Executive Summary

The Digital Accountability and Transparency Act of 2014, the “DATA Act,” is poised to create, in May 2017, the first open data set covering all federal spending, promising to vastly improve both public transparency and internal efficiency. [1] The DATA Act’s required steps should be completed by 2022.

Interviews with more than 20 DATA Act stakeholders – from the Office of Management and Budget (OMB), the Treasury Department, and other agencies as well as contractors and congressional staff members – provided insights into the law’s immediate future and ultimate aims. Initial expectations are modest, given the outlook for incomplete reporting and varying data quality at first. Even so, the data to be reported should provide a valuable resource for agencies and the public.

The interviews illuminated an ambitious vision for the transformation of federal spending. If properly implemented, new data standards and new reporting requirements will provide federal agencies, watchdog groups, grant and contract recipients, and the public with standardized, readily comparable spending data from across the federal government. Tasks that today require time and money will be automated. Waste, fraud, and abuse will be illuminated and potentially reduced. The government’s responsiveness to its people will improve, perhaps breeding greater trust.

Our interviews identified key challenges to this vision of freely flowing and interoperable spending data – and ways to mitigate them. The challenges fall into two broad categories: cultural, since the DATA Act requires agencies to change the way they have “always” reported spending, and technical, since the DATA Act’s standardized formats and fields need key improvements to support its ambitious vision.

The DATA Act is the nation’s first open data law – and the first phase of a larger effort to transform all aspects of federal reporting to produce freely flowing, interoperable and readily comparable data. Lessons from the law’s implementation can inform similar transformations in areas beyond spending, in performance, regulatory reporting, records management and more.

State governments and other nations are watching. As they attempt to realize the vision of the DATA Act, the law’s supporters in the executive branch and Congress can expect their efforts to reverberate across government, society, and history.
Introduction

When the Digital Accountability and Transparency Act, or DATA Act, became law on May 9, 2014, its ambition was clear: to change how the U.S. federal government collects, tracks and uses information about spending. [2] The law promises to benefit federal agencies, funding recipients, Congress, agency inspectors general and, most importantly, the American public, by providing open, standardized data on federal spending online, starting in May 2017.

The DATA Act establishes three basic mandates. First, it charges the Department of the Treasury (Treasury) and OMB with creating and maintaining standard data elements and formats for all federal spending reports submitted by agencies, [3] and directs the agencies to follow these data standards. [4] Secondly, it directs Treasury and OMB to assemble the information reported in this way as a unified, open data set. [5] Finally, OMB must conduct a pilot program (the “Section 5 Pilot”) to test the feasibility of imposing data standards on the reports submitted by recipients of federal grants and contracts.

Treasury, OMB, and personnel from across the federal government have worked to meet these mandates for nearly three years. They have agreed to an initial set of standard data elements and used them to create the DATA Act Information Model Schema, or DAIMS. Treasury has built open-source software, the DATA Act Broker, to help agencies report their data encoded in the DAIMS. As of May 2017, all federal agencies will begin using the DATA Act Broker to report their spending data. [6]

OMB split the Section 5 Pilot into two parts, one focused on grantee reporting, managed by the Department of Health and Human Services (HHS), and the other an OMB project for contractor reporting. The grantee reporting pilot is expected to meet its deadlines, but OMB has delayed the contractor portion while shrinking its scope. [7]

The DATA Act’s final deadline falls in November 2021; unless the law is amended, implementation will be fully complete in 2022. [8] This paper considers what happens next, and recommends ways to address the challenges to come.
A Vision for Federal Spending in 2022

By 2022, when the statutory deadlines are complete, the DATA Act’s ultimate impact should be evident: federal spending data should flow automatically from its originators to all users.

What might this vision look like? If the challenges described in this paper are successfully addressed, federal spending could change in the following ways.

Federal chief financial officers (CFOs) and their offices will not need to manually compile spending information received from multiple sub-units of their agencies. Instead, these roll-ups will occur automatically.

Auditors will not need to perform manual calculations of any kind. Instead, all calculations and models will be applied automatically, within their software. Federal agency heads will be able to see the destination of funding from any given appropriations account, the funding sources for any given program, payments made under any grant or contract — and how all these things would change under any congressional appropriation scenario.

The administration may be able to generate and compare financial metrics without manual research. Government-wide analyses that currently require either sampling or extensive data calls, such as category management and the Government Accountability Office’s (GAO) annual duplication reports, will become trivial exercises.

Inspectors general will use “dashboard” interfaces that automatically receive and retrieve all spending information down to the individual payment, from their own agencies and from comparable programs at other agencies.

Congress will craft appropriations bills electronically, with analyses automatically showing legislators how each choice would affect grants, contracts and funding levels in highly granular detail.

Grantees and contractors’ financial systems will automatically prepare and submit the information for agency and government-wide reports. And citizens will be able to quickly view breakdowns of federal spending in their communities, in grants, contracts and direct spending on offices and installations — and visualize the impact of each vote their legislators take on the whole picture.

Anyone, legislators and policy wonks and people who just want to know, will be able to track a program’s reported performance against its funding, period by period.

By 2022, if all goes well, spending information will be interoperable: consistent and comparable, across programs, bureaus and agencies. As a result, the systems needed to manage federal finances and awards will be cheaper to build and upgrade, perhaps with open-source software. Decisions about spending, performance, and program management will be driven by built-in analytics, without the need for special data calls. The American public will be able to electronically navigate from an enterprise-wide view of federal spending all the way down to programs, awards, and even individual transactions.

By 2022, if all goes well, spending information will also be automated: reported, exchanged, and audited without manual interference. Systems will deliver reports instantly; grantee and contractor software will communicate automatically with agency and government-wide systems.

