



December 2018

OPEN DATA

Treasury Could Better Align USAspending.gov with Key Practices and Search Requirements

GAO Highlights

Highlights of [GAO-19-72](#), a report to congressional addressees

Why GAO Did This Study

Open data can foster accountability and public trust by providing citizens with information on government activities and their outcomes. It can also promote private sector innovation. The DATA Act requires that the federal government collect and present open data on roughly \$4 trillion in annual federal spending. The DATA Act also includes a provision requiring GAO to review its implementation.

This report (1) identifies key practices for transparently reporting government data; and (2) evaluates the extent to which USAspending.gov is consistent with those key practices and other requirements. GAO developed the key practices by systematically evaluating and synthesizing information from literature on open data, as well as interviews with open data experts and good governance groups. GAO used these key practices as well as existing federal website standards and applicable laws to evaluate USAspending.gov.

What GAO Recommends

GAO is making five recommendations including that Treasury (1) establish a process to ensure that additions to USAspending.gov meet security requirements, (2) provide structured metadata and licensing information on the website, and (3) ensure that users can search for awards by city and program source as required by law. Treasury agreed with GAO's recommendations.

View [GAO-19-72](#). For more information, contact Triana McNeil (202) 512-6806 or McNeilT@gao.gov

December 2018

OPEN DATA

Treasury Could Better Align USAspending.gov with Key Practices and Search Requirements

What GAO Found

GAO identified five key practices for transparently reporting government data, as well as actions to implement each practice. These key practices and actions can assist managers of open government data programs in the transparent presentation of their data. Open data are information that can be freely used, modified, or shared by anyone for any purpose.

USAspending.gov aligns with several key practices. However, the Department of the Treasury (Treasury) has not fully aligned the website with all of the key practices, the requirements of the Federal Funding Accountability and Transparency Act of 2006 (FFATA), and Office of Management and Budget (OMB) guidance (see table.) FFATA, as amended by the Digital Accountability and Transparency Act of 2014 (DATA Act), directed Treasury to develop and manage USAspending.gov to provide detailed information on federal spending.

USAspending.gov Alignment with Key Practices and Other Requirements

Key practices	USAspending.gov characteristics
 <p>Provide free and unrestricted data</p>	<p>Consistent with this key practice, USAspending.gov, as a general matter, provides unrestricted data to the public and Treasury has made all the data available for free. In the course of its work, GAO found that users could only download the complete database after registering for an account with a third-party database host and incurring a charge. In July 2018, in response to GAO's inquiries, Treasury provided an option for users to download the entire database for free without creating an account.</p>
 <p>Engage with users</p>	<p>Consistent with this key practice, Treasury conducts user research and engages with users of USAspending.gov online and in person. For example, Treasury officials said they have conducted more than 150 interviews with users. Treasury officials also told GAO that they have directly engaged with various audiences about the website through a hackathon and a partnership.</p>
 <p>Provide data in useful formats</p>	<p>Consistent with this key practice, USAspending.gov provides detailed, machine-readable data in standardized formats and includes several options for downloading the data. However, Treasury lacks a process to ensure that any new pages added to the website are hosted on a .gov domain and encrypted with the secure HTTPS protocol required by OMB. Without such a process Treasury does not have assurance that new pages added to the website meet federal security requirements.</p>
 <p>Fully describe the data</p>	<p>USAspending.gov does not fully reflect this key practice because while it discloses data sources and timeliness, the website does not provide structured metadata. Providing metadata—structured descriptive information about the data—is also an OMB requirement. In addition, while OMB guidance provides that agencies must apply open licenses, the website does not provide comprehensive licensing information for the data. As a result, users may not have the information they need on how the data can be used.</p>
 <p>Facilitate data discovery for all users</p>	<p>Consistent with this key practice, Treasury facilitates data discovery by providing tools such as interactive maps and visualizations to help users interpret the data. However, users cannot search for city and program source (the unique accounting code used to track federal spending) as required by FFATA. As a result, users may be unable to find certain financial or geographic information related to recipients or federal programs.</p>

Source: GAO analysis of USAspending.gov. | GAO-19-72

Contents

Letter		1
	Background	4
	Key Practices for Transparently Reporting Government Data	8
	USAspending.gov Aligns with Several Key Practices, but Does Not Fully Meet Licensing, Search Functionality, and Other Requirements	26
	Conclusions	45
	Recommendations for Executive Action	46
	Agency Comments, Third-Party Views, and Our Evaluation	47
Appendix I:	Objectives, Scope, and Methodology	50
Appendix II:	Comments from the Department of the Treasury	54
Appendix III:	GAO Contact and Staff Acknowledgments	56
Related GAO Products		57
Table		
	Table 1: Key Practices and Key Actions for Transparently Reporting Open Government Data	9
Figures		
	Figure 1: Timeline of Federal Spending Transparency Laws and Versions of USAspending.gov	6
	Figure 2: Key Actions for Providing Free and Unrestricted Data	10
	Figure 3: Key Actions for Engaging with Users	12
	Figure 4: Key Actions for Providing Data in Useful Formats	15
	Figure 5: Example of Transaction-Level Spending Data from Ohio's Online Checkbook	16
	Figure 6: Key Actions for Fully Describing Data	18
	Figure 7: Metadata for Kansas City, Missouri's, Budget Data	19
	Figure 8: Key Actions for Facilitating Data Discovery for All Users	21
	Figure 9: "OhioCheckbook" Mobile App Showing Payments to Cuyahoga Community College	22

Figure 10: Summary Data from the Montgomery County, Maryland's, spendingMontgomery Website	23
Figure 11: Data Discovery Tools on New York City's Open Data Portal	25
Figure 12: Example of a Treasury User Profile	28
Figure 13: Treasury Contractor Interviewing a Member of the Public at the Capitol Visitor Center	29
Figure 14: Example of Detailed and Disaggregated Data from a Contract Summary Page on USAspending.gov	31
Figure 15: Download Options Available on USAspending.gov	32
Figure 16: No Award Information in Spending Explorer for the Vaccines for Children Program Activity	36
Figure 17: Example of a Map of Advanced Search Results for Social Security Retirement Insurance (as of September 2018)	41
Figure 18: Example of a Contract Explorer "Sunburst" Visualization from the Data Lab (as of September 2018)	42
Figure 19: Example of Plain Language (left) and Official Definition (right) on USAspending.gov	43

Abbreviations

API	Application Programming Interface
CFDA	Catalogue of Financial Domestic Assistance
CFO Act	Chief Financial Officers Act of 1990
DAIMS	DATA Act Information Model Schema
DATA Act	Digital Accountability and Transparency Act of 2014
DUNS	Data Universal Numbering System
FFATA	Federal Funding Accountability and Transparency Act of 2006
FOIA	Freedom of Information Act
GSA	General Services Administration
NAICS	North American Industry Classification System
OECD	Organisation for Economic Co-Operation and Development
OMB	Office of Management and Budget
PII	personally identifiable information
TAS	Treasury Account Symbol
Treasury	Department of the Treasury

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



December 13, 2018

Congressional Addressees:

The increasing availability of open government data—that is, government-produced information that can be freely used, modified, and shared by anyone for any purpose—holds great potential to promote government transparency and private sector innovation. In recent years, governments around the world have increased the availability of open data by providing information on a wide range of topics such as health, education, transportation networks, budgets, and performance. Such data can foster accountability and public trust by providing citizens with information on government activities and their outcomes. In addition, according to a 2014 study by the Department of Commerce, U.S. government data have helped private firms in data-intensive industries generate revenues of at least \$24 billion and upwards of \$221 billion annually.¹ However, for open data to achieve their full potential, government websites must ensure that their data are effectively presented to a wide range of users.

To provide more transparency for roughly \$4 trillion in federal spending, on March 2, 2018, the Department of the Treasury (Treasury) released a new version of its USAspending.gov website.² The new USAspending.gov displays data submitted by federal agencies to comply with the Federal Funding Accountability and Transparency Act of 2006 (FFATA), as amended by the Digital Accountability and Transparency Act of 2014 (DATA Act), and upon release became the official source of spending data required to be published under the DATA Act.³ Because of federal spending data's broad appeal to users such as Congress, grant recipients, journalists, and the general public, it is important that USAspending.gov follow key practices for reporting this information as transparently as possible.

¹See Department of Commerce, Economics and Statistics Administration, *Fostering Innovation, Creating Jobs, Driving Better Decisions: The Value of Government Data* (July 2014).

²Treasury first released data submitted under the Digital Accountability and Transparency Act of 2014 (DATA Act) in May 2017 on an earlier version of the website, Beta.USAspending.gov. USAspending.gov was initially launched in 2007.

³Pub. L. No. 109-282, 120 Stat. 1186 (Sept. 26, 2006); Pub. L. No. 113-101, 128 Stat. 1146 (May 9, 2014) *codified at* 31 U.S.C. § 6101 note.

The DATA Act includes a provision for us to review implementation of the act.⁴ Over the past 4 years, we have issued 13 reports assessing various aspects of DATA Act implementation.⁵ This report builds on our body of work on the DATA Act and (1) identifies key practices for transparently reporting government data on a centralized website, and (2) evaluates the extent to which the new USAspending.gov is consistent with these key practices, as well as existing standards for federal websites.

To identify key practices for transparently reporting open government data on a centralized website, we conducted a literature search and systematically reviewed articles on open government data programs and practices. We evaluated and synthesized information from the literature review to identify commonly-reported key practices for transparently reporting open government data. We also interviewed experts on open data and good governance.⁶ We first had open-ended conversations with experts to obtain their views on what key practices exist for transparently reporting open government data. After developing an initial list of key practices, we then obtained experts' feedback to finalize the list. We shared a draft of the key practices with Treasury, the Office of Management and Budget (OMB), and the General Services Administration (GSA).

To obtain illustrative examples showing how those key practices can be implemented, we selected open data practitioners from six state and local governments:

- Connecticut;
- Kansas City, Missouri;
- Los Angeles, California;
- Montgomery County, Maryland;

⁴FFATA, § 6(b).

⁵A list of related products, including our prior work on the DATA Act and USAspending.gov, is included at the end of this report.

⁶Specifically, we obtained the views of 18 individuals and organizations from a variety of sectors and backgrounds, including the Data Coalition (which advocates for the publication of government information as standardized, machine-readable data), Code for America (which organizes people to build technology that helps local governments), and the IBM Center for the Business of Government (which researches government management issues, including the use of technology). A full list of the individuals and organizations we interviewed is included in appendix I.

-
- New York City, New York; and
 - Ohio.

We selected these practitioners because they have well-regarded open data websites that include a general open data portal as well as budget or spending data visualizations, and because they represent different locations and levels of government, including cities, counties, and states. We reviewed these practitioners' open data websites and related documentation, and interviewed cognizant state and local government officials.

To assess the extent to which the new USAspending.gov is consistent with key practices and existing standards for federal websites, we compared the website to applicable laws and policies—such as FFATA and OMB guidance—as well as the key practices in this report. We also reviewed documentation related to the website, observed Treasury's participation in a hackathon, and interviewed Treasury officials.⁷ To clarify policies and practices for federal websites, we interviewed GSA officials and OMB staff.

To evaluate the extent to which specified federal award data elements on USAspending.gov are searchable, as required by FFATA, as amended by the DATA Act, we obtained and tested a nongeneralizable, random sample of award data from USAspending.gov.⁸ Specifically, we downloaded contract and financial assistance data from USAspending.gov, generated a random sample of awards, and used the website's search tools to search for the required data elements from 15 contracts and 15 financial assistance awards.

For the purposes of this report, we focused on the presentation of data and related contextual information on USAspending.gov. Our review did not include data quality or data governance.⁹ We reported on the quality

⁷A hackathon is an event in which people work in groups on projects over a short period of time, with the goal of creating a functioning product by the end of the event.

⁸FFATA, §§ 2(a)(7), (b)(1), (c); 4(b)(2).

⁹For the purposes of this report, data governance is defined as the processes, policies, standards, organization, and technologies required to manage and ensure the availability, accessibility, quality, consistency, auditability, and security of data in an organization. See Panian, Zeljko. "Some Practical Experience in Data Governance." *World Academy of Science, Engineering, and Technology*, no.62 (2010): 939-946.

of the data submitted under the DATA Act in November 2017.¹⁰ In 2019, we will report an updated assessment of the quality of these data and, separately, report on OMB and Treasury's efforts to develop a governance process for maintaining the integrity of the data standards established under the DATA Act. Detailed information on our scope and methodology is included in appendix I.

