DATA
COALITION

Committee on Oversight and Government Reform
Subcommittee on Government Operations
Subcommittee on Health Care, Benefits and Administrative Rules

United States House of Representatives
December 1, 2016

Testimony
Hudson Hollister
Executive Director, Data Coalition

[Slide XX - Introductory Slide]

Chairman Meadows, Ranking Member Connolly, Chairman Jordan, Ranking Member
Cartwright: thank you for inviting me to testify.

In April of 1802, President Thomas Jefferson wrote to Albert Gallatin,[1] the fourth Secretary of
the Treasury, supporting Gallatin's plan to "simplify our system of finance, and bring it within the
comprehension of every member of Congress." Jefferson believed that federal spending
information had become so complex and so fragmented that only the experts could understand
it.

Jefferson's solution to this problem was a "simplification in the form of accounts in the treasury
department, and in the organization of it's [sic] officers, so as to bring every thing to a single
center." By expressing federal spending as "one consolidated mass," he wrote, "we might hope
to see the finances of the Union as clear and intelligible as a merchant's books, so that every
member of Congress, and every man of any mind in the Union, should be able to comprehend
them, and consequently, to control them."

Two hundred and fourteen years later, we face the same problem and we need the same
solution. The federal government is the largest and the most complex organization in human
history. But by expressing all federal spending as one consolidated data set, we can use
commercially-available software to make it "clear and intelligible," so that Congress and the
people can comprehend it, and control it.

Federal spending information is complex and fragmented. Agencies must report their receipts
and account balances to Treasury, their budget information to the White House Office of

---

[1] http://founders.archives.gov/documents/Jefferson/01-37-02-0132#TSJN-01-37-02-0132-kw-0001

Management and Budget, and their contracting details to the General Services Administration. This information is handled by thousands of incompatible software systems.

But two years ago this Committee unanimously approved, and President Obama signed, the reform necessary to express federal spending information as a single data set. The Digital Accountability and Transparency Act, or DATA Act, directs Treasury and OMB to create a single, government-wide data structure for all federal spending information.

[Slide XX - DATA Act Timeline]

The most important deadline is five months away. In May of 2017 every federal agency must begin to report spending information using the same data format, creating a single electronic picture of all spending. In the next few minutes I am going to show you what that picture is going to look like.

I know the Committee is particularly interested in fees, fines, penalties, and settlements that agencies receive outside the appropriations process. So far, the data structure that Treasury and OMB are using under the DATA Act focuses on expenditures, not on receipts, so our single electronic picture of all spending won't be able to provide full detail on how fees, fines, penalties, and settlements are spent.

But Congress can direct Treasury and OMB to expand it. And Congress should.

When I served as counsel to this Committee, I worked on the first version of what became the DATA Act. I resigned from the Committee staff in 2012 in order to start the Data Coalition and help then-Chairman Issa and Ranking Member Cummings get the DATA Act passed. The Data Coalition is a trade association of nearly 40 companies whose commercially-available software can inform decisions, illuminate fraud, and automate compliance - but only if we have a single data set to work with.

Earlier this year we established the Data Foundation, a nonprofit organization whose mission is to illuminate the benefits of open government data through research and education. For more on the DATA Act's history, goals, and potential, you can read the Data Foundation's first research paper, *The DATA Act: Vision & Value*,[2] co-published with MorganFranklin last July.

[Slide XX - Life Cycle of Federal Spending Information]

Federal spending information is not just fragmented across agencies, systems, and reporting requirements. It is also separated into stages. Here is the full life cycle of federal spending information.

---

[2] http://www.datafoundation.org/data-act-vision-and-value-report/

Federal funds flow into receipt accounts. Some receipts are available for spending immediately. Other receipts are unavailable to be spent until Congress appropriates the funds.

In 2006, the Federal Funding Accountability and Transparency Act, or FFATA, required the executive branch to begin publishing a summary of each federal grant and contract on the USASpending.gov website. The light gray arrow shows the information that is published under FFATA.

The DATA Act - shown here with the dark gray arrow - requires the government to publish more of the life cycle and match appropriations to the grant and contract awards that are paid out of them. One of the most important expansions in transparency is not shown in this diagram: the DATA Act requires the publication of the direct spending that doesn't go out in the form of grants or contracts, like salaries and benefits.

By requiring Treasury and OMB to set up a single government-wide data structure for appropriations, grants, and contracts, the DATA Act gives us a single electronic picture of all that information - or, to use President Jefferson's phrase, "one consolidated mass" that allows Congress and the public to "comprehend … and consequently, to control" spending.

The rest of my testimony has three parts. First, I will explain the information that is available to Congress and the public before the DATA Act comes into effect this May. Second, I will preview the single electronic picture of spending that will become available after agencies begin reporting appropriations, grants, and contracts using the standardized data structure. Third, I will show how even after the DATA Act, there will still be a need to expand the data structure to include more information for a more complete picture, and suggest how Congress might do that.

[Slide XX - Federal Spending Information Before the DATA Act]

Before the DATA Act, federal spending is mostly available as static documents, not as searchable data.

[Slide XX - Monthly Treasury Statement Example]

The Monthly Treasury Statement of Receipts and Outlays, a document published every month by the Bureau of the Fiscal Service, is the most complete breakdown of federal funds received and spent by the government.

The Monthly Treasury Statement summarizes all receipts, both those that are available to be spent immediately and those that cannot be spent until a Congressional appropriation. But because it is a static document, there is no way to access more detail or follow a particular category of receipts to see what happened to it.

[Slide XX - USASpending.gov]

Under the Federal Funding Accountability and Transparency Act of 2006, federal agencies must report a summary of every grant and a summary of every contract, and that information is published on the USASpending.gov portal. It is fully searchable and interactive.