Automation will reduce the cost of government at every level while improving its reliability. More importantly, perhaps, everyone – government officials, public-interest groups, academics, the press and private citizens – will be using the same information resources, a factor that ultimately may improve and solidify trust between government and society at large.
Seven Challenges to the DATA Act’s Success

The work spurred by the DATA Act has come a long way since it was signed into law in 2014, but the vision of freely flowing and interoperable spending information faces significant challenges. These fall into two basic categories: cultural and technical. The cultural challenges require full engagement with stakeholders in the development of new processes and applications. The technical challenges are related to the DAIMS architecture and the creation of a complete picture of federal spending.

Of course, the cultural and technical challenges are equally significant. Representing the entire federal spending cycle in a unified, open data set will be a significant technical breakthrough, for example. But the programmatic changes needed to accomplish that will depend largely on the engagement of all the stakeholders involved.

Cultural Challenges

1: Some agencies may perceive the DATA Act as a bureaucratic requirement rather than a truly new approach.

According to our interviews, many agencies are having trouble seeing the DATA Act as anything more than an elaborate compliance exercise. Their financial officers are focused much more on meeting the DATA Act’s immediate requirements than on how the resulting data ultimately can be used. This focus could pose a significant threat to the project’s future.

If agencies approach the law as a bureaucratic burden rather than a genuine internal benefit, public interest organizations, journalists and researchers are likely to distrust the data — and with good reason.

The original online spending transparency platform, USASpending.gov, provides an important lesson. The Federal Funding Accountability and Transparency Act of 2006 (FFATA) required OMB to publish the contents of government-wide grant and contract databases assembled from agency reports. [10] Because the reporting agencies did not use these government-wide databases themselves, however, they had no incentive to guarantee the accuracy of the data — and its quality suffered. [11]

Case study: DATA USA

Spending data as part of a larger intelligence capability

Standardized federal spending data on its own is useful for agencies, watchdog groups, and the public. But it is even more valuable in combination with other data.

Consider DATA USA, a collaborative project from Deloitte, DataWheel and Macro Connections. DATA USA combines public data sourced from US federal and state agencies to visualize jobs, skills and education by geography. A unified federal spending data set, if combined with DATA USA and similar multi-source, geospatial data platforms, could show relationships between federal spending and overall economic and demographic trends.

The DATA Act’s structure, with its common definitions for agencies, programs, and other concepts, could provide a model for the standardization of other federal data, easing future integrations.

Similarly, the DATA Act should improve the offerings of companies that help contractors navigate the federal procurement system. These companies leverage spending information to provide insights and strategic advice, and are excited about the improved capabilities the DATA Act will provide.

For more information: https://datausa.io/
The national public-interest community – transparency advocates, academics, journalists and others – provided vital early support for the DATA Act by exposing deficiencies in the existing USASpending data, lobbying Congress, and connecting the nascent idea of open data with the complexities of federal spending.

Unfortunately, the implementation of the DATA Act is so complex in detail and expansive in scope that many stakeholders simply can’t provide as much input and support as the Act’s implementers want. Several interviewees within the federal government expressed their desire to engage more extensively with stakeholders, and expressed disappointment that they have not been able to do so.

Our interviewees acknowledged that the governance structure set up by the DATA Act, with Treasury and OMB solely responsible for developing and managing data standards, achieved the law’s basic mandates. They also agree that a different approach may be needed to ensure the DAIMS’ flexibility and stability in the future. Similarly, GAO has noted that, despite initial OMB and Treasury efforts to convene “a committee to maintain established standards and investigate new standards,” much more should be done. [12]

Simply put, Treasury and OMB do not have the resources needed to permanently manage the DAIMS. Treasury funded its initial stand-up as part of the Treasury Fiscal Service, and Congress appropriated dedicated funding in fiscal years 2016 and 2017 for the implementation. To expand the DAIMS as contemplated later in this paper, however, these resources may be inadequate.

Case study: OhioCheckbook.com

Improving inside and out

Until 2015, Ohio had been a low performer among states in digital budget transparency, according to the Public Interest Research Group’s surveys, most recently Following the Money 2016, published in April 2016. That changed quickly with the launch of OhioCheckbook.com, which boosted Ohio to the top of the list. The interactive spending transparency platform, an initiative of Ohio Treasurer Josh Mandel, allows the public to navigate from budget categories all the way down to specific expenditures. It has also delivered significant value to those working inside government. The Treasurer’s office reports that its own staff uses OhioCheckbook to perform tasks, rather than relying on internal tools.

In addition, the Treasurer’s office made the platform available to Ohio municipalities free of charge, allowing local governments to deliver transparency to their residents and a new management tool for themselves.

Ohio’s transformation was greatly eased by the fact that the state’s financial data had already been standardized by the implementation of a single, state-wide financial management system several years earlier. Government-wide system consolidation is impossible on the federal level – hence the need for the DATA Act’s mandate to standardize data while leaving disparate systems in place.

OhioCheckbook.com perfectly illustrates the virtuous cycle that can begin with increased spending transparency. Better data, interactively presented, leads to more engagement from both external and internal stakeholders, which can help identify data quality problems and inefficient spending, which can lead to even better data and analysis.

For more information:
http://ohiocheckbook.com,
http://www.uspirg.org/reports/usp/following-money-2016-0
Technical Challenges

4: Legacy reporting systems may reduce data accuracy.

The DAIMS assembles two broad categories of previously siloed spending information: financial information maintained in financial management systems and award information maintained in grant- and contract-writing systems.

The DAIMS and its reporting processes require agencies to extract financial information directly from their source systems. The DAIMS does not require the same level of detail for award-related information, however. Rather than extracting award information directly from grant- and contract-writing systems, agencies need only submit the same compilations they already provide to legacy databases predating the DATA Act. [13] These government-wide databases, primarily the Federal Procurement Data System-Next Generation (FPDS-NG) and the Federal Assistance Awards Data System (FAADS), have well-documented accuracy problems. [14]

GAO has found that Treasury and OMB’s decision to rely on the FPDS and FAADS databases – rather than requiring agencies to extract award information directly from source systems – “raises concerns about the quality of the data” that will result. [15] In addition, agencies may find it difficult to combine source-system data with previously compiled award reports.

