRL are dominant in other disclosure modernization efforts). Treasury then built an open-sourced data broker, which is a data collection and validation tool, to consume the CSV files and validate and structure the data for publication. However, knowing that one of the core requirements and strategic values of the DATA Act is establishing data standards to ensure the reporting of reliable, consistent federal spending data for public use,

Treasury followed open data principles to develop the DATA Act Information Model Schema (DAIMS) and used it as the foundation of the data broker as well as the various consumption models. The DAIMS gives an overall view of the hundreds of distinct data elements used to tell the story of how federal dollars are spent. The DAIMS also provides clarity on how the public can better understand the inherent complexity of the data.

How the DATA Act could be used as a model for state and local entities

It would be impossible to build a system large enough to collect data from hundreds and thousands of organization in order to link the public sector finance data together. However, an approach similar to that taken by the DATA Act could be adopted at the state and local government levels to standardize and link their financial and performance data given their highly decentralized nature.

Although no one expects all U.S. governments to follow the same data standards—because data standards often have roots in policies as well as legal authority—it is reasonable to expect our governments to embrace the open and non-proprietary data standard principles and practices. As a result, data from disparate systems can be integrated efficiently and reliably to provide meaningful information. Over time, the public sector financial systems could become increasingly connected digitally. The possibilities will then become endless in terms of spurring innovation, such as creation of new technology products or new industry; benchmarking of performance across local, state, and federal governments; and making more informed decisions on policy and service delivery for the benefit of citizens and their communities.

The innovative approach taken by Treasury could be a model for future reforms and other entities, including state and local governments. State and local governments may have similar technical and cultural challenges. As illustrated in this case study, significant incremental progress in disclosure modernization could still be made if an entity first drives strategic and foundational changes and limits the impact of the changes to accommodate the reality of longstanding institutional processes and culture. For example, states and local governments that still have many legacy systems could use less technologically advanced formats such as CSV to extract the data first and then build modern tools to validate and publish the data. The key principle that cannot be compromised is the application of open data principles when developing the data standards.

Pending legislation and its role

Although more possible than ever, there are significant challenges to connecting public finance data, including lack of data standards, limitations in human capital, and procurement limitations which inherently hinder a government's ability to access the latest technology products and services. Organizations need to see the incentives of value before justifying difficult decisions. Given CAFR's value-driven reporting history, state and local governments should be leading the way in connecting and using the data to improve outcomes.

Nonetheless, congressional mandates are needed to achieve consensus among diverse stakeholders in embracing open and non-proprietary data standard principles. The chart below illustrates the relationships among regulators, standards-setting bodies, and reporting entities in both the private and public sectors for disclosure purpose. Congress is responsible for enacting laws which provide regulators and Executive Branch agencies legal authority to set specific reporting requirements for the relevant reporting entities. At times, Congress can mandate specific statutory reporting requirements (e.g., DATA Act).

The standards-setting bodies are, for the most part, independent of the reporting entities and are responsible for defining technical standards (e.g., accounting) to be applied by the reporting entities in compliance with the regulatory or statutory reporting requirements. For example, SEC, as a regulator, established the reporting requirements for publicly-traded companies such as annual and quarterly financial filings. FASB, as a standards-setting body, defined the accounting standards known as GAAP. A number of years ago the SEC also set the reporting requirements for public companies to submit their quarterly and annual financial filings (i.e., 10-Q and 10-K) in a data encoding language called XBRL as part of its disclosure modernization efforts. FASB currently also is responsible for updating and maintaining the GAAP taxonomy (in XBRL). Although many improvements are still needed, the disclosures of the SEC filers are clearly more advanced than the public sector. In the public sector, as noted in previous sections, financial and performance information are not integrated. Therefore, the opportunities are great to further modernize disclosures in the public sector and improve transparency and accountability as well as to enhance decision-making.