We conducted this performance audit from February 2017 to December 2018 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

Federal Open Data Policy

Recognizing the federal government's role as a major supplier of data, the 2018 President's Management Agenda announced the creation of a Federal Data Strategy.¹¹ According to the agenda, this strategy promises to leverage data as a strategic asset to grow the economy, increase the effectiveness of the federal government, facilitate oversight, and promote transparency. It proposes improving data dissemination by making data available more quickly and in more useful formats, maximizing nonsensitive data shared with the public, and enabling external users to access and use government data for commercial and other public purposes.

The Federal Data Strategy builds on existing policy governing the federal government's websites and data. In 2016, OMB memorandum M-17-06, *Policies for Federal Agency Public Websites and Digital Services*, established policy for the federal government's online information resources, such as the need to ensure that information is searchable and to inform users about information quality issues. In addition, in 2013,

¹⁰See GAO, *DATA Act: OMB, Treasury, and Agencies Need to Improve Completeness and Accuracy of Spending Data and Disclose Limitations*, [GAO-18-138](#) (Washington, D.C.: Nov. 8, 2017).

¹¹See OMB, *The President's Management Agenda* (Mar. 19, 2018).

OMB memorandum M-13-13, *Open Data Policy—Managing Information as an Asset*, established an information management framework to promote interoperability and openness at each stage of the information life cycle.

These efforts are consistent with the international Open Government Partnership, which aims to make governments more inclusive, responsive, and accountable. Seventy-five countries have committed to the Open Government Partnership by endorsing the Open Government Declaration. In doing so, these countries have committed to increasing the availability of information about government activities, supporting public participation in government, and using new technologies for openness and accountability, among other things.

DATA Act and USAspending.gov

Enacted in 2006, FFATA requires agencies to report information on federal awards—such as contracts, grants, and loans. In 2014, the DATA Act expanded on FFATA by establishing new requirements intended to allow policymakers and the public to more effectively track federal spending, including:

- **Reporting additional data.** Agencies are required to report additional financial data from different points in the spending life cycle.
- **Setting government-wide standards.** OMB and Treasury are responsible for establishing government-wide financial data standards for any federal funds made available to or expended by federal agencies.¹² These standards define and describe the data elements that agencies must report.
- **Reporting consistently.** Agencies reporting financial information are required to comply with the standards established by OMB and Treasury so that information can be compared across the government.¹³
- **Improving data access.** The data must be made available in machine-readable and open formats, to be downloaded in bulk, and—to the extent practicable—for automated processing.¹⁴

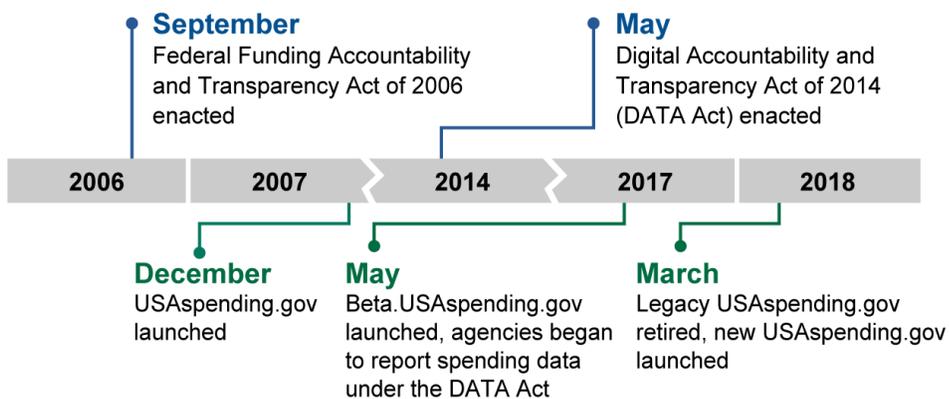
¹²FFATA, § 4(a), (b).

¹³FFATA, § 4(c)(2)(A).

¹⁴FFATA, § 2(c)(7).

The DATA Act required agencies to begin reporting data in accordance with the data standards issued by Treasury and OMB within three years of its enactment, and required that those data be displayed on USAspending.gov or a successor system. USAspending.gov has been the platform to provide federal spending information to the public since 2007 (see figure 1). In May 2017, Treasury released a new website, Beta.USAspending.gov, where it began to publish information submitted under the DATA Act. In March 2018, this new website assumed the USAspending.gov web address and Treasury retired the old version of USAspending.gov.¹⁵

Figure 1: Timeline of Federal Spending Transparency Laws and Versions of USAspending.gov



Source: GAO analysis of applicable laws, USAspending.gov, and beta.USAspending.gov. | GAO-19-72

Data on USAspending.gov come from a variety of sources, including files that agencies began submitting quarterly for DATA Act reporting in May 2017. When agencies submit quarterly data, Treasury’s DATA Act Broker ingests the data and validates certain information before the data are published on the website.¹⁶ Agency Senior Accountable Officials certify that the agency’s submission is valid and reliable. In addition to agencies’

¹⁵In this report, we refer to the version of USAspending.gov that was available prior to March 2018 as “legacy USAspending.gov.” In contrast, “USAspending.gov” refers to the website that was released as Beta.USAspending.gov in May 2017 and assumed the USAspending.gov web address in March 2018.

¹⁶The DATA Act Broker is an information technology system that collects and validates agency data, and extracts information from government-wide reporting systems. For more information on the sources and validation process for DATA Act data, see [GAO-18-138](#).

quarterly DATA Act reporting files, USAspending.gov includes data from government-wide systems. Government-wide procurement data on the website are updated daily, while government-wide financial assistance data are updated biweekly. The new USAspending.gov also includes older award data that had been available on the prior version of the website.

Our Prior Report Identified Data Quality and Website Issues

In November 2017, we issued our first report on data quality as required by the DATA Act.¹⁷ We found issues with the completeness and accuracy of the data that agencies submitted for the second quarter of fiscal year 2017 as well as the use of data elements. For example:

- Of the 78 agencies that submitted data on time, 13 submitted the data file intended to link budgetary and award information without providing any data in the file.
- Between 56 and 75 percent of the newly-required budgetary records were fully consistent with agency sources, but only between 0 and 1 percent of award records (such as grants, contracts, and loans) were fully consistent.¹⁸
- Agencies differed in how they interpreted and applied OMB's definitions for two data elements—*Primary Place of Performance* and *Award Description*—raising concerns regarding data consistency and comparability. These two award data elements are particularly important to achieving the transparency goals envisioned by FFATA because they provide the public with information on where the federal government spends money and what it spends it on, respectively.

We also found issues with the presentation of the data on Beta.USAspending.gov, including fragmented or incomplete search results and insufficient disclosure of data limitations. Among other things, we recommended that Treasury disclose known data quality issues and limitations on the new USAspending.gov website. We provide an update

¹⁷[GAO-18-138](#).

¹⁸These estimates came from a statistical sample which we selected from the population of all 24 Chief Financial Officers Act of 1990 (CFO Act) agencies' USAspending.gov data submitted from January through March of 2017. The spending data for the 24 CFO Act agencies represented approximately 94 percent of all records on Beta.USAspending.gov. A record was considered "fully consistent" if the information it contained matched agency sources for every applicable data element. We found that the consistency of individual data elements varied.

on actions Treasury has taken to address this recommendation later in this report.

Key Practices for Transparently Reporting Government Data

We identified five key practices that managers of open government data programs can consider to help ensure the transparent presentation of their data. We also identified key actions for implementing each key practice. We identified these key practices and key actions by systematically evaluating and synthesizing information from literature on open data as well as interviews with open data experts and good governance groups. These key practices and key actions are listed in table 1.

Table 1: Key Practices and Key Actions for Transparently Reporting Open Government Data

Key Practices	Key Actions
 <p>Provide free and unrestricted data</p>	<p>Make government data open by default, while protecting sensitive or restricted information</p> <p>Do not charge users for access to the data</p>
 <p>Engage with users</p>	<p>Identify data users and their needs</p> <p>Solicit and be responsive to user feedback</p> <p>Reach out to potential users to encourage data use</p>
 <p>Provide data in useful formats</p>	<p>Provide users with detailed and disaggregated data</p> <p>Provide machine-readable data that can be downloaded in bulk and in selected subsets</p> <p>Provide data downloads in a non-proprietary format</p> <p>Make the data interoperable with other datasets</p>
 <p>Fully describe the data</p>	<p>Disclose known data quality issues and limitations</p> <p>Disclose data sources and timeliness</p> <p>Clearly label data and provide accompanying metadata</p> <p>Publish data under an open license and communicate licensing information to users</p>
 <p>Facilitate data discovery for all users</p>	<p>Provide an interface that enables intuitive navigation and ensure that the most important information is visible</p> <p>Provide users with appropriate interpretations of the data, such as visualizations or summaries</p> <p>Ensure that the website's content is written in plain language</p> <p>Provide a search function that is optimized for easy and efficient use</p> <p>Use central data repositories and catalogues to help users easily find the data they are looking for</p>

Source: GAO. | GAO-19-72

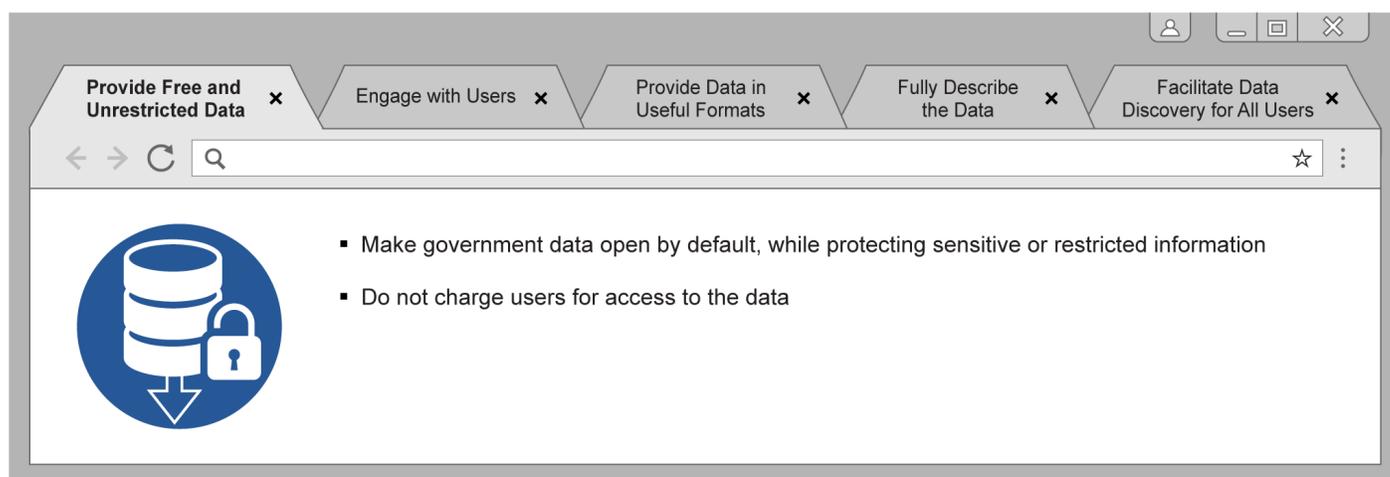
These key practices and actions are intended to be used in tandem with requirements for federal government websites and open data programs, such as relevant laws and OMB guidance. They are not intended to replace or supersede any applicable requirements. When considering an

individual open government data program, some key practices and actions may be more relevant than others because the purpose and characteristics of open government data programs may vary. In addition, while this report focuses on the presentation of open government data, open data practitioners should also consider other elements—including data quality and data governance—to ensure that the public has access to high-quality information.

Provide Free and Unrestricted Data

To promote transparency, we found that open data should be freely and equally available to users without restrictions. As such, we identified two key actions for providing free and unrestricted data (see figure 2).

Figure 2: Key Actions for Providing Free and Unrestricted Data



Source: GAO. | GAO-19-72

Make government data open by default, while protecting sensitive or restricted information. Making government data open by default ensures that the data are equally open to all types of users; in contrast, when government information is available by request, it may favor citizens with greater information about and access to government institutions. In addition, some open data practitioners we spoke with said that providing open data can minimize the burden of responding to information requests. For example, according to Connecticut officials, before the state’s open spending data website was launched, payroll data were the most frequently requested information under the state’s Freedom of Information Act (FOIA). Officials said that providing open payroll data on the website

significantly reduced FOIA requests, which allowed state officials to spend time and resources on other activities.

However, not all government information is appropriate to publish. Some datasets may contain sensitive information such as personally identifiable information, information that is classified or similarly not subject to disclosure, or intellectual property. Other legal restrictions may also prohibit the disclosure of certain information. In some cases the information in an individual dataset may not pose a risk of identifying sensitive information, but may pose such a risk when combined with other available information. For that reason, when considering whether or not information may be disclosed, OMB M-13-13 requires agencies to determine whether it may be combined with existing publically available data to identify an individual or otherwise pose a security concern.¹⁹ In such situations, agencies must conduct a risk-based privacy analysis to determine whether the information can be made publicly available that accounts for the nature of the information, the availability of other information, and the technology in place that could facilitate the process of identification.