When agencies begin to report data under the DATA Act in May 2017, the data should be generally consistent with the originating systems. Since GAO already questions the accuracy of agency submissions to legacy award databases, we can expect to see similar issues may exist with DATA Act award reporting.

Accuracy is, of course, the key component of data quality, and vital for decision-making, research and oversight. Unless the mismatch between the way agencies compile financial and award information is corrected, accuracy problems may persist for years.

5: Overly vague definitions may imperil data comparability.

GAO also noted that several of the established DATA Act definitions could lead to inconsistent reporting. [16] In particular, GAO singled out the definition for Award Description and Primary Place of Performance as “imprecise and ambiguous.” For example, the “primary place of performance” for highway projects could be interpreted as the mileage marker closest to

---

Case study: Small Business Administration

Better Visibility Leads to Quality Improvements

The Small Business Administration (SBA) worked with Treasury and the General Services Administration’s (GSA) 18F team to test the DATA Act Broker. This gave the SBA an early look at the power of a single, unified set of financial and award data. For the first time, the SBA’s Chief Financial Officer (CFO) has visibility into award-specific details and the ability to analyze award characteristics across all SBA programs.

This increased visibility has already helped the SBA identify and address several data quality issues. Funds that had been improperly identified as specifically allocated when they really belonged to a discretionary account have been reclassified. Issues with miscoded place of performance information that indicated the agency was spending millions of dollars in the wrong location have been fixed. These problems would have gone unnoticed if the SBA had not begun to embrace and analyze its own data.

The SBA’s Chief Financial Officer was initially skeptical of the DATA Act’s transformative power, but he quickly changed his mind when he saw the immediate results. When all agencies begin reporting in May 2017, they should learn from the SBA and consider using DATA Act data to improve their overall data quality.

For more information:
the actual work; the contractor’s headquarters; the state capital; or a city hall. [18] A second definitional issue concerns the DATA Act’s mandate for a “single award identifier,” a common data element, centrally managed across all agencies, to identify every award of a grant or contract. This objective has not been met, again due to deference to existing award reporting systems.

6: The government may make future changes to the data field that is used to identify entities that receive grants, contracts, and other awards.

Under a General Services Administration contract, the federal government currently uses the Data Universal Numbering System (DUNS) Number as the exclusive identification code for grantees and contractors receiving federal funds. The DUNS Number is proprietary. In February 2017, the GSA issued a Request for Information asking for input on whether it should adopt an identification code with fewer licensing restrictions. [19] When building the DAIMS, Treasury and OMB chose to continue using the DUNS Number as its official identification data field, but an alternatives analysis may be conducted in the future to evaluate other possibilities. [20]

7: The DAIMS does not reflect all stages of the federal spending lifecycle.

Federal spending information can be viewed as an annual lifecycle, beginning with the receipt of revenue from taxes and other sources and proceeding through the President’s proposed budget and congressional appropriations, through Treasury’s processes and finally resulting in payments to grantees and contractors. At the end, financial and performance reports drive future decisions by the White House and Congress.

After May 2017, the data set will reflect some but not all of the stages of this lifecycle.
DATA Act 2022: Changing Technology, Changing Culture

Information Flows Covered by the DATA Act Information Model Schema (DAIMS)

Simplified Federal Spending Information Lifecycle:

President / OMB

Congress

Budget Reporting
Budget Direction

Agencies

Financial Reporting
FAADS (Grants) Reporting

Treasury

Expenditures

Award, Obligation, Outlay

FPDS (Contracts) Reporting

Recipients

Subaward Reporting

GSA
Realizing the vision

The DATA Act could represent the most ambitious transformation of U.S. public-sector management ever attempted. Despite its challenges, the DATA Act begins with a strong foundation: a comprehensive data model, support from Congress and civil society, and a commitment from the new administration. [21]

Those implementing the DATA Act have shown determination and creativity so far. Building the DAIMS, the DATA Act Broker and a new version of USASpending.gov was an enormous challenge, but Treasury employed an agile development model that now serves as a leading practice across government.

Similarly, each solution to the challenges described above represents an opportunity to broadly improve federal management and inform future innovations in the federal government and beyond.

We believe these challenges can be addressed by 2022. The following recommendations are directed primarily to Treasury and OMB, and secondarily to Congress, the executive branch, and private stakeholders.

Our recommendations will follow the “cultural” and “technical” distinction, but again the two are ultimately inseparable. For example, it will be much easier to address data quality problems if agencies use the DAIMS open spending data for internal purposes. Similarly, our recommended technical improvements to the DAIMS will require a strong, permanent, and inclusive governance structure.

Cultural Solutions

1: Help agencies navigate and gain value from the unified data set.

Individual agencies will compile their spending information, align it with the DAIMS, and report the standardized data that comprise the promised DATA Act data set. The best way to ensure the data set continues to improve in quality and scope is to make the agencies themselves its primary users.

Many federal agencies saw USASpending.gov as an external transparency platform, not an internal management tool. Unsurprisingly, the site’s grant and contract award data were questionable at best. Its application programming interface (API) was inadequate for large-scale use, while the data sets were not suitably structured to answer common user questions. [22] But the DATA Act can avoid its earlier counterpart’s fate if federal agencies use the DAIMS and their DAIMS-aligned data for internal management.

At present, most federal managers must navigate a series of siloed systems – for financials, budgeting, resource management, awards and more – deployed within divisions, bureaus and program offices. To assemble spending information for their own decisions, managers rely on data calls and static, purpose-built reports. But the DATA Act directs each agency to create a spending data set reaching across all its divisions and programs. Even with the DAIMS’ current, limited scope, this will be valuable.