Fortunately, legislation already is pending in some areas where disclosure modernization is needed and can be

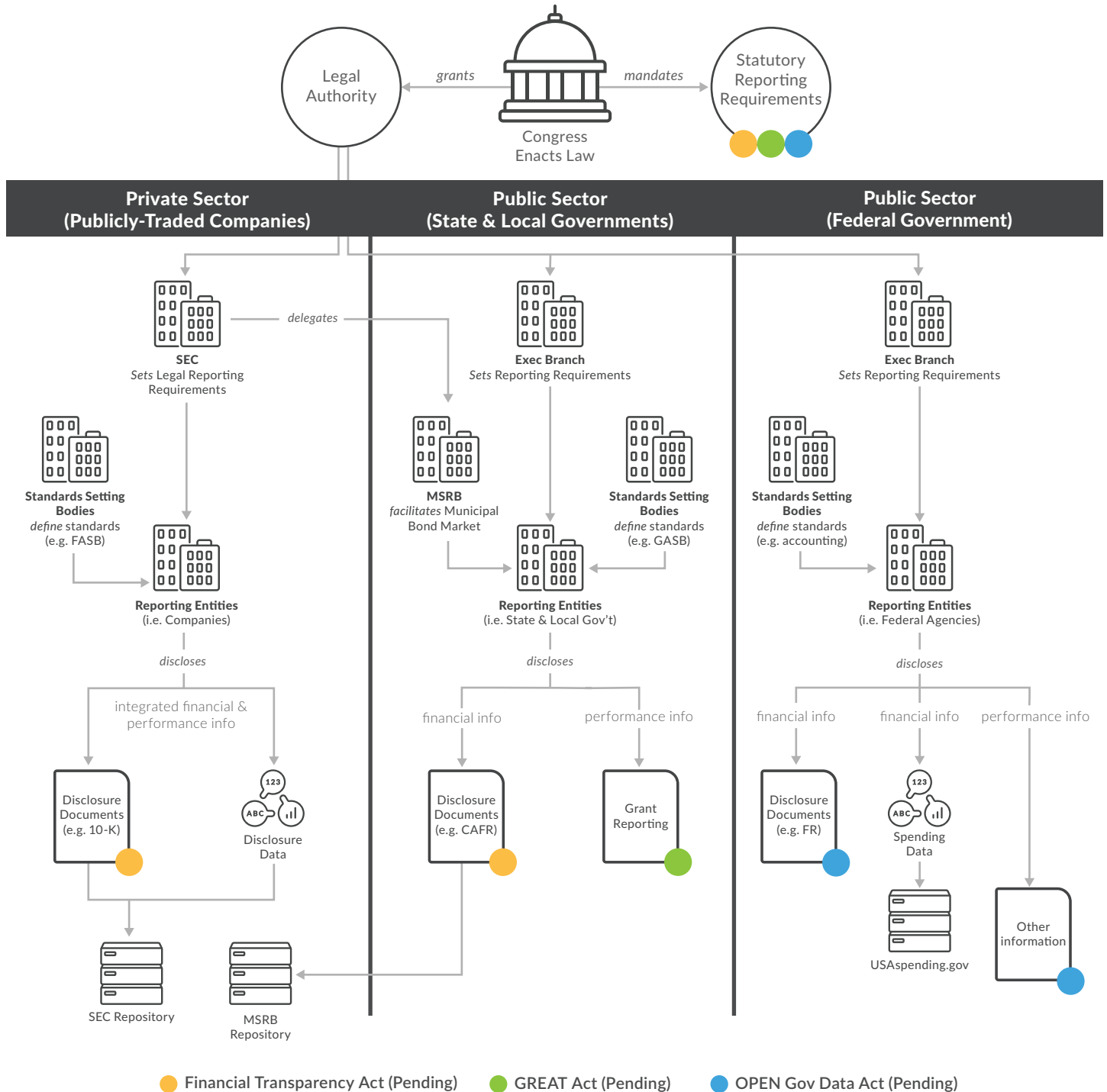
accelerated with a congressional mandate. Similar to the DATA Act, these pending legislations are advocated by Data Coalition⁷ on behalf of the private sector and the public interest for the publication of government information as standardized, open data.

Pending legislation: The Financial Transparency Act

In March 2017, the Financial Transparency Act (FTA) was introduced in the 115th Congress.⁸ The FTA requires that financial regulatory disclosures be filed and published

using consistent open data standards and formats. Similar to the DATA Act, the FTA has the potential to transform reporting and unleash the value of data. As part of the FTA, filings required and published by the MSRB will be subject to the data standards mandated in the FTA.