As an example of how open data practitioners can balance these types of considerations, Montgomery County, Maryland, applies safeguards such as a review by internal departments. Additionally, if Department officials request a secondary fact-specific review, the Office of the County Attorney will review the information to further ensure that protected information is not published. In some cases, sensitive information is removed from a dataset prior to publication. For example, according to county officials, the names of housing assistance recipients are removed from spending data to protect resident privacy. Users can see nonsensitive aspects of these data, such as the amount spent, with identifying details removed.

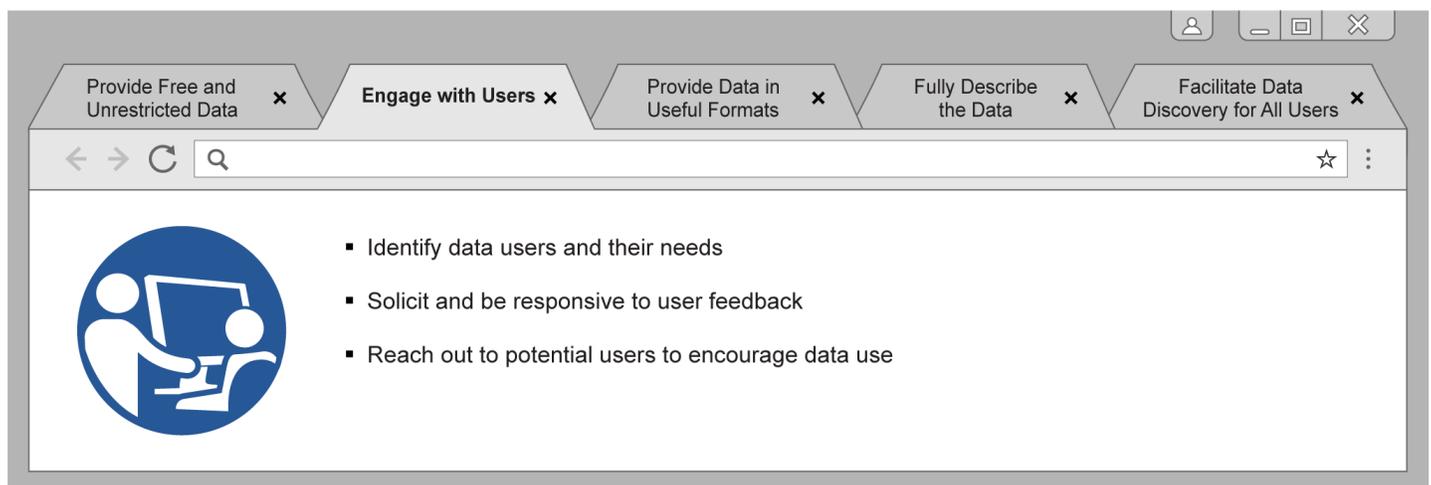
Do not charge users for access to the data. Providing data for free can help ensure equitable access to users independent of their ability to pay. Lowering barriers, such as cost, increases the value of open data, as more users are able to access it.

¹⁹See OMB, *Open Data Policy—Managing Information as an Asset*, M-13-13 (May 9, 2013).

Engage with Users

Open government data only create value to the extent that they are used. With that in mind, we identified three key actions for engaging with users (see figure 3).

Figure 3: Key Actions for Engaging with Users



Source: GAO. | GAO-19-72

Identify data users and their needs. By identifying who is using the data and what content or features are important to them, data providers can better prioritize their efforts to present information to data consumers. Open data experts we spoke with emphasized that data providers should engage with users both inside and outside of government, including groups that may typically have less access to government institutions. For example, to further New York City's Open Data for All vision to provide open data for people from all walks of life and all five of the city's boroughs, Columbia University students conducted user research on behalf of the city to better understand the extent to which community organizations use open data and what barriers they face, according to the capstone report for this project. By surveying and interviewing these organizations, the students learned that users found the city's data portal interface difficult to use. In response, the city worked with users to design and test a new, more streamlined portal.

Solicit and be responsive to user feedback. Soliciting and being responsive to user feedback—both when the website is being developed and on an ongoing basis—can help ensure that the website meets users'

needs. Feedback can also surface issues with the functionality of the website and the quality of the data, thus enabling the data provider to make corrections when needed. User feedback mechanisms vary and can include online comment forms, forums, and discussion boards, as well as in-person public forums. Open data experts we spoke with said it is particularly helpful to list the contact information for a responsible official on the website in case users have questions about the data. In addition, timely response to feedback encourages engagement by assuring users that their voices have been heard.

Monitoring how the public is using the data can also help practitioners determine which content and features are most useful. Web analytics can show how the data are being used, such as by identifying commonly-used search terms and datasets, and showing trends over time.

Web Analytics

Web analytics is the collection, reporting, and analysis of website data, such as the number of users who visit the website.

Source: GAO. | GAO-19-72

According to a city official, Los Angeles uses web analytics to measure how frequently its datasets are accessed. Web analytics data showed that some datasets were often accessed on certain dates or in conjunction with current events, while other datasets were rarely used. City officials use this information to adjust how data are presented on the website, which has increased overall use of the city's data. For example, the city created data visualizations and links to data—including its City Revenue and City Budget Expenditures datasets—to accompany the release of its Comprehensive Annual Financial Report.

Reach out to potential users to encourage data use. Actively engaging potential users can provide an opportunity to educate them on how the data can be appropriately used and encourage innovation. Data trainings can provide potential users with important context and information, which can include teaching users how the data can be used. Resources such as how-to guides can also encourage data use. For example, as shown on the website, New York City's open data portal includes a "How To" page with a step-by-step guide to help users get started with open data, and directs them toward additional resources such as data dictionaries.

We previously found that open data collaboration and prize competitions or challenges are two strategies that agencies can use to harness the ideas, expertise, and resources of those outside of their organization.²⁰

²⁰See GAO, *Open Innovation: Practices to Engage Citizens and Effectively Implement Federal Initiatives*, [GAO-17-14](#) (Washington, D.C.: Oct. 13, 2016).

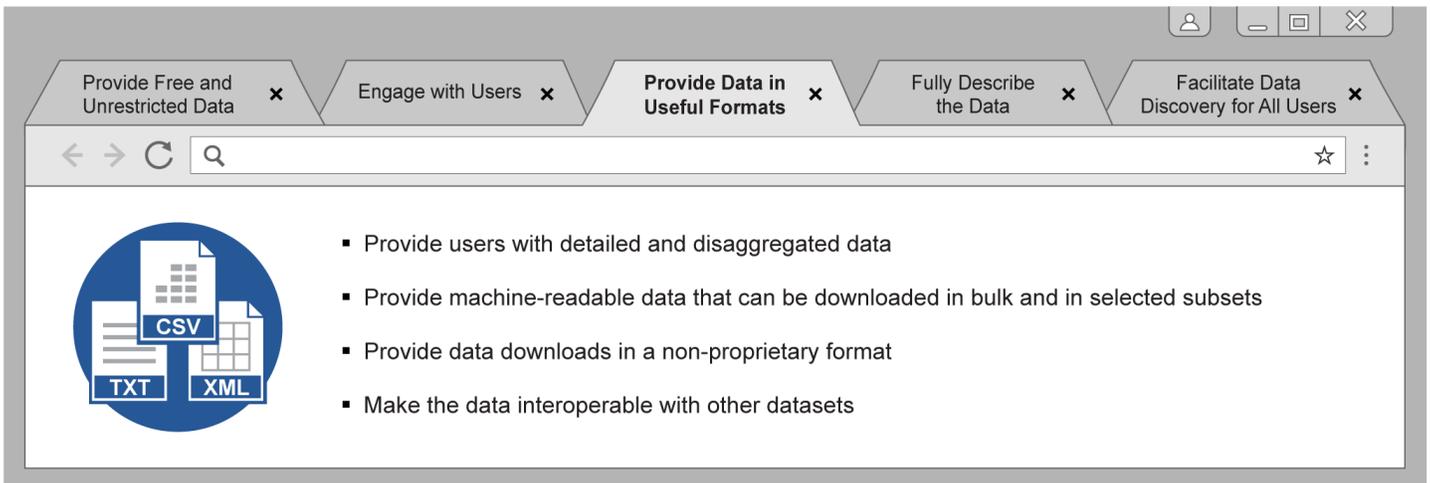
Agencies engage in open data collaboration by mobilizing participants to use their open data in innovative ways, such as sharing, exploring, and analyzing publicly-available datasets; using the data to conduct research; designing data visualizations; or creating web and mobile applications and websites that help people access and use the data. In addition, agencies use prize competitions or challenges for help solving a problem or reaching a specific goal by asking members of the public to submit potential solutions. The agency then evaluates these proposals and provides a monetary or nonmonetary award to selected winners.

New York City encourages the use of its open data using these strategies by hosting data literacy trainings, hackathons, and contests. For example, in the spring of 2018 the city hosted a contest to recognize projects that effectively use its open data and showcase the diversity of potential uses, according to city officials and contest documentation. Winning projects were posted to a gallery on the city's open data website. As shown on the city's open data website, one winner—a project called "Plan(t)wise"—predicts various tree species' likelihood of survival in locations throughout the city based on tree census data, and recommends which type of tree to plant at a given address.

Provide Data in Useful Formats

Data are most useful when they are provided in formats that allow them to be analyzed in a variety of ways. We identified four key actions for providing data in useful formats (see figure 4).

Figure 4: Key Actions for Providing Data in Useful Formats



Source: GAO. | GAO-19-72

Provide users with detailed and disaggregated data. Data are most useful when they are provided in as much granularity as possible.²¹ For example, Ohio’s online checkbook allows users to view detailed, disaggregated data in a user-friendly checkbook format, as shown on the website (see figure 5). The representation of the expenditure is displayed as a check, and includes the vendor’s name and address, the amount paid, payment date, check number, and contact information for the appropriate state office.

²¹While providing disaggregated data is optimal, we acknowledged earlier that, in some cases, aggregation may be necessary to protect sensitive information so that the dataset can be released to the public. For example, sensitive data—such as the names of individuals receiving housing assistance—can be aggregated by providing summary statistics rather than individual data points.

Figure 5: Example of Transaction-Level Spending Data from Ohio’s Online Checkbook

Close

Department of Education	5/3/2018	Check# 0006231528
Pay to the order of CINCINNATI PUBLIC SCHOOLS		
Eight Million Nine Hundred Two Thousand One Hundred Sixty Dollars and Thirty Five Cents		\$8,902,160.35
Memo Grant - Paid to Government or Non-profit Entity		
<small>Department of Education > 2018 > Subsidies Shared Revenue > Subsidies & Shared Revenue > Grant - Paid to Government > Grant - Paid to Government or Non-profit Entity > Property Tax Reimbursement Education > Homestead Property Tax Allocation > General Revenue</small>		

Who's getting paid?	Where's the money coming from?
<p>This \$8,902,160.35 payment for Grant - Paid to Government or Non-profit Entity is part of a larger check total of \$13,040,380.99 from the Department of Education. If you would like to see a list of the other payments included in Voucher ID 02382499, click here.</p>	
<p>For questions regarding this payment, contact: Department of Education Email: contact.center@education.ohio.gov Phone: (614) 995-1545 Mail: 25 South Front Street Columbus, Ohio 43215</p>	
<p>Paid By: Department of Education Payment Amount: \$8,902,160.35 Voucher ID: (EDU01) 02382499 Check Number: 0006231528 Payment Date: 5/3/2018 Expense Type: Grant - Paid to Government or Non-profit Entity</p>	
<p>Paid To: CINCINNATI PUBLIC SCHOOLS Vendor City: CINCINNATI Vendor State: OH Vendor Zip: 45201-5384</p>	

This is not an image of a real check. This check is an illustration of payments made using public funds.

Source: Ohio's Online Checkbook (<http://Ohiotreasurer.gov/Transparency/Ohios-Online-Checkbook>), accessed September 14, 2018. | GAO-19-72

Provide machine-readable data that can be downloaded in bulk and in selected subsets. Providing data in machine-readable formats makes them easier to process and analyze, which is particularly important for large datasets. For example, Kansas City officials said the city has been working to convert information from the PDF format to machine-readable formats such as CSV because PDF documents are challenging for the city to update and for users to navigate. In one instance, officials said that converting the city’s list of vehicles for sale in its tow lot from PDF to CSV format allowed the city to update the data more frequently so that users can see what vehicles are for sale at any given time.