With a bit of effort, agencies can employ DATA Act data sets for powerful analytics, program evaluation, budgeting and management decisions. Early pilot projects at the SBA demonstrated that DAIMS-aligned data can support views and analyses that previously required time-consuming translation. [23] SBA worked with Treasury and the GSA’s 18F technology development team to test the DATA Act Broker. The resulting data set provided award-specific details and the ability to analyze grant and contract characteristics across all SBA programs for the first time. [24]

Beyond the inherent value of the initial DATA Act data set itself, agencies can benefit from reusing parts of the system to standardize the outputs of their award and performance systems. The data set will become a resource for their inspectors general as well. [25]

Finally, agencies choosing to use DAIMS-aligned data for management may be more likely to notice and fix errors; they may also be more likely to accept future increases in the DAIMS’ scope and role.

How can we encourage agencies to use DAIMS-aligned data in their own operations? Our interviews highlighted the need for education as implementation rolls forward. Treasury is considering how to educate users as they explore the new version of USASpending.gov; contractors are thinking about how to generate value from the new data sets. And congressional leaders are pondering how to keep attention on the DATA Act and encourage agencies to stay involved in its evolution. [26]
Some peer evangelism will play a key role. SBA has already seen success in sharing its experience, helping other agencies understand the benefits of a unified, trustworthy data set. And outside groups’ educational programming can help provide agencies with the necessary skills and motivation.

One key point: Congressional use of the DAIMS data should give agencies an incentive to use the same data for managerial tasks. If congressional budgeters and appropriators begin relying on dashboards powered by the DAIMS to make spending decisions, agency managers will naturally gravitate to the same data for budget submissions – and, eventually, for other management activities.

Many organizations heavily involved in passing the DATA Act have not yet participated in its implementation, possibly because they lack the technical expertise to comment effectively on the specifics of the resulting data structures.

For example, the Treasury team in charge of developing the new USASpending website runs an open process at openbeta.USAspending.gov, an experimental platform designed to engage a wide range of stakeholders on new ways to visualize spending data. The Treasury team reports significant feedback from agencies, but much less from external stakeholders. According to a key public interest representative who worked diligently in the congressional fight for passage of the DATA Act, philanthropic attention shifted away from spending data during the 2016 election cycle and subsequent political transition. [27]

To achieve the full promise of the DATA Act by 2022, the public-interest community must become reengaged.

Education will be needed to further this engagement. Treasury can begin by creating spending-data visualizations that match the old financial statements and budget documents public interest advocates already know well. For example, an electronic mockup of a federal financial statement with added interactivity – expanding line items and drill-downs, for instance – might help advocates better grasp the power of open data and the benefits of the DATA Act. Treasury could build, or encourage other agencies to build, sector-specific spending data visualizations that appeal to specific public-interest communities – foreign aid, energy exploration leases, education programs, etc.

Cost savings through analytics

The potential for cost savings based on the effective use of data analytics is clear. The United States Postal Service Inspector General (USPS IG) has already used data modeling to identify $924 million in savings and improve operations in numerous business areas of their business. For example, the USPS IG created a data model to help the Postal Service achieve cost savings and efficiencies related to the more than 23,000 leased buildings across the country. In one geographic area alone, this model was able to identify $2.2 million in savings associated with 77 expiring leases. Another data model predicts the possibility that a given contract may involve fraud. According to former Inspector General Dave Williams, in an interview with DATA Act 2022 authors, “Investigators at the U.S. Postal Service Inspector General opened 270 investigations in 2015 and recovered $20 million directly attributable to data analytics. The auditors using their models booked about $100 million in savings opportunities.”

The USPS IG, an independent office, does not own the data it leverages, but must source it from the Postal Service. The USPS IG had to procure and curate spending data sets, at significant time and financial investment, in order to start its analytics practice. But the USPS IG’s analysis has now been demonstrated to be so valuable to the Postal Service that Postal Service officials are now working actively to deliver data sets and standardize them to make them more useful for analytics.

By standardizing disparate sources of spending information and by matching separately-generated financial and award information together, DATA Act compliance opens the possibility of similar projects for every inspector general.

For more information:
3: Establish a permanent governance structure for the DAIMS.

Treasury and OMB should not be expected to manage the DAIMS and its data elements indefinitely; they are responsible for financial management, not policy-driven transformation. And if the DAIMS is expanded to state and local spending, performance, awardee reporting, and other areas, as we recommend, the project will reach far beyond these agencies’ jurisdictions.

Treasury and OMB therefore should establish a permanent, sustainable governance structure for data standards.

It is clear that the community of spending-data consumers extends far beyond the executive branch. Logically, then, the ongoing management of the data structure will require its users’ participation and support.

It is also clear that a governance structure for DATA Act standards should be financially sustainable, able to respond quickly to advances in technology and to remain independent from the whims of politics.

There are at least two possible models for DATA Act governance. The Open Contracting Data Standard is a common data model for contracting that touches stakeholders around the globe. [28] Its governance structure consists of an independent secretariat, the Open Contracting Partnership (OCP), and a technical working group representing stakeholders from various governments, the private sector and civil society. [29] The Global Legal Entity Identifier Foundation (GLEIF) takes a similar, supra-national approach to the development and implementation of the non-proprietary Legal Entity Identifier (LEI), a unique code identifying entities reporting to participating agencies. [30] The LEI’s broad utility has attracted a critical mass of users to make it a practical success.

Both OCP and GLEIF spun out of existing public-sector efforts. The OCP initially was conceived and hosted by a public-sector entity. In 2015, it became an independent program housed within its fiscal sponsor, the Fund for the City of New York. [31] The GLEIF was created by the Financial Stability Board after a call from the G20 for leadership “in developing recommendations for a global LEI and a supporting governance structure.” [32] Both OCP and GLEIF show that multinational governance structures can be spun out of existing governmental or quasi-governmental programs.