Although the state and local governments who issue municipal bonds are not regulated by the MSRB, they are required to file their annual CAFRs to MSRB as part of the continuing disclosure. Currently, the Electronic



Municipal Market Access (EMMA) website, which is provided by the MSRB, provides visitors access to all issuer disclosures, but most of these disclosures are presented as PDF documents. These include CAFRs with financial tables composed of scans of paper pages printed from spreadsheets; that is, they are bit-mapped pictures of spreadsheets. Surely, bond investors and analysts could use better data.

If the CAFR can be published as open data, investors will be better able to perform in-depth analytics to make well-informed investment decisions. Further, regulators will be able to monitor compliance and identify irregularities more efficiently and effectively. Finally, filers can reduce compliance costs by leveraging technology for data management and multi-purpose reporting. For example, in addition to CAFRs, states and local governments that receive more than \$750,000 in annual federal financial assistance have to comply with the OMB A-133 Single Audit Act. If source data for CAFR and Single Audits was the same, the compilation and audit processes for both could be significantly streamlined.

By mandating that federal spending be published using consistent and open data standards in compliance with the DATA Act and implementing that mandate successfully, Congress and the Executive Branch initiated and made significant progress toward the disclosure modernization effort in the public sector. These successes would not have occurred without legislation because transformational and large scale changes require a clear and uncompromising goal. Similarly, the FTA, if passed, will provide that goal and accelerate actions among diverse communities. Specifically, it will mandate regulators to use the act's legal authority to promulgate open data standards practices and motivate the issuers (i.e., state and local government) to implement them.

Pending legislation: The GREAT Act

In January 2018, the Grant Reporting Efficiency and Agreement Transparency Act (the "GREAT Act")⁹ was introduced into the 115th Congress. If passed, it will ensure the modernization of reporting by grant recipients. First, it mandates the creation of a standardized data structure for the information that recipients must report to federal agencies. And second, it defines goals for this new standard, including searchability, consistency with accounting principles, and the use of a non-proprietary syntax.

The executive branch (via the OMB) is assigned the task of establishing these data standards, and this legislation mandates that they materially consult with the grant recipient community and software providers to achieve this. The GREAT Act also mandates Congress' support for the standards already established by the DATA Act. As noted in the chart above, grant reporting is currently the only means the Federal Government has to collect performance data and to ensure that the grant recipients are achieving the results intended by the public investments of nearly \$600 billion per year. By standardizing the expenditure and performance data for all grant recipients, the GREAT Act would allow the Federal Government to provide transparent, timely information about how that grant money is used.

Pending legislation: The OPEN Government Data Act

Lastly, also introduced into the 115th Congress is the Open, Public, Electronic and Necessary Government Data Act (the "OPEN Government Data Act").¹⁰ It provides a sweeping, government-wide mandate for federal agencies to publish all their information as open data—using standardized, non-proprietary syntax. It also sets an official presumption that "Government data assets made available by an agency shall be published as machine-readable data... in an open format, and...under open licenses."

As illustrated in the chart above, this impacts all information produced by the federal agencies, including financial and performance information, in addition to what was already published by Treasury in compliance with the DATA Act.

Open data movement

The already passed and pending federal legislation is evidence that open data will reach all three levels of government: federal, state, and local. A closer look at the concepts of open data principles will introduce some operational and technical considerations for their implementation.