Application Programming Interface (API)

An API sets up machine-to-machine communication, which can allow users to obtain real-time data updates.

Source: GAO. | GAO-19-72

Making data available to download in bulk allows users who need the full dataset to easily access it rather than retrieving information record-by-record. If the full dataset is large, allowing users to download selected subsets can make it easier for them to work with only the data they need. Data can also be provided to users through an Application Programming Interface (API), which allows users to connect directly with the dataset by enabling machine-to-machine communication. APIs can be particularly useful for large, frequently updated, or highly complex datasets because

they offer users flexibility to obtain the data they need. In addition, developers can use APIs to build applications based on the data.

Non-Proprietary File Formats

File formats describe what type of information a file contains, as well as how that information is stored and structured. Some file formats are proprietary, meaning that they can only be opened by specific commercial software applications. In contrast, non-proprietary formats are publicly available and can be used by all software developers. Examples of non-proprietary file formats include:

- CSV, which stores tabular data;
- RDF, which stores metadata;
- TXT, which stores unformatted text; and
- XML, which stores both the format and content of data.

Source: GAO. | GAO-19-72

Provide data downloads in a non-proprietary format. To ensure broad and equitable access, data downloads should be available in formats that do not require specific commercial software to access, and therefore do not exclude users who do not have access to such software. Non-proprietary data formats include, but are not limited to, CSV, RDF, TXT, and XML. For example, Kansas City, Missouri's, open data portal allows users to export spatial data in an open format that does not require proprietary mapping software, according to city officials and the city's open data portal website. Open data experts we spoke with said that practitioners should consult stakeholders when determining which format is appropriate for a given program, and that the appropriate format may change over time as technology advances.

Make the data interoperable with other datasets. Making data interoperable with other datasets can make them more useful because users may want to create new opportunities for analysis by linking datasets together. This can be done by standardizing the way that the data are reported. For example, using standard definitions for the specific items included in a set of data—known as data elements—can promote consistency with other datasets. Additionally, documentation such as data dictionaries can help ensure that definitions are clear and avoid misunderstandings.

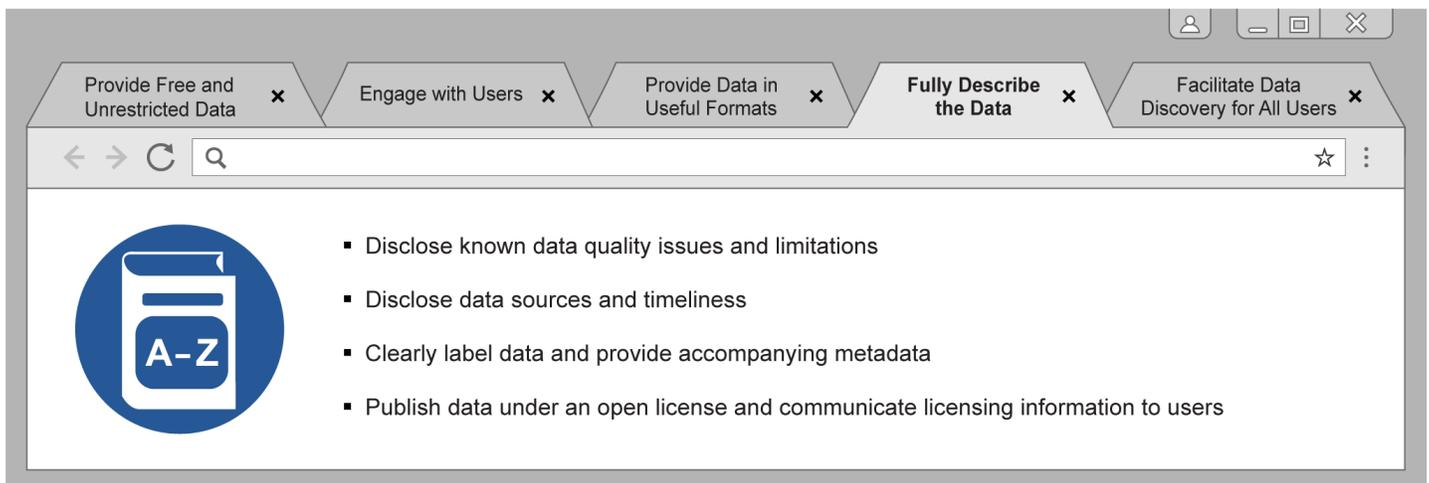
To promote interoperability between datasets that use geographic information, Kansas City uses standard land parcel identification numbers across departments. This allows different datasets that contain location information to be used in combination. For example, officials said that the city is linking different datasets that use those identification numbers—including building code violation data, 311 calls, and dangerous buildings data, among others—to build a model to prioritize code enforcement inspections.²²

²²311 is a non-emergency phone number that people can call in many cities to find information about services or report problems such as road damage, among other things.

Fully Describe the Data

Providing information about a dataset allows users to determine whether it is suitable for their intended purpose, and make informed decisions about whether and how to use it. With that in mind, we identified four key actions for fully describing the data (see figure 6).

Figure 6: Key Actions for Fully Describing Data



Source: GAO. | GAO-19-72

Disclose known data quality issues and limitations. Disclosing data quality issues and limitations helps users make informed decisions about whether and how to use the data. Disclosure of data quality issues and limitations can include descriptions of the completeness, timeliness, and accuracy of the data, such as an explanation of why certain data may not be disseminated. For example, we observed that Connecticut’s “OpenCheckbook” website includes an “About” page explaining that some information is excluded to protect privacy, or because it is not processed through the state’s financial system, such as the state’s Airport Authority, jury duty payments, and unclaimed property.

Disclose data sources and timeliness. Disclosing where the data come from and how frequently they are updated provides context that helps users judge their quality and determine whether they can be appropriately used for the intended purpose. Without this information, users may view, download, or use data without full knowledge of the extent to which they are timely, complete, or accurate, and therefore could inadvertently draw inaccurate conclusions from the data.

Metadata

Metadata provide descriptive information about a dataset in a structured, machine-readable format. They describe aspects of the dataset—such as the source of the data and when it was last updated—in clearly delineated fields.

Source: GAO. | GAO-19-72

Clearly label data and provide accompanying metadata. In addition, data should be clearly labeled and accompanied by structured metadata so that users can easily find information about the dataset. Metadata describe the characteristics of data in clearly defined, machine-readable fields, which can include attributes such as the date the data were created or modified, or the license used, among other things. Structuring metadata in clearly defined fields makes it possible for search tools to filter and match content pertaining to those fields. As shown in figure 7, Kansas City’s budget data are accompanied by metadata showing when they were last updated, the source of the data, and the name and contact link for the dataset owner, among other things.

Figure 7: Metadata for Kansas City, Missouri’s, Budget Data

Budget for budget.kcmo.gov Budget

View Data Visualize Export API ...

Updated June 7, 2018

Data Provided by City of Kansas City, MO - Budget Office

About this Dataset

Updated	Topics
June 7, 2018	
Data Last Updated June 7, 2018	Category Budget
Metadata Last Updated June 7, 2018	Tags budget, open budget, budget.kcmo.gov, open
Date Created February 10, 2016	Licensing and Attribution
Views 385	License Public Domain
Downloads 1,263	Source Link https://budget.kcmo.gov
Data Provided by City of Kansas City, MO - Budget Office	Dataset Owner Eric Roche

Contact Dataset Owner

Source: “Open Budget KC” website <https://data.kcmo.org/Budget/Budget-for-budget-kcmo-gov/whg3-khkh>. | GAO-19-72

Publish data under an open license and communicate licensing information to users. Documentation for a dataset should also specify what license applies to the data because a data license provides users with information about how they may use the data, including whether

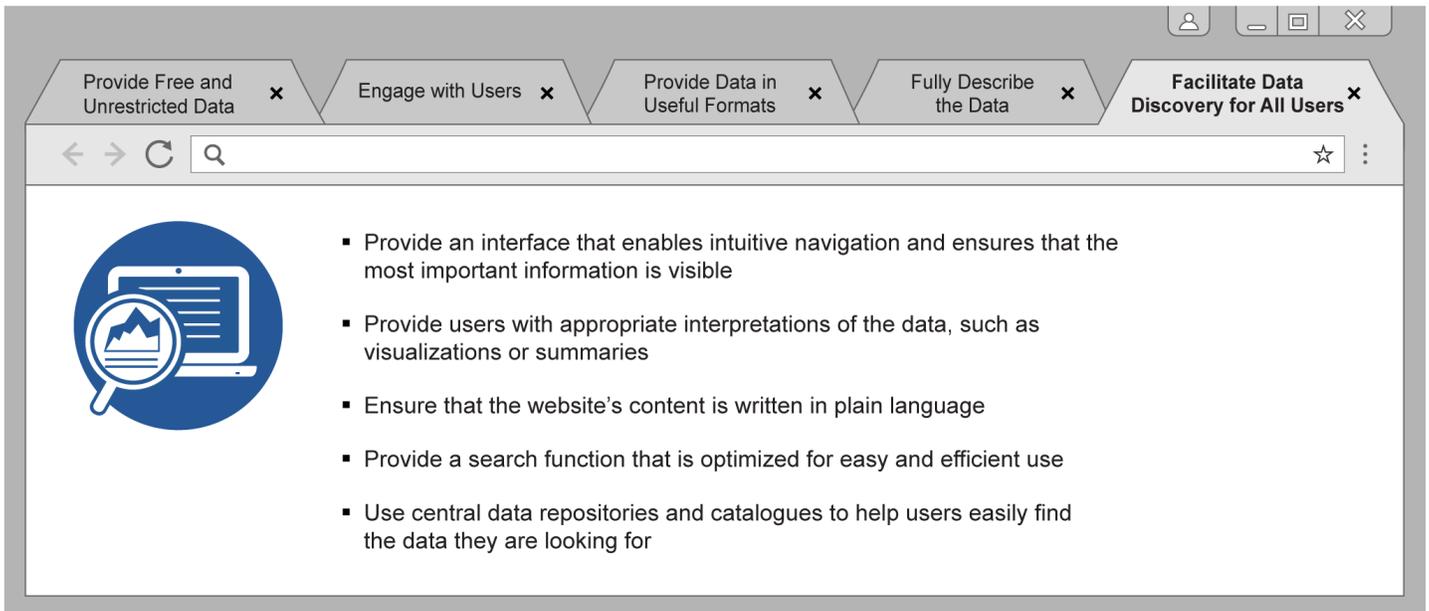
there are any restrictions, such as copyrights.²³ An open license indicates that there are few to no restrictions on how the data may be used. An open license can encourage innovation, for example, by assuring users that they are permitted to use the data to develop commercial applications. To realize these benefits, licensing information should be clearly communicated to users, ideally in machine-readable and human-readable formats. As shown in figure 7, metadata can be used to communicate licensing information in a clear and structured way. Including licensing information in metadata can help ensure that it is machine readable—which makes it easier to process and analyze—as well as help users discover the licensing information and compare it across datasets.

Facilitate Data Discovery for All Users

Data discovery is facilitated by presenting the data in a way that enables users to easily explore them. We identified five key actions for facilitating data discovery for all users (see figure 8).

²³While OMB M-13-13 generally requires federal agencies to apply open licenses, it also specifies that “when information is acquired or accessed by an agency through the performance of a contract, appropriate existing clauses shall be utilized to meet these objectives while recognizing that contractors may have proprietary interests in such information, and that protection of such information may be necessary to encourage qualified contractors to participate in and apply innovative concepts to government programs.”

Figure 8: Key Actions for Facilitating Data Discovery for All Users



Source: GAO. | GAO-19-72

Provide an interface that enables intuitive navigation and ensures that the most important information is made visible. To facilitate data discovery for all users, practitioners of open government data should ensure that the data are provided on a website that is simple and intuitive so that users can easily navigate it to find the information they need. Obtaining user feedback and conducting usability testing can help practitioners assess whether the website is easy to use, and identify any aspects that do not work well for users. In addition, websites designed to work on mobile devices, as well as mobile applications such as Ohio's "OhioCheckbook" app (see figure 9), can allow users to access data on a variety of devices, according to the state's website and our observations.