The federal government should consider the long-term establishment of an independent, nonprofit entity to maintain data structures for public-sector spending, perhaps modeled on the GLEIF or the Open Contracting Data Standard. By 2022, a governance structure should be in place to ensure the DAIMS and its data elements are maintained and expanded.

Technical Solutions

4: Update the DAIMS to extract award data directly from source systems.

The decision to rely on legacy compilations for the DAIMS’ award information made it easier for agencies to meet the May 2017 deadline. Moving forward, however, it is a weakness.

As GAO has pointed out, this made the resulting data set subject to these systems’ documented accuracy problems. Moreover, so long as the DAIMS relies on FPDS-NG and FAADS reporting, Treasury and OMB will not be able to add greater detail to, or improve the definitions within, award data.

For both reasons, Treasury and OMB’s first major revision to the DAIMS should cut pre-existing compilations out of the process and require agencies to extract this information directly from their source award systems. Going straight to the source for both financial and award data should lead to more efficient processing, boost data quality, and save agencies time and money. It will also allow the legacy compilations to be eliminated.
Before 2022, then, Treasury, OMB and Congress should take the steps needed to eliminate other spending-related reporting requirements and replace them entirely with the DAIMS.

After agencies begin extracting contract information from their source systems, FPDS-NG and FAADS reporting requirements should be eliminated.

5: Improve the DAIMS’ data definitions.

If the DAIMS is to govern all spending reporting, each data element involved must fit the needs of its end users. Consider again the previously-mentioned “Place of Performance” data element. In the real world, the place of performance might be most naturally identified by a mailing address, a political or administrative region or a geographic area. A good data definition will accommodate such real-world variation while ensuring that each potential expression contains as much specificity as possible. For example, a Place of Performance not based on a mailing address should be denoted with accepted electronic geolocation standards. Legacy systems simply cannot allow this flexibility; the FPDS-NG’s Place of Performance field, for instance, requires a standard U.S. mailing address.

The development of the DAIMS allows the government to deploy richer data definitions that better reflect the real world. If the DAIMS relies on legacy systems, it will likely, instead, continue to employ their limited and often unrealistic definitions.

6: Work with all stakeholders to select a recipient identifier that maximizes transparency for data users, efficiency for agencies, and convenience for recipients, and then stick to it.

As it ponders the future of recipient identification, the government should involve all stakeholders, including vendors of identification services, the users of procurement and assistance data, and the recipients of grants, contracts, and other awards. These stakeholders have different needs. Vendors of identification services need to be able to recover the costs of managing recipient registration; federal and non-federal users of contract and grant data want quick and easy access; and grantees, contractors, and other recipients want to be able to be identified conveniently and cheaply.

Decisions about recipient identification are further complicated by the multiple legal regimes and agencies involved. For example, the GSA manages the government-wide contractor registration system; the GSA, the Defense Department, and NASA manage the Federal Acquisition Regulation; and Treasury and OMB are responsible, under the DATA Act, for government-wide data standards for spending.

Once a government-wide consensus about the future of entity identification has been reached, Treasury and OMB should ensure that the DAIMS conforms to it, to provide certainty and ensure interoperability.

7. Expand the DAIMS to reflect the full lifecycle of spending information.

As the DATA Act’s unified data set becomes the preeminent source of spending information, Treasury and OMB should expand the DAIMS to reflect all federal spending data throughout the lifecycle, and eventually to accommodate state and local spending information and performance reports. Congressional and White House leadership will be needed in some cases. The DAIMS ultimately should include the following stages of the lifecycle:

Post-award recipient reporting. Grantees and contractors are subject to a vast range of reporting requirements, some agency-specific and others government-wide. Most reports are legally public information but are not available in any central location. Incorporating these reports into the DAIMS would lower the compliance burden on grantees and contractors while adding a wealth of information to the unified spending dataset.

Significant delays in the contracting portion of the Section 5 Pilot could slow this effort. Regardless of the results, however, the executive branch and Congress should consider efforts to standardize post-award recipient reporting. These standards should be developed along with the DAIMS and eventually incorporated directly into the schema.

The President’s proposed budget. The proposed presidential budget, usually issued in late winter for the upcoming fiscal year, is a lengthy document. President Barack Obama’s fiscal 2017 budget request, issued in February 2016, was the first published as machine-readable data, in partnership
with Socrata. [33] If the DAIMS were expanded to encompass the president’s budget request, and if the new administration continues to issue budget requests as machine-readable data, the public and Congress could easily compare requests with previous-year spending. Executive-branch priorities – and areas of agreement and disagreement with Congress – would become clearer.

The White House could further this goal by continuing to release the president’s budget request as open, machine-readable data, and particularly by producing a DAIMS-compatible version. Treasury and OMB could provide technical assistance for this effort or even construct an appropriate extension of the DAIMS themselves.

Non-appropriated receipts. The DAIMS focuses on expenditures rather than on receipts, and does not distinguish federal funds by source. [34] The House Committee on Oversight and Government Reform recently has taken interest in receipts from sources other than taxes, some of which are not subject to the appropriations process. [35] An expansion of the DAIMS to reflect federal funds by source, as well as expenditure category, could deliver a more complete picture of federal spending.

Thus, we recommend that the DAIMS be expanded to include data about government receipts. Congress should consider amending the DATA Act to confer the necessary authority on Treasury and impose the necessary reporting mandates on federal agencies.

Payments. The DAIMS categorizes expenditures by appropriation, account, budget item, program activity, and award, but does not include – nor does the DATA Act require – checkbook-level disclosure of each federal payment. Many state governments, by contrast, now offer checkbook-level data. [36] Expanding the DAIMS to identify and track payments as they flow out of the government’s coffers could greatly enhance its level of detail. [37]

We recommend that the DAIMS be expanded to include Treasury’s existing Payment Information Repository, the most comprehensive source of transaction-level spending data.