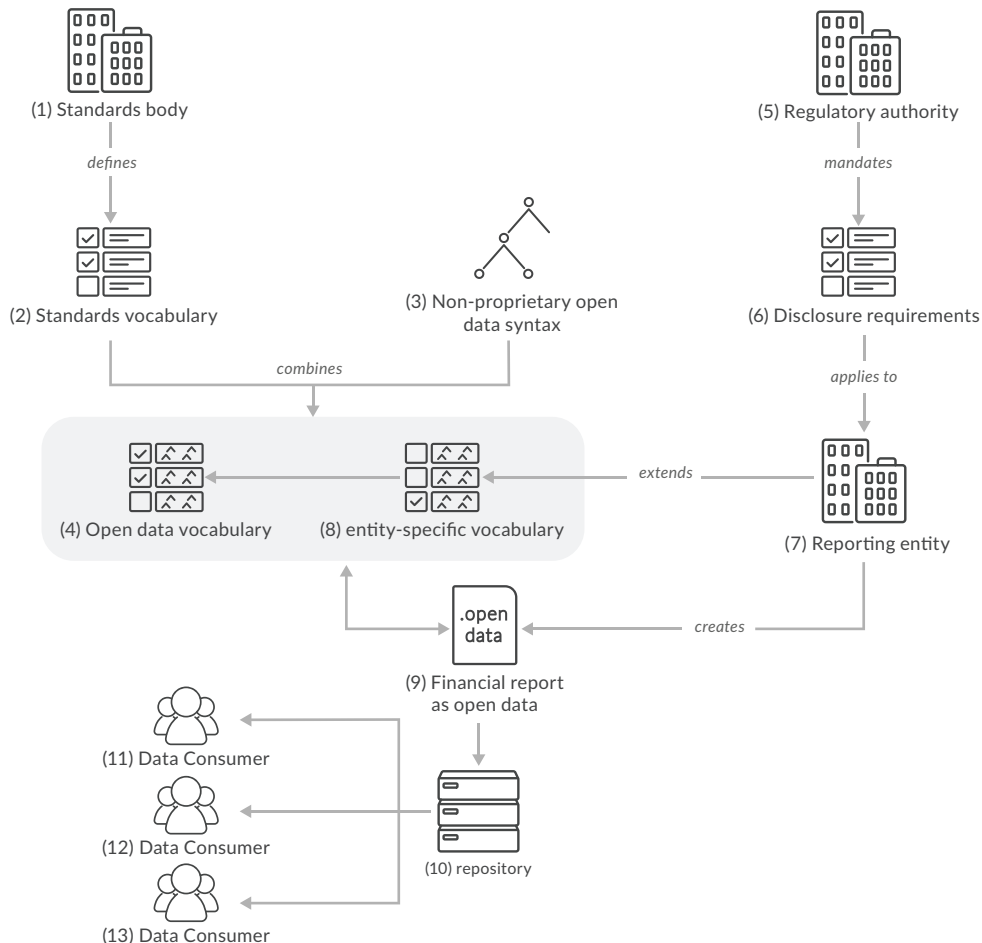
The idea of a movement is one that reorients perspectives and thereby reshapes the appearance of the landscape. We believe that the change in perspective from proprietary data to open data meets the definition of a movement because it will cause those who work with data—or wish they could—to see data differently. The open data movement may be expressed as responding to the follow expectations:

- a *shared understanding* of data meaning (unambiguous)
- *data accessibility* in a machine-readable form as data (not text)
- *automated data integration*
- *non-proprietary syntax* for the data and the data model

The following diagram and its description illustrates how these expectations may be met within a financial reporting regime.

- (1) Standards body. A standards body (authoritative or by common consensus) is the start of creating a shared understanding for open data. The body defines the standards vocabulary. Note: In some cases, the standards body also is the regulatory authority (5).
- (2) Standards vocabulary. Defined by the standards body, this vocabulary is more accurately portrayed as an ontology because it also captures the properties of and relations between the terms that comprise the vocabulary. These may be codified (as is U.S. GAAP financial reporting standards promulgated by FASB) or

- (3) Non-proprietary open data syntax. Not all syntax is created equal. By requiring that a syntax support open data we also and necessarily require a machine-readable syntax with expressive power to sufficiently represent a standards vocabulary (2) (which is why CSV files certainly are not satisfactory), and capable of modifications and extensions that minimize data incompatibilities as the standards vocabulary changes (which is why XML can be a problematic technical choice). The requirement that the syntax be non-proprietary reflects the governmental interest in transparency in data as well as in operations.