Figure 9: "OhioCheckbook" Mobile App Showing Payments to Cuyahoga Community College

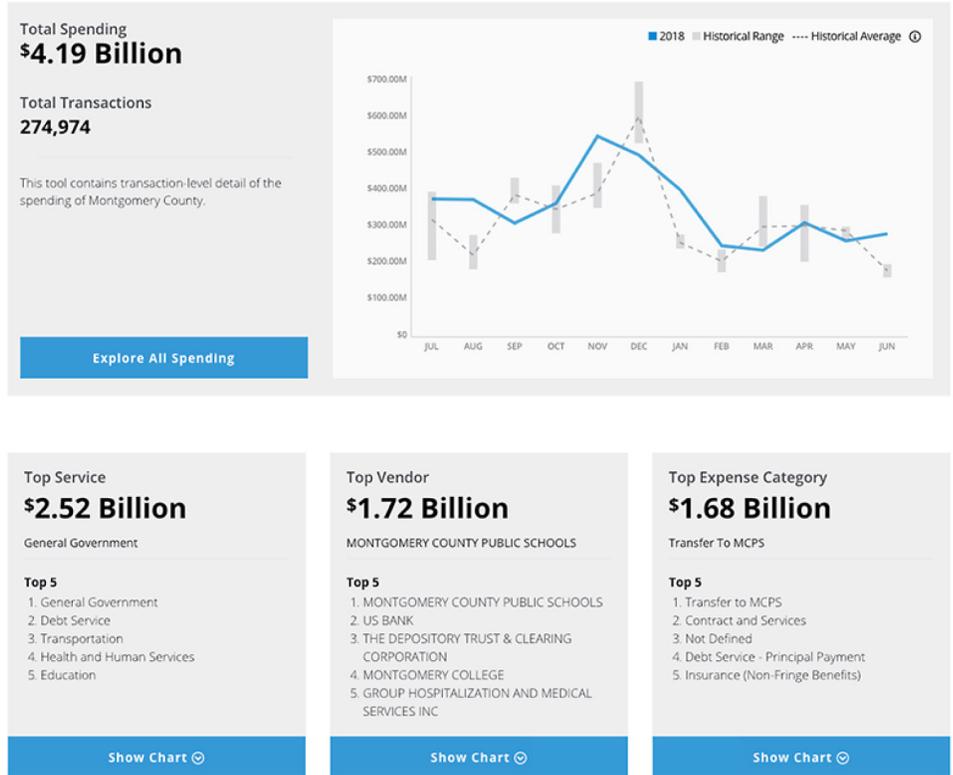


Source: "OhioCheckbook" Mobile App, accessed September 14, 2018. | GAO-19-72

Provide users with appropriate interpretations of the data, such as visualizations or summaries. Summaries and visualizations can help users explore data. For example, our review of Montgomery County, Maryland's, "spendingMontgomery" website found that the website provides summary data of the top five services, vendors, and expense categories with the greatest amount of spending, as well as a chart of annual spending along with historical averages, as shown in figure 10. This summary information provides a starting point for users, who can then navigate through the website to access more granular data.

Figure 10: Summary Data from the Montgomery County, Maryland's, spendingMontgomery Website

Spending Summary 2018 ▾



Source: Montgomery County, MD "spendingMontgomery" website (<https://spending.data.montgomerycountymd.gov>), accessed September 14, 2018. | GAO-19-72

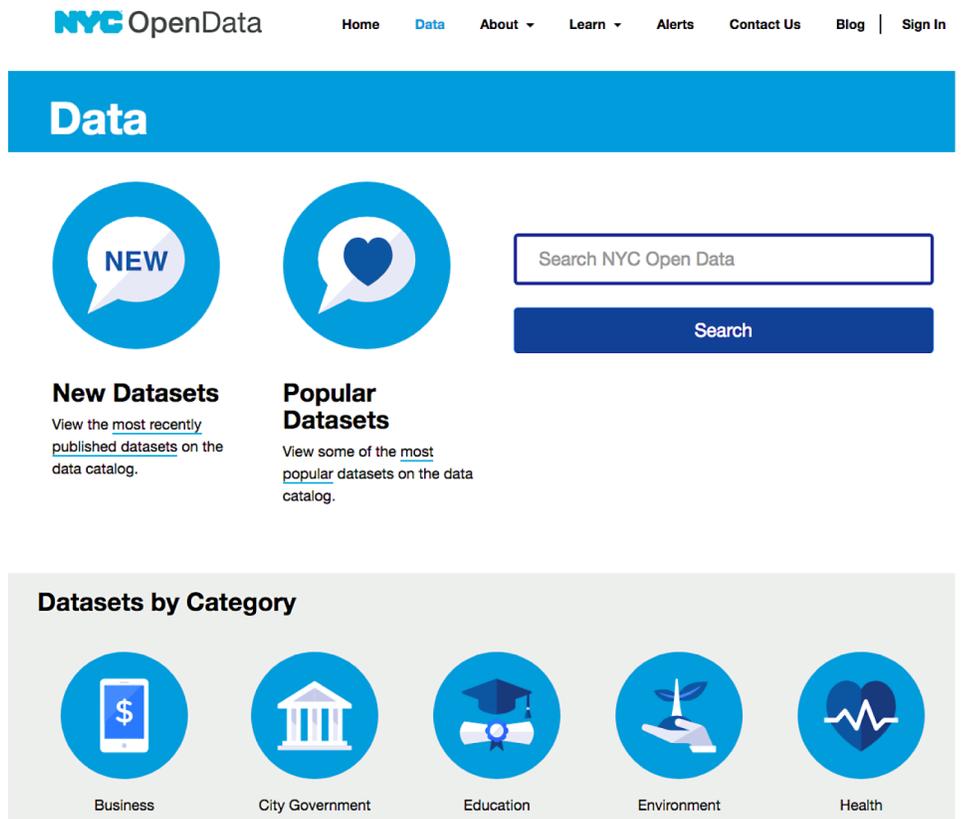
Ensure that the website's content is written in plain language. The content of an open data website should also be written in a way that is clear and direct, using plain language. Using commonly understood terms rather than technical jargon can help users understand the information provided. For example, to use well-understood terms when communicating budget information, Kansas City officials told us they participated in plain language training and applied that knowledge to the city's open budget data website. In addition, we found that in cases where it is necessary to use technical language, providing a glossary that defines key terms can help make the information understandable to users.

Provide a search function that is optimized for easy and efficient use. Open data websites should also include a search function that is optimized for easy and efficient use so that users can find information that is relevant to them. For example, Connecticut officials said that the search function on the state’s open spending data website is designed so that users do not need to be familiar with the state government’s structure or terminology to find meaningful results. When a user enters a search term, the search bar will return a list of items that include this term and a description of what they are. For example, when we typed “Education” into the search bar, the website suggested Department of Education spending, bilingual education programs, and a vendor called Family Life Education.

In addition, Connecticut officials told us that they track the most commonly-used search terms—such as “housing” and “voter turnout”—on the state’s open data portal, and test them to verify that the information is discoverable. Similarly, Ohio’s online checkbook includes a “Popular Searches” tool that provides presaved searches that allow users to see expenditures for a variety of categories—such as travel, roads and highways, or parks—by clicking a single button. In addition, officials told us that if a user’s search returns a large volume of results, a pop-up appears prompting the user to narrow their search, which could help them focus on more relevant information.

Use central data repositories and catalogues to help users easily find the data they are looking for. Central data repositories and catalogues—also known as data portals—are websites that provide a “one-stop shop” for users to access a variety of datasets. These websites host the data directly, link to other websites where users can access the data, or a combination of the two. They typically provide descriptions of the datasets, as well as structured metadata, to help users find data suitable for their purpose. New York City’s open data portal also includes a number of tools to help users find datasets, including a search function as well as lists of new datasets, popular datasets, and datasets by category, as shown in figure 11.

Figure 11: Data Discovery Tools on New York City's Open Data Portal



Source: "NYC OpenData" website (<https://opendata.cityofnewyork.us/data/>), accessed September 14, 2018. | GAO-19-72

USAspending.gov Aligns with Several Key Practices, but Does Not Fully Meet Licensing, Search Functionality, and Other Requirements

We found that USAspending.gov aligns with the key practices of providing free and unrestricted data and engaging with users. However, Treasury does not fully describe the data and two data elements required by law are not searchable. In addition, Treasury lacks a process to ensure all pages on the website are secure, consistent with federal requirements.

USAspending.gov Provides Free and Unrestricted Spending Data

Key Practice

Provide Free and Unrestricted Data



Source: GAO. | GAO-19-72

Spending data are open by default and sensitive information is protected. The Federal Funding Accountability and Transparency Act of 2006 (FFATA) requires the website displaying the data that agencies must provide to be accessible to the public at no cost.²⁴ In response to requirements in FFATA, as amended by the Digital Accountability and Transparency Act of 2014 (DATA Act), in May 2014, OMB and Treasury developed standard definitions for data elements for agencies to report, and Treasury displays these open data on USAspending.gov.²⁵ Agencies should not report classified or protected information, such as personally identifiable information (PII). However, they are required to aggregate some awards containing PII at the county or state level if they are unable to report spending at the individual level.

All data are available for free. Treasury has made all of the data on USAspending.gov available to users at no cost, as required by the DATA Act and FFATA. During the course of our work, we found that users could only download the complete database after registering for an account with the database host—Amazon Web Services. Further, we also found that users would incur a charge when attempting to download the entire database. Treasury officials said they intended for the data to be available

²⁴FFATA, § 2(b)(1).

²⁵FFATA does not require the public disclosure of information that is exempt from disclosure under sections 552 and 552a of title 5, United States Code (commonly known as the Freedom of Information Act and the Privacy Act of 1974, respectively) or taxpayer return information protected by section 6103 of title 26, United States Code. FFATA, § 7. We previously reported on OMB and Treasury's implementation of data standards for DATA Act reporting. See GAO, *DATA Act: Data Standards Established, but More Complete and Timely Guidance Is Needed to Ensure Effective Implementation*, [GAO-16-261](#) (Washington, D.C.: Jan. 29, 2016).

for free and were unaware that users were being charged to access the data. In response to our inquiries on this issue, in July 2018, Treasury resolved this issue and provided an option for users to download the entire database for free without creating an account.

Treasury Engages Users by Encouraging Feedback and Data Use

Key Practice

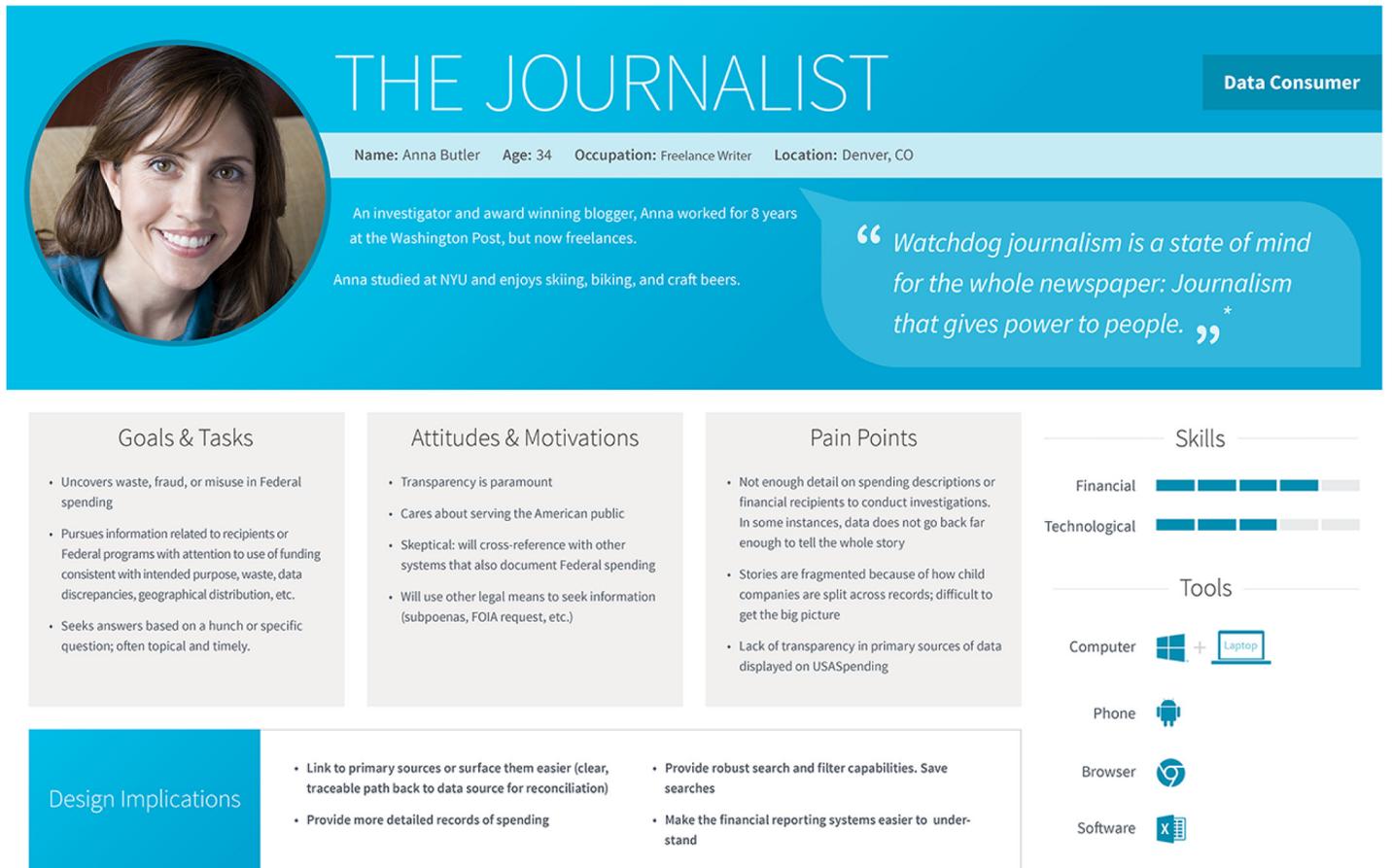
Engage with Users



Source: GAO. | GAO-19-72

Treasury identifies data users and their needs through user research. Treasury researches users to understand their needs when working with USAspending data. Treasury has developed profiles for eight types of users ranging from data consumers like “Citizen” or “Journalist” (see figure 12) to budget analysts or chief financial officers. These profiles are part of Treasury’s user-centered design process in which officials told us they learn from users, make changes to the website, and test whether those changes make the website more useful and intuitive to users.