Appropriations bills and laws. Congress directs spending by passing appropriations laws that dictate annual spending levels for each executive-branch agency and division. By expressing appropriations bills and laws as machine-readable data, Congress could give agencies, recipients and the public an instant view of the impact of each bill before its passage.

We recommend that Congress publish appropriations bills and laws in a machine-readable format that conforms to the DAIMS.

Performance data. Several interviewees cited the combination of performance and financial data as an important next step for the DAIMS – one that could help determine the return on investment for specific programs, better target spending, and more.

The Government Performance and Results Act Modernization Act of 2010 (GPRAMA) provided a mandate for agencies to report performance electronically through a machine-readable list of all programs, but the executive branch has never complied with the act by compiling an authoritative list of programs. [38] The lack of a standard definition of “program,” the siloed nature of performance.gov and the continued struggle to create a federal program inventory stand in the way of meaningfully connecting performance and financial data. [39]

By accommodating program-level performance data, the DAIMS could realize the intent of GPRAMA and connect spending and performance throughout the federal government. Lessons learned from the DATA Act and GPRAMA should be combined to help achieve full connection between government financial and performance data.
Simplified Federal Spending Information Lifecycle:

- **President / OMB**
  - Proposed Budget
  - Performance Reporting

- **Congress**
  - Appropriations Laws

- **Agencies**
  - Payment Requests
  - Post-Award Recipient Reporting (Grant, Contract, Loan, & Foreign Aid Reports)

- **Recipients**
  - State and Local Financial Data

- **Treasury**
  - Receipts

- **GSA**

Information Flows NOT Covered by the DATA Act Information Model Schema (DAIMS)
State and local financial data. While state and local entities are not directly connected to the federal spending cycle, a large portion of their revenue comes from the federal government. Many states are already working to improve the quality of their own financial data; bringing their information into harmony with DATA Act standards would allow for a complete picture of all public-sector spending information nationwide.

Lastly, then, but certainly not least, we recommend that policymakers consider how to conform state and local spending cycles to DATA Act standards and connect their spending transparency structures to the DAIMS.

**Conclusion**

The May 2017 release of the first-ever unified federal spending data set will be a hard-earned first step. But the DATA Act promises far more. If the DATA Act community – legislators, Treasury and OMB leaders, agency financial managers, contractors and civil-society advocates – continue their work, the vision of freely-flowing and interoperable spending data could be achievable by 2022.
Appendix I: Surveying the Status of DATA Act Implementation

On May 9, 2017, federal agencies are required to begin reporting spending data to Treasury in accordance with the DAIMS. The pilot programs required under Section 5 of the FFATA, as amended by the DATA Act, are required to finish their data collection at the same time. [40]

This appendix briefly describes the inspiration for these mandates and some of the related technical details.

Agency Reporting

The DATA Act represents the first attempt to unify three broad categories of reporting requirements for federal spending: cash-based agency budgets; accrual-based accounting data required by the Chief Financial Officers Act of 1990 (the CFO Act); and award data required by FFATA and the American Recovery and Reinvestment Act of 2009 (ARRA). [41] To this end, the law requires the federal government to first apply standard data elements and a government-wide data format (or schema) to all federal spending and second to publish these standardized data as a unified open data set. OMB and Treasury already have achieved consensus on standard data elements and built a schema; in May 2017, it will become clear if their efforts have in fact produced a unified open data set.

The DAIMS, built upon 57 data definition standards finalized by OMB and Treasury, is intended to give “an overall view of the hundreds of distinct data elements used to tell the story of how federal dollars are spent.” It provides technical guidance to federal agencies as they decide what data to report to Treasury. [42]

OMB directs agencies to report seven distinct files of data: Appropriations Account, Object Class and Program Activity, Award Financial, Procurement, Financial Assistance, Additional Awardee Attributes and Sub-Awardee Attributes. The first three files are pulled directly from agency systems and deal specifically with financial information, while the final four represent awards data and rely on existing, government-wide systems – the FPDS-NG and the FAADS, as well as the Award Submissions Portal (ASP) and the Federal Funding Accountability and Transparency Act Subrecipient Reporting System (FSRS).

Agencies will report their spending data sets to the DATA Act Broker, open-source software designed to receive them; validate their syntax; store them; and eventually provide them to USASpending.gov (or its successor website) for public use. As part of the process, senior officers at federal agencies (usually the chief financial officer) must certify that their datasets meet certain requirements. The DATA Act Broker can flag discrepancies, helping agencies identify and fix potential data-quality issues. This should eventually help agencies gain more control over their data and gain the confidence to use the DAIMS for internal purposes.

To create the DAIMS, Treasury embraced an “agile” development approach, based on design cycles that incorporated user feedback (gathered online via Github) in successive iterations. This process resulted in a delay in the release of the finalized schema (version 1.0), which GAO has said may make it difficult for some agencies to begin reporting through the DAIMS by the deadline. [43] On the other hand, interviews with Treasury staff as well as contractors working on DATA Act compliance indicate that this iterative, stakeholder-centric approach helped ensure that the DAIMS is both technically sound and functional for agencies. [44]

Representatives from the Center for Organizational Excellence praised Treasury for improving stakeholder engagement and change management by “leaps and bounds” with its DATA Act program management office. [45] The Treasury team working on DATA Act implementation has been honored for its approach to updating the USASpending.gov website. [46]

Our interviews confirm that DAIMS 1.0 is designed with agencies in mind and that, despite some delays, the data agencies report will be better for it.

The DAIMS structure and policy approach take into account agencies’ concerns about resource allocation, technical capacity and time constraints. [47] For example, the DATA Act Broker was designed to help agencies validate their data and begin addressing quality issues with a minimum of additional effort. [48] Interviewees generally agreed they believe most agencies would be ready to begin reporting data in May. [49] But there is wide agreement that some data quality problems are to be expected in the initial release.