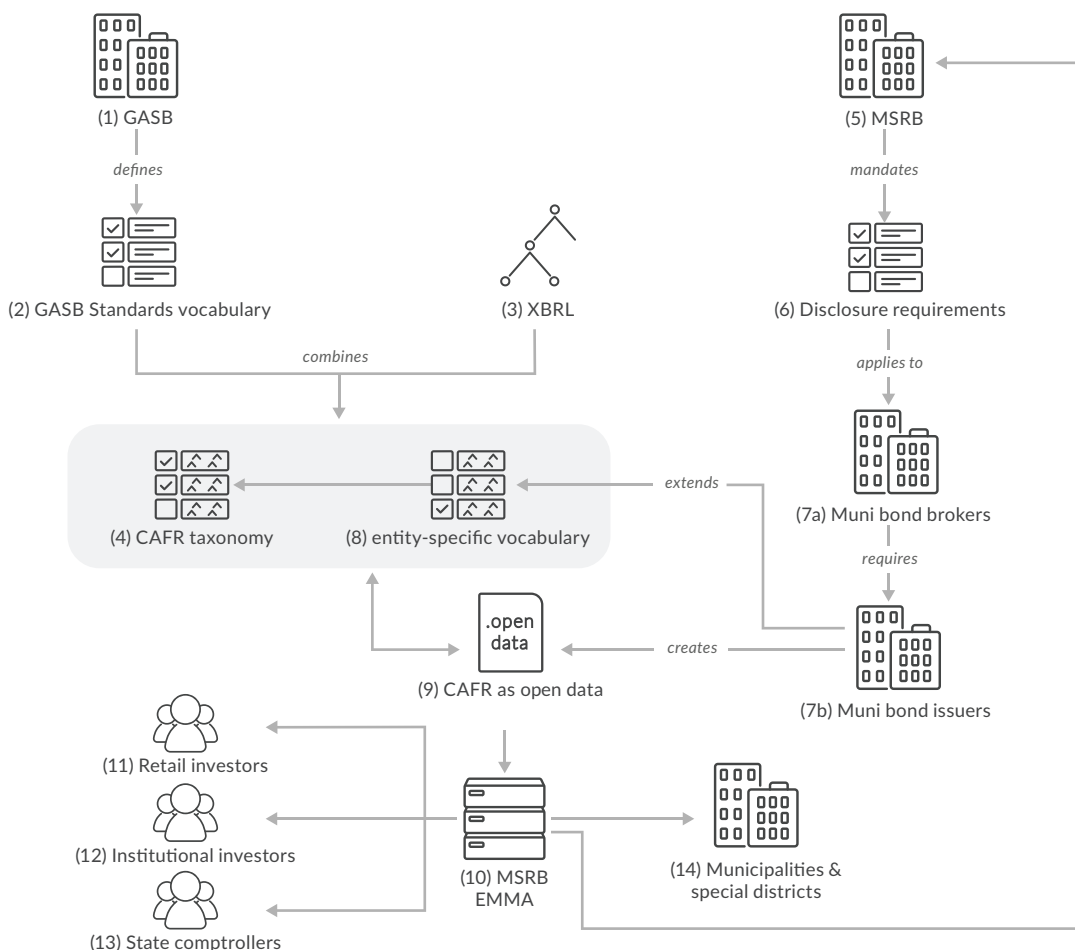


- (4) Open data vocabulary. This is the *reference implementation* of the standards vocabulary (2) in a non-proprietary open data syntax (3). The vocabulary is then used to create financial reports as open data (9).
- (5) Regulatory authority. The entity that mandates disclosure requirements (6) practices.
- (6) Disclosure requirements. The rules detailing the criteria requiring and the contents of required disclosures.
- (7) Reporting entity. The entity that is the subject of the disclosure requirements (6).
- (8) Entity-specific vocabulary. In some reporting regimes, the reporting entity (7) is permitted to extend the open data vocabulary (4). This may be necessary to capture entity-specific disclosure, such as the specific governmental activities (of a CAFR) or the business segments (in a public company's 10-K).

- (9) Financial report as open data.
- (10) Open data repository. May be hosted by the original intended recipient of the report, e.g., regulatory authority (4).

The following diagram and its description illustrates how these expectations may be met in a hypothetical reporting regime for CAFR to the MSRB.

- (1) GASB. In this hypothetical case GASB is the standards body.
- (2) GASB standards vocabulary. This is the body of rules defining Government GAAP.
- (3) XBRL. XBRL is a purpose-built language used worldwide in disclosure modernization projects. It is managed by the XBRL International standards body. Use of XBRL is royalty-free and non-proprietary.
- (4) CAFR taxonomy. An open data vocabulary implemented in XBRL, representing government GAAP as applied to CAFR disclosures.