Figure 12: Example of a Treasury User Profile



* <http://www.poynter.org/2005/slapping-a-headline-on-watchdog-journalism/67427/>

Source: Treasury's GitHub website (<https://fedspendingtransparency.github.io/>), accessed June 12, 2018. | GAO-19-72

Treasury obtains and responds to user feedback. Treasury officials told us that they track user feedback, which informs improvements they make to the website. We previously found that Treasury has a variety of user feedback mechanisms, including a community forum, one-on-one interviews, and a “contact us” link that allows users to provide feedback by email.²⁶ As of July 2017, Treasury officials said they had interviewed more than 130 users, such as citizens, funding recipients, and federal agency officials, regarding USASpending.gov website features. They have

²⁶GAO-18-138.

since conducted 20 additional interviews about the user experience and received feedback from another 130 users about the Data Lab, a related website that offers visual interpretations of the spending data. Treasury has also conducted “intercept” interviews where interviewers go to a location with large groups of people and request feedback about the website from random individuals. For example, figure 13 shows a Treasury contractor interviewing a visitor about an early version of USAspending.gov at the U.S. Capitol Visitor Center.

Figure 13: Treasury Contractor Interviewing a Member of the Public at the Capitol Visitor Center



Source: Booz Allen Hamilton website (<https://www.boozallen.com/e/insight/thought-leadership/making-usaspending-gov-agile-and-open-source.html>), accessed September 14, 2018. | GAO-19-72

Treasury officials said they respond to user feedback about USAspending.gov on two websites. They respond directly to user comments on the USAspending.gov Community website, where users can share feedback and find answers to frequently asked questions. Treasury officials told us they also track users’ issues as “stories” on an open development platform called JIRA, which is their primary way of documenting website development decisions and tracking potential improvements to the website.²⁷ For example, Treasury added new functionality to the Application Programming Interface (API) based on user feedback from agencies. Officials said this feature allowed some agencies to more efficiently manage their quarterly DATA Act submissions. Treasury announces updates to the API and other changes to the website via an email newsletter.

²⁷Treasury tracks users’ issues on the following website: <https://federal-spending-transparency.atlassian.net/secure/Dashboard.jspa>.

Treasury reaches out to potential users to encourage data use.

Treasury educates the public about the use of the spending data on USAspending.gov and the Data Lab through how-to guides and outreach activities. For example, the website offers an “API Guide” for users seeking to utilize computer programs to request and receive the data, and the Data Lab features an “Analyst Guide” that answers questions about using the data.

Treasury officials told us that they have directly engaged with various audiences about USAspending.gov. For example, they have engaged with the Syracuse University Maxwell School of Government to map the use of federal funds in New York State. In April 2017, we observed Treasury’s participation in a hackathon where participants developed ways to use federal spending data, including using the spending data to evaluate block grant formulas and track the economic impact of stimulus money. Treasury officials said they have held information sessions with Congress, federal agencies, and nongovernment organizations.

USAspending.gov Data Are Detailed and in Standardized Formats, but Treasury Lacks Certain Security Processes

USAspending.gov provides users with detailed and disaggregated data. As shown in figure 14, an award summary page on the website displays information on specific awards, including the awarding agency, recipient, award amount, description, and location. These pages also include transaction histories so that spending can be tracked over time. As of October 2018, we found that USAspending.gov listed more than 53 million pages of prime awards representing more than \$33 trillion in obligated funds between fiscal year 2008 and 2018.

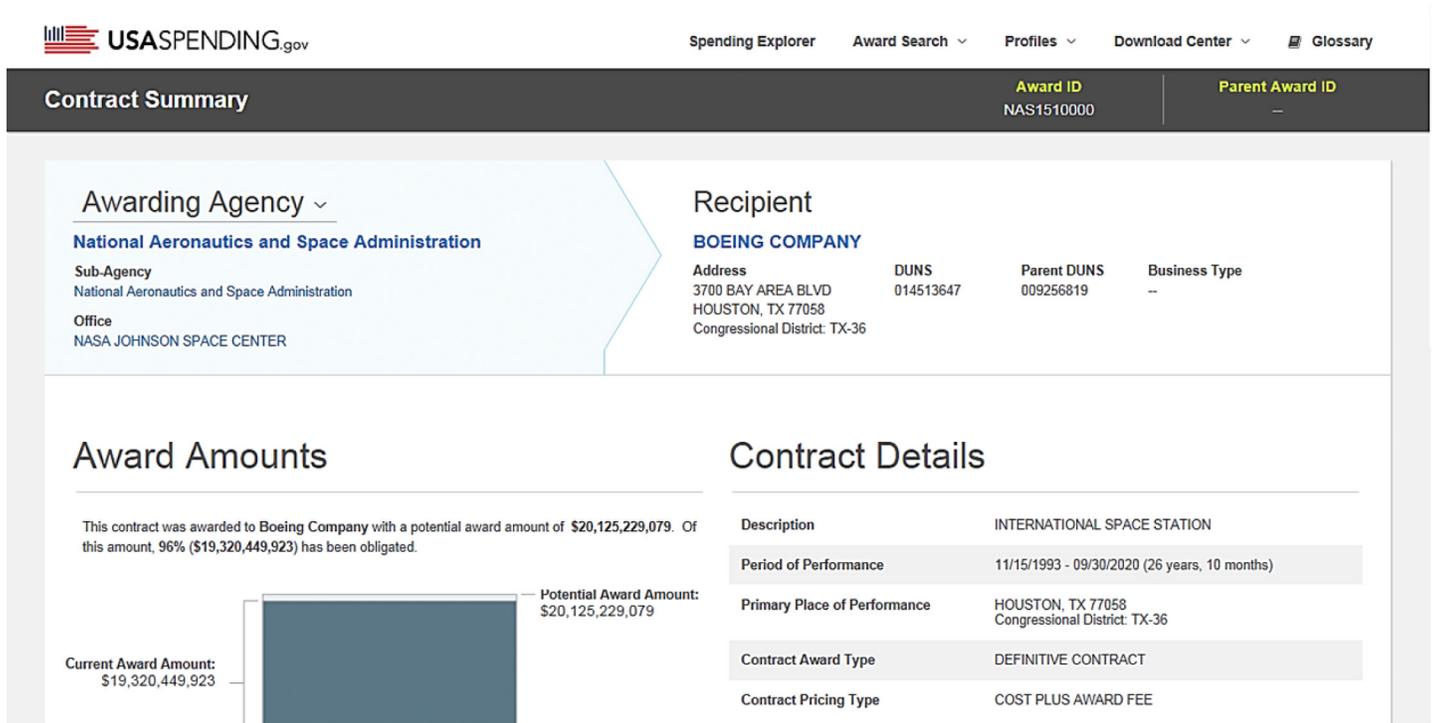
Key Practice

Provide Data in Useful Formats



Source: GAO. | GAO-19-72

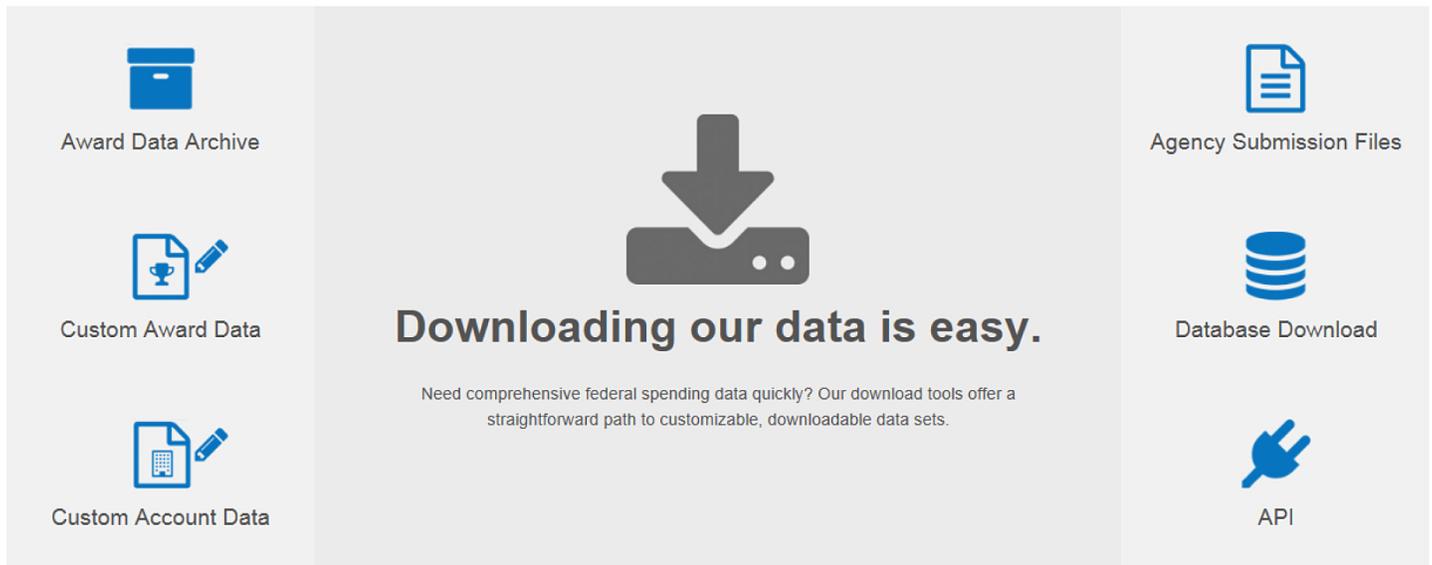
Figure 14: Example of Detailed and Disaggregated Data from a Contract Summary Page on USAspending.gov



Source: USAspending.gov Contract Summary (<https://www.usaspending.gov/#/award/27117525>), accessed September 14, 2018. | GAO-19-72

Data are machine readable and can be downloaded in bulk and in selected subsets, but Treasury lacks a process to fully ensure security. As shown in figure 15, USAspending.gov provides six ways for users to download the data, including subsets of the data or the complete database. An API is also available, which enables machine-to-machine communication that allows real-time updates.

Figure 15: Download Options Available on USAspending.gov



Source: USAspending.gov website (<https://www.usaspending.gov>), accessed September 14, 2018. | GAO-19-72

During the course of our review, we found that some of the data download web addresses did not point to a government domain and were unsecured. OMB guidance requires that federal web pages be hosted on a .gov domain and be encrypted by the secure HTTPS protocol.²⁸ In response to our inquiries on this issue, Treasury updated USAspending.gov in October 2018 so that the web pages for the database download and agency submission files use the secure HTTPS protocol and are on a government domain. As a result, the users requesting this information from USAspending.gov now have better assurance of the integrity of the data requested, the privacy of their connection to USAspending.gov, and that the website they are using is a trusted government domain.

Treasury officials said they take steps to ensure they meet federal information security requirements, but had not noticed that the web pages were unsecured or on a nongovernment domain. According to Treasury officials, the agency has a process for the team developing a website to vet whether proposed pages are secured and hosted properly, but they

²⁸OMB Memorandum M-17-06.

acknowledged unintended gaps in how the process was applied in this case, which caused some pages to be omitted. *Standards for Internal Control in the Federal Government* states that management should design control activities to achieve objectives and respond to risks.²⁹ Control activities are the policies, procedures, techniques, and mechanisms that enforce management's directives to achieve the entity's objectives and address related risks.