These expected quality problems have several causes. As noted above, the source systems for award data (FPDS-NG, FAADS, ASP and FSRS) have known data quality problems. Despite their shortcomings, OMB and Treasury decided to continue relying on these systems for the time being, to “[promote]
timeliness in reporting and [provide] an initial solution to challenges imposed by the absence of new funding streams.” [50]

Secondly, as GAO has noted, some DAIMS data definitions leave enough room for interpretation that agencies are likely to report data in varying ways. [51] Agreeing on data standards and building DAIMS was a complex social exercise with numerous stakeholders. Certain decisions made to ease agency compliance in the short term may not be suitable to the long-term vision of the DATA Act. Future changes to the DAIMS and its accompanying standards should ensure better, more useful data for agencies and the American public.

A third theme in our interviews was the need to settle on a permanent governance structure – particularly for decisions concerning the maintenance and improvement of the DAIMS – as the system continues to evolve.

The initial technical and policy decisions made by Treasury and OMB have led to the creation of a long-envisioned unified spending data set, but more work is required to ensure that it is as accurate, useful and ultimately as transformative as possible.

### Recipient Reporting

The recipient reporting pilot program mandated by the DATA Act stems from lessons learned from ARRA, which took an ambitious, data-centric approach to overseeing the $800 billion of stimulus funding it directed into the nation’s economy. [52] ARRA created a Recovery Accountability and Transparency Board that required stimulus recipients to report via a government-wide data standard. [53] The board conducted a small pilot program to determine whether this approach, if adopted for all grants, could reduce grantees’ total compliance costs. [54] The results suggested that standardized recipient reporting could, in fact, reduce compliance costs through automation. [55]

Section 5 of the DATA Act requires OMB to conduct another pilot program to test standardized recipient reporting. [56] OMB divided the Section 5 pilot into two programs, one for grants and other forms of assistance and the other for contracts. HHS managed the grant reporting pilot. Less information has been made available about the contracting pilot, but OMB has testified that several agencies are collaborating on it. [57] The grants portion of the pilot appears to be on schedule. The contracts portion, however, is nearly a year behind schedule and may derail the entire timeline for Section 5.

### Grantee Reporting Pilot

HHS, the most significant federal grant-making agency by dollar amount, was already considering how to standardize its spending data before passage of the DATA Act, and had already been working with OMB in this area, so its participation in the pilot project was a natural fit. [58]

As a tie-in to the pilot, HHS launched its Common Data Element Reporting (CDER) Library, a “federal-wide, online searchable repository for grants-specific data standards, definitions and context.” It includes the 57 DATA Act data standards as well as a growing set of thousands of data elements unique to grant reports. [59] The CDER Library is a powerful tool that should be adopted for grantee reporting throughout the federal government.

Again, the grantee reporting portion of the pilot appears to be going smoothly and should be ready to issue its portion of the required report to Congress by the August 2018 deadline specified in the law.

### Contractor Reporting Pilot

The contractor portion of the pilot, by contrast, has encountered significant problems.

GAO noted its concerns about the contractor reporting pilot at the end of 2016, reporting that data collection under the pilot would not begin until January or February 2017, at least seven months late and only a few months before the pilot was supposed to conclude. The pilot, furthermore, applies only to certified payroll reporting and “does not clearly document how findings related to the centralized … portal will be applicable to other types of required procurement reporting.” [60]

A congressional staffer praised OMB for trying to design an ambitious pilot but expressed serious concerns about the delays. The staffer noted every piece of the pilot program must be complete before this section of the DATA Act can move forward. The delay in the contractor portion will delay congressional oversight and, ultimately, any decision OMB makes about standardized recipient reporting. [61]

The Section 5 Pilot has been hampered by delays and internal politics, but this should not prevent the government from embracing the goal of standardized recipient reporting. The CDER Library shows it is possible to map the range of data points recipients may need to consider, while the success of ARRA transparency requirements remind us of the power of standardization. Standardization can lower the burden on recipients by allowing them to automate their reporting. It will also produce better data for federal agencies, Congress and the public.
Appendix II: Interviews

- Gary Bass, executive director, The Bauman Foundation
- Andrew Brown, data scientist, Govini
- Autumn Carter, Government Affairs, OpenGov, Inc.
- Congressional Oversight Staff
- Kaitlin Devine, Department of the Treasury
- Dr. Paul Eder, lead consultant, Center for Organizational Excellence
- Amy Edwards, Department of the Treasury
- Bill Eggers, executive director, Deloitte Center for Government Insights
- Shane Engel, senior manager, Deloitte Risk and Financial Advisory, Deloitte & Touche LLP
- Eric Gillespie, CEO, Govini
- Steve Goodrich, president and CEO, Center for Organizational Excellence
- Rob Gramss, principal, Deloitte Risk and Financial Advisory, Deloitte & Touche LLP
- Dick Gregg, managing director, H.J. Steininger PLLC
- Timothy Gribben, Chief Financial Office, Small Business Administration
- Christina Ho, Department of the Treasury
- Debbie Kramer, Department of Health and Human Services
- Karen Lee, Office of Management and Budget
- Derick Masengale, managing director, Deloitte Consulting
- Marco Mendoza, visual designer, Govini
- Seth Metcalf, deputy treasurer, Ohio
- Timothy Miller, public safety specialist, Socrata
- Michael Peckham, Department of Health and Human Services
- Christian Peratsakis, consulting services manager, Socrata
- Tim Richardson, CTO, Govini
- David Schmidtknecht, principal, CBYEONData
- Chuck Simpson, principal consultant, Center for Organizational Excellence
- Scott Straub, director of Federal Markets, Lexis Nexis Risk Solutions
- Kelly Tshibaka, chief data officer, USPS Office of the Inspector General
- Peter Viechnicki, strategic analytics manager and data scientist, Deloitte Services LLP
- Christopher Zeleznick, Department of Health and Human Services
Tasha Austin (Deloitte & Touche LLP) is a Senior Manager in Deloitte’s Federal practice and has more than 16 years of professional services experiences involving commercial and federal financial statement audits, fraud, dispute analysis and investigations, and data analytics. Tasha leads the DATA Act offering for the Risk and Financial Advisory practice, where she introduces innovative insight-driven solutions to her clients to help them transform financial management. Most of Tasha’s time is spent providing strategic direction to executives and management across the federal community to help them solve their agencies’ most complex and unique data challenges. Tasha delivers innovative data solutions to drive business process reengineering, internal controls assessments, financial statement account reconciliations, and implementation of corrective actions to help agencies create a repeatable, sustainable, and auditable environment. Tasha has moderated panel discussions with executives across the federal community and facilitated agency-level training to demonstrate the value in leveraging data as a strategic asset to deliver key insights.