- (5) MSRB. The MSRB is the authority mandating disclosure requirements for brokers (and others) participating in the municipal bond market.
- (6) Disclosure requirements. These are the rules detailing what must be disclosed and how to do so.
- (7a) Municipal bond brokers. Bond issuers (7b) use municipal brokers for access to capital markets, and in this relationship agree to reporting requirements.
- (7b) Municipal Bond issuers. The issuers are the reporting entities subject to the reporting requirements as agreed to with bond brokers (7a).
- (8) Entity specific vocabulary. For CAFR reporting, issuers would be expected to extend the vocabulary to support government-wide reports and fund statements for specific government activities and components. Some state governments may build state-specific extensions to cover state-specific reporting requirements.
- (9) Financial report as open data. A CAFR could be delivered in a format called iXBRL which combines the human-readability of HTML with the machine-readability of XBRL. iXBRL is a disclosure modernization language standard used worldwide. There exists a competitive market of software vendors offering products that enable financial reporting personnel to create XBRL and iXBRL formatted financial reports. Service companies that do this work on behalf of reporting entities are more numerous and also competitively priced. This report-creation technology has been in use, in some cases, for nearly a decade.
- (10) Electronic Municipal Market Access (EMMA) website. It would be expected that the MSRB would extend the capabilities of EMMA to support receiving open data financial reports (9).
- (11–14) Consumers of open data.

The data that state and local governments need to drive performance has complementary uses by other entities. The use of open data technology standards is critical to enable this modernization effort to proceed in a decentralized manner, over time.

Open data technology

Open data technology standards support most of the components in the above diagrams: the non-proprietary open data syntax (3); software to produce the open data vocabulary (4) and each entity-specific vocabulary (8); software to create open data financial reports (9); software to provide the functionality of an open data repository (10); software to extract, transform, and load (ETL) from open data reports (9) for use by data consumers (11–14); and optionally, as a matter of efficiency and design integrity, software to organize the standards vocabulary to facilitate the creation of the open data vocabulary (4). The first item mentioned—the choice of the open data syntax—is the primary consideration. The technology choices available for the other components depend upon the syntax.

Requirements for an open data syntax

Computer languages are numerous; however, in practice, XBRL and XML are, at present, the two languages suitable for open data reporting because they are the ones that best address the following requirements:

The syntax must be machine-readable in a way that contains explicit meaning. All languages have a syntax but in this case the computer must be able to read and act upon the encoded data as within the context of the subject-matter—the explicit meaning. This is quite different from the fragile representation of financial reports in spreadsheets (see requirement below) or for purposes of a Google search (i.e., text that is consumed and structured through machine-learning systems). For example, if a state wished to compare capital leases to operating leases across all its school districts, their data analysis will be much more dependable when performed on reports as open data (especially with the data vocabulary’s references to authoritative disclosure requirements) than for a machine-learning system to infer the data from a collection of PDF documents provided by various school districts in the state. In order to have trustworthy transparency and clarity, control should be applied to the way that data is labeled, categorized, and read by the machine. Reliance on a standard open data vocabulary enables each reporting entity to control the interpretation of its report.

The syntax must be suited to express financial reporting information. First, it is important to distinguish between computer languages that instruct the operations of a

computer from those that describe a data vocabulary. The former are programming languages while the latter are data encoding languages. In this case, the language chosen must be the latter—a language suitable for data encoding financial reporting information. The language should enable the description of core finance-relevant properties such as balance type; period type (e.g., as-of date, or a period of time); and properties for how a value reported in one term (e.g., assets) foots with other terms (e.g., current and non-current assets).

The syntax should provide for a non-fragile representation of the data. Interpretation of a data point should not be dependent on the location of the data within a document. Most spreadsheets create a condition of data fragility in which the interpretation of a value in a spreadsheet cell depends upon the human reader associating that value with contextual information in adjacent cells. Thus, moving that value to another location dislocates that value from its context. In a standard open data system, data should be encapsulated with its meaning so that even if the data value moves to a new location its context travels with it. Ideally, the data representation itself is *normalized*. Normalization maintains data integrity by isolating assertions of fact from how facts are related to one another. For example, you may want to sequence assets, current assets, and noncurrent assets (in that order). The sequence should be able to be reversed without reorganizing the data. Simply, the description of relationships between facts are kept separate from the fact assertions.