Until the gaps in Treasury's information security process are addressed, the agency does not have assurance that any new pages that may be added to USAspending.gov will conform to federal information security requirements. In response to our inquiries, Treasury provided documentation showing that the agency is in the process of addressing the issue to prevent future unintended gaps. The agency has taken initial steps to revise its process to ensure that all pages on USAspending.gov are secured and hosted properly. We will continue to monitor Treasury's efforts to develop and implement this new process.

Downloads are available in standard, non-proprietary formats.

Downloads of search results, agency files, and subsets of the USAspending.gov database are available in file formats that can be opened using common office software, including CSV and XML files.

Spending data are potentially interoperable with other datasets. The data in the USAspending.gov database are organized according to a government-wide standard which can potentially support interoperability with related government datasets. The DATA Act Information Model Schema (DAIMS) provides standardized definitions for federal spending information, including 57 data standards that federal agencies are required to report for DATA Act implementation. These standards come with technical specifications describing the format and structure of each data element, which are intended to facilitate consistent data reporting across the federal government, and allow for interoperability between agencies' data submissions.

According to Treasury's DAIMS Architecture document, DAIMS could eventually support interoperability between USAspending.gov and related nonfederal datasets such as state, local, and international spending

²⁹GAO, *Standards for Internal Control in the Federal Government*, [GAO-14-704G](#) (Washington, D.C.: September 2014).

data.³⁰ For example, state governments could make their data interoperable with the federal spending data on USAspending.gov by developing their own data standards and definitions aligned with DAIMS as appropriate. In addition, the DAIMS Architecture document specifies that future DAIMS content could include federal receipt and financing balances with accounts and sources, as well as performance measures and outcomes linked to federal grants, awards, or other financial assistance.³¹

Treasury Improved Descriptions of Data, but USAspending.gov Lacks Metadata and Complete Licensing Information

Key Practice

Fully Describe the Data



Source: GAO. | GAO-19-72

Website still does not completely disclose data quality issues.

Treasury has improved the disclosure of data quality issues and limitations, but other issues have not yet been described to users. In November 2017, we found that the website did not sufficiently disclose known limitations affecting data quality.³² We recommended that Treasury disclose known data quality issues and limitations. Treasury agreed with the recommendation and took the following steps to disclose limitations:

- By May 2018, Treasury had added a “Learn More” box to the website with information about the data, including an explanation that the Department of Defense reports its data later than other agencies.³³
- In June 2018, Treasury added information on unreported spending to the Spending Explorer tool that visualizes federal spending, clarifying that information reported on the website does not capture the totality of federal spending. Treasury explains to users that data might not be reported when an agency reports incomplete data, has a submission deadline extension, is not required to submit certain data elements, and for accounts that are not reported to Treasury.

³⁰The current version of the DAIMS Architecture and related documentation is available on the website of Treasury’s Bureau of the Fiscal Service:

<https://www.fiscal.treasury.gov/data-transparency/data-act-v1.2.html>.

³¹While the DATA Act does not require that DAIMS be expanded to include interoperability with state, local, and international spending or links with performance measures and outcomes, Treasury and OMB could choose to do so as a matter of discretion.

³²[GAO-18-138](#).

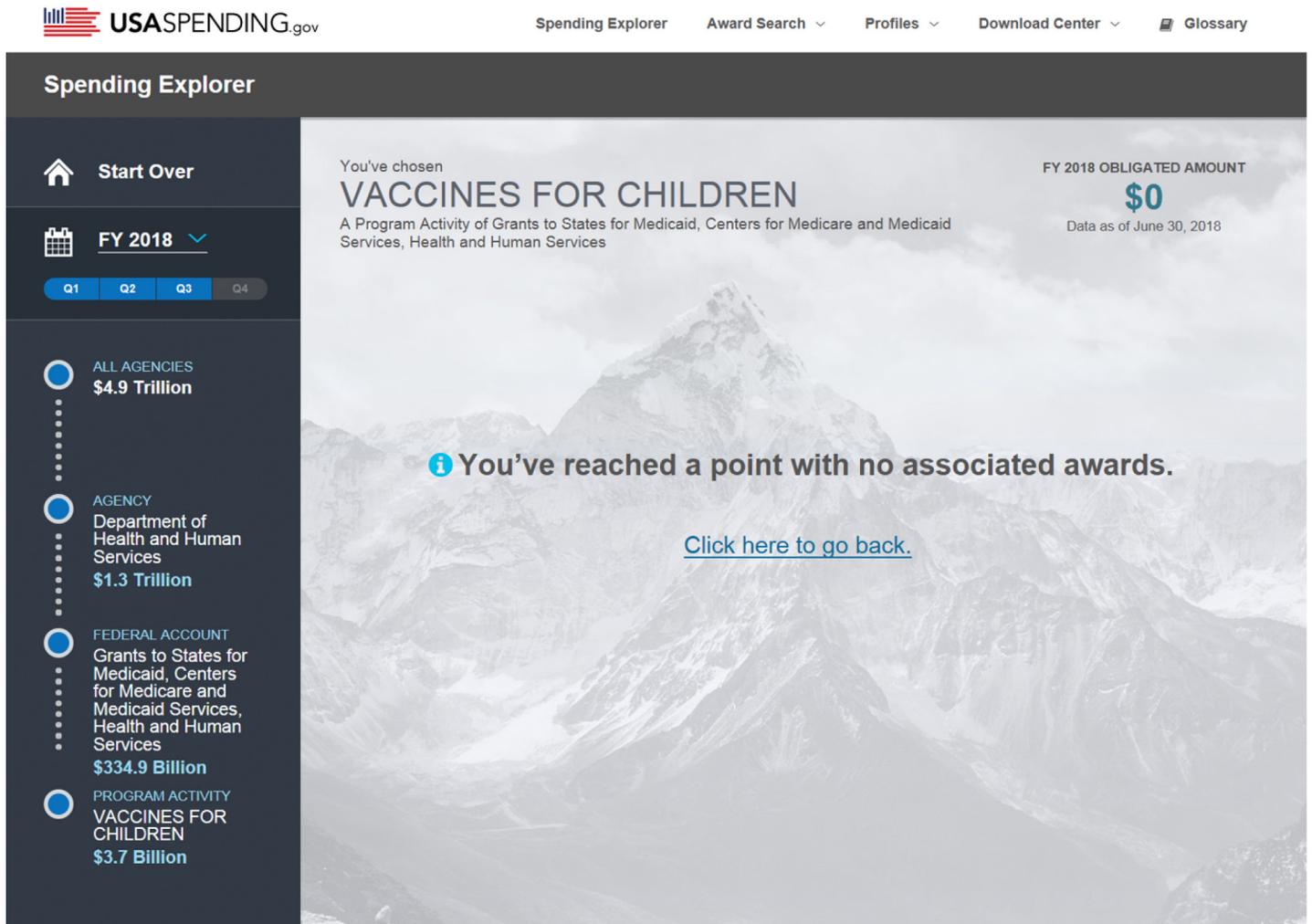
³³The Department of Defense’s procurement data are released to the public after a 90-day waiting period, which results in a 90-day delay in reporting these data to USAspending.gov. In addition, the DATA Act specifies that OMB may grant up to three extensions for the Department of Defense for reporting financial and payment data, each lasting up to 6 months. FFATA, § 4(c)(2)(B).

While the steps Treasury has taken are positive, they do not fully address our recommendation. This is because one purpose of the DATA Act is to allow users to track federal spending more effectively by linking specific awards to financial budgetary information. However, we found that award data do not appear in the Spending Explorer for combinations of certain agencies and program activities.³⁴ For example, as figure 16 shows, there are “no associated awards” for the program activity “Vaccines for Children” within the Department of Health and Human Services account for Medicaid grants to states.³⁵ However, the account page for this program elsewhere on USAspending.gov shows approximately \$3.6 billion in obligations and various associated awards for the first three quarters of fiscal year 2018. There is no context for a user to understand whether this information is required for this federal account, missing, or searchable elsewhere on the website.

³⁴A program activity is a specific activity or project as listed in the program and financing schedules of the President’s budget. For more information on budget terminology, see GAO, *A Glossary of Terms Used in the Federal Budget Process (Supersedes AFMD-2.1.1)*, [GAO-05-734SP](#) (Washington, D.C.: Sept. 1, 2005).

³⁵The Department of Human Service’s Grants to States for Medicaid account consists of three program activities: Vaccines for Children, Medicaid Vendor Payments, and State and Local Administration.

Figure 16: No Award Information in Spending Explorer for the Vaccines for Children Program Activity



Source: Spending Explorer on USAspending.gov (<https://www.usaspending.gov/#explorer>), accessed October 9, 2018. | GAO-19-72

Note: In the Spending Explorer, no associated awards are attributed to the Vaccines for Children program activity (i.e., spending project) within the Department of Health and Human Services account for Medicaid grants to states. However, the account page for this program elsewhere on USAspending.gov shows approximately \$3.6 billion in obligations and various associated awards for the first three quarters of fiscal year 2018.

Treasury officials informed us of a number of data limitations that could cause spending data and award data to be disconnected in the Spending Explorer, but these issues are not disclosed on the website. According to Treasury officials, agencies might not currently report certain data fields

as some fields are optional, there are inconsistencies between several agency data systems, and some agencies have not been able to link financial and award data. As a result, the Spending Explorer does not consistently provide a clear and complete presentation of federal spending, and because Treasury does not disclose these limitations, it could limit the ability of taxpayers and policy makers to fully track federal spending with this tool.

More broadly, we have raised concerns that USAspending.gov does not sufficiently disclose other, broader government-wide data quality issues. For example, we found in November 2017 that only between 0 to 1 percent of awards were fully consistent with agency records.³⁶ While the consistency of individual data elements varied, our prior report found inconsistencies with agency records in at least 41 percent of the data for *Award Description*, *Current Total Value of Award*, and *Primary Place of Performance Address* from the second quarter of Fiscal Year 2017.

Website discloses data sources and timeliness. The “About” page and “Frequently Asked Questions” describe data sources, data quality, and legal requirements. There is also a diagram on the “About” page showing how the data go from agencies to the database for USAspending.gov, and the frequency with which the data are updated, which is a useful way to visualize how the types of award data are updated either daily, bi-monthly, or every quarter.

Website labels some data, but lacks structured metadata. Treasury labels some of the data on USAspending.gov in tables and data visualizations, and describes it in narrative form. The website also includes data dictionaries that provide definitions for the data elements. However, the website lacks structured, machine-readable metadata. OMB guidance requires agencies to use metadata to describe their datasets so that all users can understand and process open data.³⁷ Agencies must consult with the best practices from Project Open Data, OMB’s online repository of tools and schema, to help agencies meet the requirements of its open data policy. According to Project Open Data, metadata are structured information that describe, explain, locate, or otherwise make it easier to retrieve, use, or manage datasets like that displayed on

³⁶[GAO-18-138](#).

³⁷See OMB Memorandum M-13-13 (May 9, 2013).

USAspending.gov. This guidance also indicates that making metadata machine readable greatly increases their utility.

Treasury officials said the types of information found in metadata are already available in a number of separate documents on Treasury's Fiscal Service web page.³⁸ Treasury officials told us that they decided not to provide structured metadata on USAspending.gov because it is more efficient to provide external links to other websites. Further, Treasury officials asserted that providing metadata on those websites is sufficient to comply with OMB guidance.³⁹

However, the information found on these various websites does not align with best practices outlined in Project Open Data, or the key action to clearly label data and provide accompanying metadata, because it is not provided in a single place on USAspending.gov as structured metadata in a machine-readable format. Without easy access to information that fully describes the data, it may be difficult or time consuming for users of USAspending.gov to find the information available on other websites, and determine whether or how to use the data for their purposes.

Website lacks complete licensing information. While the website describes restrictions on the use of proprietary contract data from Dun & Bradstreet Inc.'s Data Universal Numbering System (DUNS), it does not

³⁸The various documents that describe metadata for USAspending.gov include the Interface Definition Document, Reporting System Specifications, and the Online Data Dictionary. The DAIMS Consumption-Focused Metadata document is available on the site of Treasury's Fiscal Service at <https://www.fiscal.treasury.gov/data-transparency/data-act-v1.2.html>.

³⁹We found structured, machine-readable metadata for USAspending.gov on DATA.gov, a government repository of databases. These metadata were updated in September 2018. Metadata for "USAspending.gov Federal Award, Subaward, and Account Data" can be found at <https://catalog.data.gov/dataset/usaspending-gov-contracts-and-purchases-a7059>. Metadata for agency's raw submission files can be found at <https://catalog.data.gov/dataset/usaspending-gov-grants-and-loans-89a71>.

include general licensing information for the rest of the data.⁴⁰ The website includes a link to a notice specifying the “Limitation on Permissible Use of Dun & Bradstreet, Inc. Data.” According to Treasury officials, most data on USAspending.gov are in the public domain, but we found that the website does not clearly indicate which data are openly available to use without restrictions.