Tasha also develops promising STEM Talent across Greater Washington by serving as an Adjunct Professor at the Northern Virginia Community College, where she teaches a combination of advance mathematics and statistics courses. Tasha has a BS and MS in Mathematics from North Carolina Central University in Durham, NC and an Executive MBA from Howard University in Washington, DC.

Christina Canavan (Deloitte & Touche LLP) is a Managing Director in Deloitte’s Federal practice and has more than 16 years of professional service experience providing innovative and advanced data analytic solutions for both Federal agencies and Fortune 500 companies. Christina specializes in leveraging data to provide insights and recommendations regarding risk, performance, and goal achievement. Christina helps organization’s leverage their data as a strategic asset by delivering insights and data-driven solutions to help clients solve long standing challenges. Currently, Christina leads the Advisory Analytics practice for Financial Risk Transactions & Restructuring, which spans multiple departments and provides professionals with a focus on bringing analytic insights to the financial management community. Christina has written articles focused on building analytics professionals within the federal community and moderated panel discussions with executives.

Dean Ritz is a subject matter expert in information modeling with over three decades of experience in various data-dominated domains including artificial intelligence, expert systems, object-oriented programming, and most recently the modeling of financial information. As a Senior Director at Workiva, he applies his expertise to product strategy for collaborative work management and the management of the company’s expanding patent portfolio. His interests extend to the topics of rhetoric and ethics, with scholarly work in these areas published by Oxford University Press (2011, 2009, 2007), Waveland (2011), and Routledge (forthcoming 2017).

Matt Rumsey is a consultant, researcher, and advocate with expertise in open data and transparency policy. He has done work for organizations including the Center for Open Data Enterprise, the Sunlight Foundation, and the Data Foundation. Previously, he managed the Sunlight Foundation’s federal level policy initiatives as senior policy analyst. He is a graduate of the American University in Washington, D.C. and currently resides in Paris, France.
References


[3] DATA Act at new Sec. 4(c)(2).


[8] The DATA Act includes reporting requirements for agency inspectors general and the U.S. Comptroller General. The last of these reports is not due until November 2021. See DATA Act at new Sec. 6.


[13] U.S. Government Accountability Office, OMB and Treasury Have Issued Additional Guidance and Have Improved Pilot Design but Implementation Challenges Remain, Figure 1.


[23] Interview with Peter Viechnicki.

[24] Interview with Timothy Gribben, chief financial officer, U.S. Small Business Administration; and Case Study: The SBA.


[27] Interview with Gary D. Bass, founder and former director of OMB Watch.


[29] Interview with Timothy Miller and Christian Peratsakis, Socrata.


[37] Interview with Timothy Gribben.


[39] Interview with Peter Viechnicki, strategic analytics manager and data scientist, Deloitte.


[44] Interviews with Steve Goodrich, Chuck Simpson, and Paul Eder, Center for Organizational Excellence, and with Timothy Gribben, SBA.

[45] Interviews with Steve Goodrich, Chuck Simpson, and Paul Eder, Center for Organizational Excellence.

[46] See for instance “Third Annual Le Hackies: Top Ten Legal Hacks of the Year,” docs.google.com/presentation/d/1LiRslih9vLmaoxHCWcf91B-YeVgZQnd3hPw2AlPXEhj4/edit#slide=id.g1b8118f566_0_183.

[48] Interview with Treasury staff.


[55] Recovery Accountability and Transparency Board, Grants Reporting Information Project, p. 3 (“The ability to produce standard machine readable report files (e.g. XML2) and to ‘bulk’ or ‘batch’ multiple grant reports in one XML file, regardless of agency or program, can significantly increase reporting efficiency”).

[56] DATA Act at new section 5 (describing scope of pilot program to test standardized reporting; DATA Act purpose no. 3 (“amplify reporting for entities receiving Federal funds by streamlining reporting requirements and reducing compliance costs while improving transparency”).

[57] For more information about the genesis of the pilot programs, see Data Foundation and MorganFranklin Consulting, Vision & Value, pp. 9-10.


[61] Interview with Congressional Oversight Staff.
The Data Foundation is the nation’s first industry-focused open data research organization. Through research, education, and programming, the Foundation illuminates the benefits of transforming government information into standardized, open data.

For more information, visit datafoundation.org.

Deloitte provides industry-leading audit, consulting, tax and advisory services to many of the world’s most admired brands, including 80 percent of the Fortune 500 and more than 6,000 private and middle market companies. Our people work across more than 20 industry sectors to deliver measurable and lasting results that help reinforce public trust in our capital markets, inspire clients to make their most challenging business decisions with confidence, and help lead the way toward a stronger economy and a healthy society.

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. In the United States, Deloitte refers to one or more of the US member firms of DTTL, their related entities that operate using the “Deloitte” name in the United States and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting.

Please see www.deloitte.com/about to learn more about our global network of member firms.

Copyright © 2017 Data Foundation. All rights reserved.
Copyright © 2017 Deloitte Development LLC. All rights reserved.

This document contains general information only and Deloitte and The Data Foundation are not, by means of this document, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This document is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor.

Deloitte and The Data Foundation shall not be responsible for any loss sustained by any person who relies on this document.