The syntax should provide for serialization. Serialization is a technical feature that enables the storage of a set of data. Data can, for example, be stored in a file. In an open data system, the requirement for serialization applies to the entirety of both the data and the data vocabulary. This is critical because open data reports may utilize entity-specific vocabulary (8), and these need to be bundled with the data if the data is to be correctly interpreted. As a practical counterexample, imagine the difficulties arising from having to submit database files from an enterprise reporting platform.

The data vocabulary represented in the syntax should be extensible. A reporting entity may wish to customize a label for when a term from the data vocabulary is presented to a data consumer. Further, a reporting entity may be required to add terms thereby creating an entity-specific vocabulary. For example, a municipality reporting an open data version of its CAFR will need to define terms

for their specific government activities, their business type activities, and their discretely-presented component units. Both examples are considered as *extensions to the reference implementation* of an open data vocabulary. An extensible data encoding language enables these extensions without threatening the integrity of the data vocabulary. Therefore, an extensible language provides a necessary combination of stability and flexibility, both of which are particularly important in a complex public system.

It should be a non-proprietary syntax with an active marketplace of supporting software and service providers. The perfect syntax would be non-proprietary, available for royalty-free use, and actively managed by a not-for-profit standards body.

XML & XBRL

XML is an acronym for “eXtensible Markup Language,” and it is used to encode documents with structured data. However, in regard to its suitability for open data reporting, it falls short in two key requirements. First, its syntax does not make it particularly suited to expressing financial reporting information. It is a general purpose data encoding language, and encoding financial reports is not its intended purpose. It can be extended to represent financial reporting ideas, but these extensions require agreement from all data consumers if they are to have reliable meaning. Second, it falls short of the requirement that it should provide for a non-fragile representation of the data. XML interlaces data and the data vocabulary, and thus changes to the vocabulary have a tendency to break the data.

XBRL is an acronym for the eXtensible Business Reporting Language. It has been in use since 2003 with widespread application for financial reporting.¹¹ It is based on XML and was designed to overcome critical limitations encountered by early prototyping efforts to use XML for quarterly and annual financial reports of U.S. public companies.

As noted earlier, it is necessary that different governmental entities be able to move independently of each other. The risks of independent action are mitigated by applying open data principles and making technical choices that respect the above requirements. These are durable investments that anticipate and prepare for meeting the open data expectations dominating the movement for disclosure modernization.

This approach enables governments to create large-scale, meaningful financial models without altering the

fundamental decentralized organizational model. That is, state and local governments could maintain their own systems and rights while also contributing interpretable, transparent data sets. These standards would allow the U.S. government to find an appropriate balance between unity and independence.

Proprietary systems, while familiar practice, have not shown to be as cost-effective or as useful as what governments worldwide have realized with open data technologies such as XBRL and XML. There are tremendous amounts of information available about these two languages which expand on these assertions.¹²

Conclusion: Open data standards should be standard

Standards create efficiency. Standards create markets. And with regard to good governance, standards convert data into critically useful information. The best standards provide for their own evolution, both in real-time via purpose-specific extensions to the standard and in the long term via an active standards body that manages the standard. Here, governments can follow the model of the Internet. The Internet, itself a technical standard, exploded in power and benefits when the networking protocol it established was extended by additional standards such as HTML and the world-wide web protocols. Before the web, the Internet was used to communicate data in proprietary formats. It is an irony probably not lost of the inventors of the Internet.

This speaks to the importance of standards. Continuing with the example of the world wide web, the explosion in benefits would not have been possible if every company's web browser had a proprietary syntax for representing a web page. In fact, sometimes companies introduce proprietary extensions to standards, deliberately introducing incompatibilities to derail the benefits of standards.¹³ Ultimately, the benefits are larger through cooperation via standards than they are with proprietary technologies.

Nearly all of the technology needed for standards based on open data principles already is in use and proven, and a substantial amount of this technology is available for royalty-free use. Standardized technology also increases price competition by both technology and service providers. It also enables governments to switch service providers as needed without having to start from scratch.

Where do standards come from?

In reporting to and by governments, standards arise from appropriate authorities. Governmental agencies already have existing authority to mandate disclosure modernization, and they should ensure that this move is based on open data principles. The disclosure modernization efforts by the Federal Government and the Securities and Exchange Commission are examples of this authority exercised. However, these efforts needed support from legislation to compel compliance to the authority. This is not unreasonable. Therefore, legislators should continue advancing legislation that compels agencies to cooperate and apply their authority for disclosure modernization so that financial data can bring performance-based benefits to federal, state, and local governments.