However, because USASpending.gov is solely focused on grants and contracts, it does not give us the full picture of federal spending. It does not allow us to navigate back and forth between larger appropriations categories and specific grant and contract awards.

[Slide XX - Federal Spending Information After the DATA Act]

Under the DATA Act, Treasury and OMB have created a government wide data structure that connects expenditure accounts to grants and contracts for the first time. Once all the information is expressed using this data structure, starting in May 2017, we will have a single, authoritative data set that shows all expenditures, broken down by account, by grant, and by contract.

Commercially available software will use this single data set to portray an electronic picture of federal spending. Several of our Data Coalition members are working on software that will provide agencies, and Congress, and the public with new ways to comprehend, and control spending.

[Slides XX-XX - Art of the Possible]

Here we see a navigation from an overall appropriation category all the way to a particular contract. This level of interactivity will be possible across the whole executive branch.

This set of visualizations was created by Booz Allen Hamilton, but we expect other software companies will compete, once federal spending is available as a single data set.

The reliability of this electronic picture of federal spending will depend on how well agencies comply with the DATA Act, starting in May 2017. They must report complete, timely, accurate, high-quality data that matches the data structure Treasury and OMB have created.

The DATA Act requires every inspector general to evaluate the data its agency reports. Last month, the inspectors general of most agencies published readiness reviews, and most came out positive, but some agencies are not ready. The Data Coalition has published a summary of the readiness reviews.[3] I understand this committee intends to conduct further oversight on agencies' DATA Act work.

[Back to Slide XX - Federal Spending Information After the DATA Act]

---

[3] http://www.datacoalition.org/blog/.

Even after the DATA Act, some limitations will remain. First, the data structure that Treasury and OMB have created focuses on expenditures, not receipts. That means we won't easily be able to differentiate between appropriated and non-appropriated funding sources.

Second, the data structure does not go all the way down to the checkbook level, with details for each payment. That means we won't yet be able to navigate all the way from the whole federal government to see the date and amount of every payment the government makes. Many states, led by Ohio, are providing this level of transparency already. In fact, Ohio allows the public to view every payment, see which accounts and budget categories the money came from, and see the name and contact information of the state official who was responsible for that payment.

[Slide XX - Federal Spending Information if the DATA Act is Amended]

Congress should amend the DATA Act to direct Treasury and OMB to expand their data structure. Here is what the life cycle of federal spending information would look like if the DATA Act covered both receipts at the beginning and payments at the end, as shown by the green arrow.

I asked our Data Coalition members to come up with a prototype visualization that shows what this expanded electronic picture of federal spending might look like.

[Slides XX-XX - Receipts Dashboard]

Here we see a navigation that itemizes the federal government's receipts between those that are unavailable, and require Congressional appropriation to be spent, and those that are available for spending immediately. We can click on a particular receipt and see how it flowed into expenditure categories, and what happened to it.

If the data structure is expanded as we are recommending, we will also be able to navigate all the way down to the payment level, just as is possible today in Ohio and other states.

[Slide XX - Conclusion: Expand the Data Structure; Eliminate Duplicative Systems; End the DUNS Monopoly]

We have prepared recommendations for the Committee on how to amend the DATA Act to expand the data structure and ensure that receipts and payments are brought into the picture. Our legislative recommendations for the DATA Act are based on the Center for Open Data Enterprise's Transition Report,[4] which explains how the next Presidential administration can work with Congress to use open data to deliver transparency, accountability, and efficiency across all government operations.

---

[4] http://opendataenterprise.org/transition-report.html

In addition to expanding the data structure to provide a more complete electronic picture of federal spending, we believe Congress should clarify how the DATA Act should streamline the federal government's reporting processes. Beginning in May 2017, agencies will be reporting their spending information spending twice - first the old-fashioned way, as static documents and through disconnected databases, using legacy systems like the Federal Procurement Data System (FPDS), and a second time as searchable, standardized data under the DATA Act.

Congress should eliminate this duplication as soon as possible. The DATA Act process, based on data instead of on documents, must become the only way that agencies report spending. Our legislative recommendations provide a foundation to accomplish this.

Finally, Congress must address the biggest obstacle to the public's access to spending information. The federal government uses an identification code called the DUNS Number to identify every grantee and contractor across all of its systems. The DUNS Number is proprietary. It is owned by Dun & Bradstreet, Inc., which is itself a contractor. Nobody can download federal grant or contract data without purchasing a license from Dun & Bradstreet. Taxpayers paid for this information to be compiled, and yet they cannot download or analyze it without paying again, every time.

Congress should end Dun & Bradstreet's monopoly by directing the government to adopt a nonproprietary, freely reusable identification code for grantees and contractors. Our legislative recommendations would phase out the DUNS Number and replace it with an identification code that everyone can freely download, such as the Legal Entity Identifier (LEI).

President Jefferson's 1802 letter to Secretary Gallatin complained that the government's financial position had taken on "the most artificial and mysterious form," "until the whole system was involved in impenetrable fog" and beyond the comprehension of Congress and the public. To anyone but a Treasury expert, the distinction between appropriated and non-appropriated receipts is indeed artificial and mysterious, and it allows the government to operate beyond Congress' comprehension and beyond public accountability.

But by tracking all the complexity of federal spending using a single, government-wide data structure, and by publishing all that information as one data set, we can realize Jefferson's vision of "one consolidated mass," "clear and intelligible." This Committee began this work by passing the DATA Act. By holding the executive branch accountable to follow the law, and by expanding the law where necessary, this Committee can finish it. Thank you.