State legislatures should be active here, too. Most states have widespread opportunities for disclosure modernization that would enhance state oversight and the performance of those subject to oversight. A disclosure modernization project could focus on one reporting regime, e.g., state utility commissions, state insurance commissions, state universities, etc. A more significant disclosure modernization project with CAFR reporting standards¹⁴ would more significantly improve states' abilities to monitor and evaluate the financial health of the state as a whole or through specific lenses such as by type of government, location, types of programs, size of government, etc.; the state could better identify those entities with on-going risk concerns. Those subject to oversight may appeal to trust but it's best to verify.

Disclosure modernization isn't just about technical standards. Reporting standards are critical too. They are the keystone to the structure of open data. The first step for states to take would be to mandate compliance with GASB's GAAP reporting. It's not unreasonable that if an entity meets the threshold for an audit then they should, by default, report based on government GAAP.

This leads to the next suggestion for action by states: advocate for the GASB to update their standards to be more easily represented as an open data vocabulary, at least through codification of the standard. As a complement, and at the same time, states should unite and require the GASB to simplify the reporting requirements. It is an enormous undertaking for small and mid-sized entities to prepare a CAFR.

States can learn from the SEC example in which disclosure modernization for public company reporting led to use the use of new technology with an impressive result: the cost of preparing the paper and XBRL versions of a quarterly and annual reports (with appropriate technology) was substantially less than the cost of preparing the paper report alone (with conventional document editors and spreadsheets). To the surprise of many companies, the move to new reporting software due to the XBRL requirement resulted in substantial process improvements and reduced the cost of preparing and filing those reports with the SEC.¹⁵

We look at data as a source of knowledge. It can support our capacity for judgment. Without it, we're left with educated guesses. Given the importance of good governance, it's better to measure than it is to guess. If each level of government adopted open data principles, over time governments could assemble the data they need to track the flow of money between the disparate systems of the numerous government entities, thereby facilitating clear, transparent models of government spending and performance.

Endnotes

- 1 <https://taxfoundation.org/states-rely-most-federal-aid/> (accessed on March 12, 2018).
- 2 <http://www.pewtrusts.org/en/research-and-analysis/analysis/2016/03/22/how-closely-are-state-economies-tied-to-federal-spending> (accessed on March 12, 2018).
- 3 Sohl, Shannon N., Waymire, Tammy R., and Webb, Thomas Z. "Determinants in Illinois' Local Governments' Reporting Lag." *Journal of Emerging Technologies in Accounting* (2018, in press).
- 4 <https://www.census.gov/newsroom/releases/archives/governments/cb12-161.html> (accessed on March 12, 2018).
- 5 https://www.congress.gov/113/plaws/publ101/PLAW_113publ101.pdf (accessed on March 12, 2018).
- 6 This assertion combines the fact that the US government is the largest national government in the world (see <https://www.cia.gov/library/publications/the-world-factbook/fields/2056.html>) with the recognition of criticality of data about these expenditures (see <https://opendatacharter.net/resource/g8-open-data-charter/>).
- 7 www.datacoalition.org
- 8 <https://www.congress.gov/bill/115th-congress/house-bill/1530> (accessed on March 12, 2018).
- 9 <https://www.congress.gov/bill/115th-congress/house-bill/4887> (accessed on March 12, 2018).
- 10 <https://www.congress.gov/bill/115th-congress/senate-bill/760> (accessed on March 12, 2018).
- 11 See this comprehensive paper on the the use of XBRL at www.datafoundation.org/xbrl-report-2017 (accessed on March 12, 2018).
- 12 Ibid.
- 13 <http://www.nytimes.com/1997/10/08/business/sun-sues-microsoft-on-use-of-java-system.html> (accessed on March 12, 2018).
- 14 Snow, Neal M., and Reck, Jacqueline L. "Developing a Government Reporting Taxonomy." *Journal of Information Systems* 30, no. 2: (2016): 49-81.
- 15 As reported by Workiva Inc., the leading company in SEC compliance reporting. As one example see <https://www.workiva.com/customers/wdesk-makes-life-easier-american-equity> (accessed on March 12, 2018).