OMB M-13-13 specifies that federal agencies “must apply open licenses, in consultation with the best practices found in Project Open Data, to information as it is collected or created so that if data are made public there are no restrictions on copying, publishing, distributing, transmitting, adapting, or otherwise using the information for non-commercial or for commercial purposes.”⁴¹ According to OMB staff, agencies should include licenses in metadata so that this information is machine readable. If data access is limited, this should also be prominently featured in the metadata. In addition, Project Open Data specifies that licensing information should be provided in metadata.

Treasury officials said that the agency is evaluating options and approaches for including open data licensing information on the website, consistent with OMB M-13-13. In addition, Treasury officials said they had only received one question from users about licensing.

However, not displaying licensing information for the majority of data elements on the website is not consistent with the key action to publish data under an open license and communicate licensing information to

⁴⁰In June 2012, we found that Dun & Bradstreet had placed restrictions on DUNS numbers that limit the purposes for which the government can use the data, and hamper the ability to switch to a new numbering system. We recommended that the Administrator of GSA initiate discussions with Dun & Bradstreet on ways to reduce restrictions on the use of DUNS numbers. GSA agreed with our recommendation and took steps to consider alternatives. OMB staff told us in April 2018 that they were conducting an analysis of alternatives to DUNS that will be completed in December 2018. See GAO, *Government Is Analyzing Alternatives for Contractor Identification Numbers*, [GAO-12-715R](#) (Washington, D.C.: June 12, 2012).

⁴¹See OMB Memorandum, M-13-13 (May 9, 2013). As stated earlier in this report, M-13-13 also specifies that “when information is acquired or accessed by an agency through performance of a contract, appropriate existing clauses shall be utilized to meet these objectives while recognizing that contractors may have proprietary interests in such information, and that protection of such information may be necessary to encourage qualified contractors to participate in and apply innovating concepts to government programs.” Treasury’s disclosure of DUNS data restrictions is consistent with this provision.

users. Without licensing information for all of the data, users will likely be unable to determine what license, if any, applies to USAspending.gov, and it will be unclear to the public whether there are any restrictions to reusing data that they can download from the website.

USAspending.gov Makes Data Easy to Discover, but Does Not Fully Meet Search Requirements

Key Practice

Facilitate Data Discovery for All Users



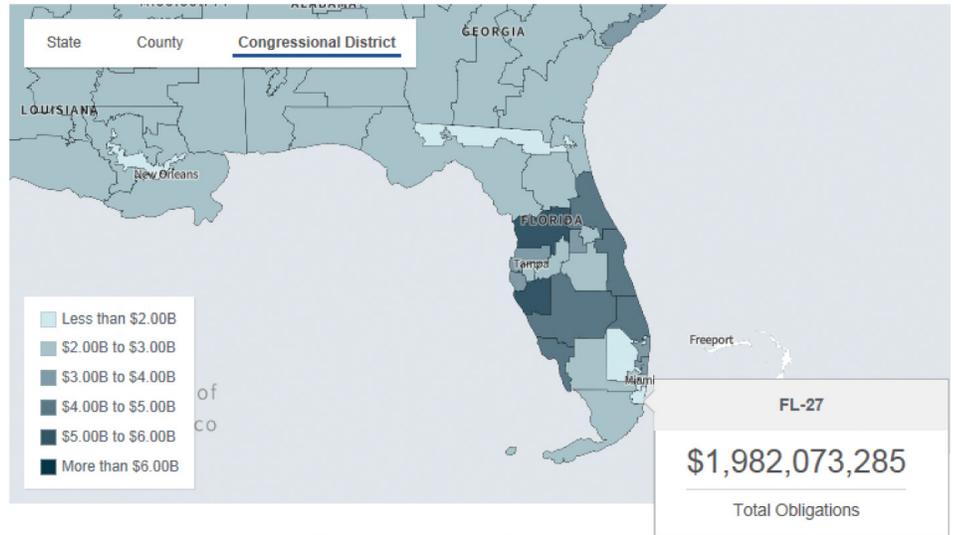
Source: GAO. | GAO-19-72

Website includes a user interface to assist navigation.

USAspending.gov's top menu gives users various ways to explore, search, download, and understand the most important information. The menu links to the Spending Explorer, Award Search, Profiles, Download Center, and Glossary. There is also "featured content" on the home page guiding users to the Data Lab, and other new features such as a download option for Federal Account data and recipient profile pages for any entity that has received federal money in the form of contracts, grants, loans, or other financial assistance.

Interactive visualizations enable exploration. Search results can be visualized by prime award or subaward aggregated in a table, in a chart showing awards over time, in a map showing the geographic distribution of awards (see figure 17 for an example of social security insurance results mapped by congressional district), or in a bar chart showing the top 10 awards by category. The visualizations show how spending has increased over time, the regional concentration of spending, and a list of the top recipients.

Figure 17: Example of a Map of Advanced Search Results for Social Security Retirement Insurance (as of September 2018)



Source: Search result on USA Spending.gov (<https://www.usaspending.gov/#/search>), accessed September 14, 2018. | GAO-19-72

We found that the Spending Explorer provides a simple, graphical interface that allows users to navigate spending data by budget function, agency, and object class.⁴² It gives users the option to drill down from these three high-level categories to specific program activities, federal accounts, recipients, or awards. It displays the total amount obligated for the selected category, and a breakdown of the amounts in dollars and as a percentage of the total.⁴³

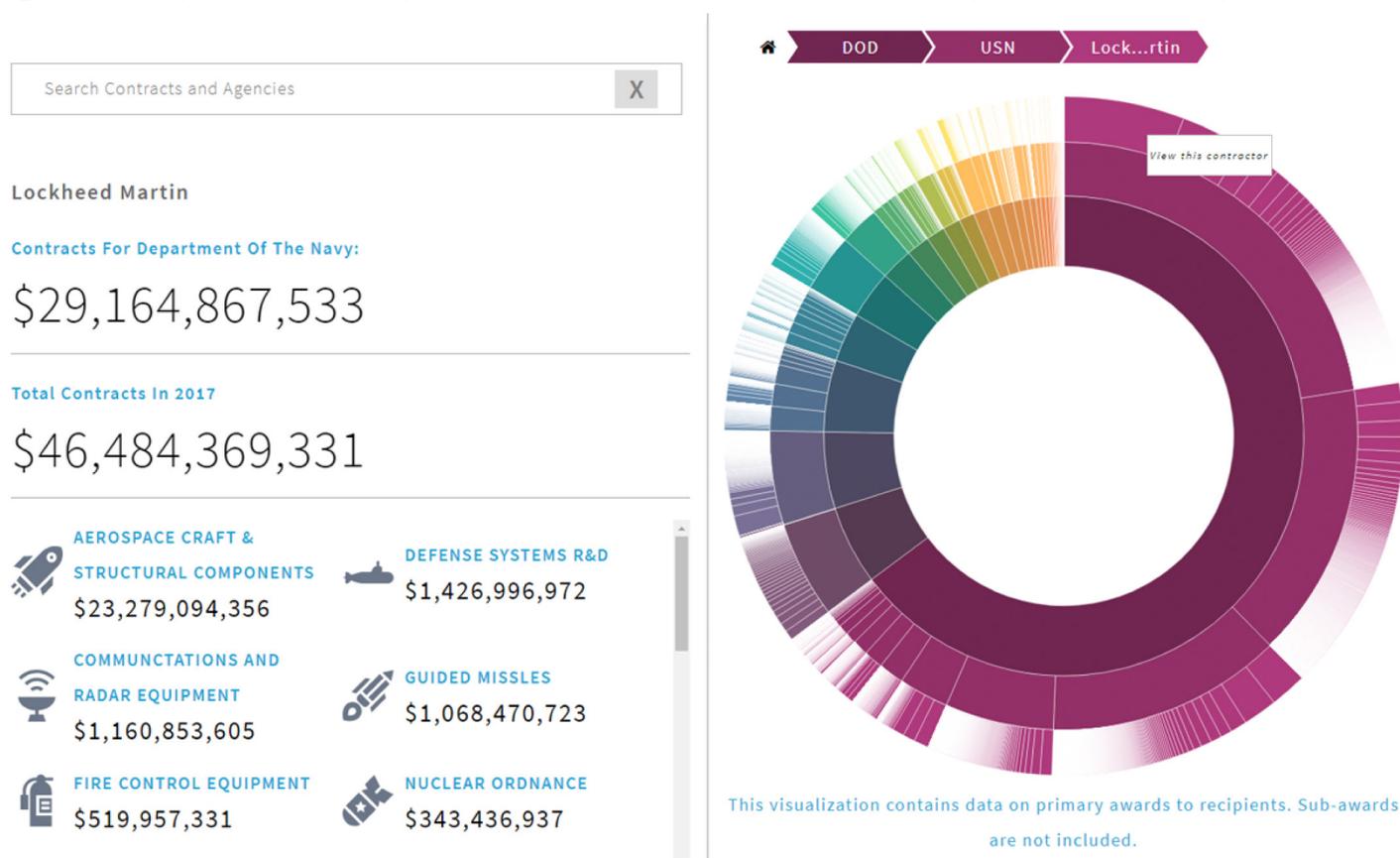
The Data Lab is a separate website linked to USA Spending.gov that offers users visual interpretations of the spending data. Treasury officials

⁴²The “budget function” classification system is a way of grouping budgetary resources so that all budget authority and outlays of on- and off-budget federal entities and tax expenditures can be presented according to the national needs being addressed, such as agriculture or national defense. The “object class” is a uniform classification identifying the obligations of the federal government by the types of goods or services purchased (such as personnel compensation, supplies and materials, and equipment) without regard to the agency involved or the purpose of the programs for which they are used.

⁴³An “obligation” is a definite commitment that creates a legal liability of the government for the payment of goods and services ordered or received, or a legal duty on the part of the United States that could mature into a legal liability by virtue of actions on the part of the other party beyond the control of the United States.

said the “Contract Explorer Sunburst” is a popular Data Lab visualization. As shown in figure 18, it provides users an interactive overview of about \$500 billion in federal contract data organized as a set of concentric circles starting from the funding agency (inner ring) to the recipients (outer ring). Treasury officials noted that analyses and visualizations in the Data Lab are updated with varying frequency because it can be a challenge to continually update some of the visualizations.

Figure 18: Example of a Contract Explorer “Sunburst” Visualization from the Data Lab (as of September 2018)



Source: Data Lab (<https://datalab.usaspending.gov/contract-explorer.html>), accessed September 14, 2018. | GAO-19-72

The website includes a glossary that provides plain language definitions of terms that describe the spending data. To help users understand the data on USASpending.gov, the website provides a “Glossary” sidebar that is available on every page of the website, and provides users both “plain language” and official definitions of financial terms that are used on the website, as shown in figure 19. According to

the key practices we identified, using commonly understood terms rather than technical jargon can help users understand the information provided.

Figure 19: Example of Plain Language (left) and Official Definition (right) on USAspending.gov

The figure displays two side-by-side screenshots from the USAspending.gov website. Each screenshot has a dark blue background with white text. The left screenshot shows the 'Plain Language' tab selected for the term 'Outlay'. The text reads: 'An outlay occurs when federal money is actually paid out, not just promised to be paid ("obligated").' Below the text is a '< BACK' link. The right screenshot shows the 'Official Definition' tab selected for the term 'Gross Outlay Amount'. The text reads: 'Payments made to liquidate an obligation (other than the repayment of debt principal or other disbursements that are "means of financing" transactions). Outlays generally are equal to cash disbursements but also are recorded for cash-equivalent transactions, such as the issuance of debentures to pay insurance claims, and in a few cases are recorded on an accrual basis such as interest on public issues of the public debt. Outlays are the measure of Government spending.' Below this text is '(defined in OMB Circular A-11)'. Both screenshots have a top navigation bar with 'Plain Language' and 'Official Definition' tabs.

Source: USAspending.gov glossary (<https://www.usaspending.gov>), accessed September 14, 2018 | GAO-19-72

A variety of search tools are available, but program source and city are not searchable. We found that the website features a variety of search tools to help users find and interpret the data. Users can search the data using generic keyword search and advanced search filters. These features allow users to explore and quickly obtain large volumes of award results. For example, we found that searching by funding agency returns all spending by that particular federal agency and by award.

However, we tested the search functionality of the website and found that two data elements required to be searchable by FFATA, as amended by the DATA Act, were not: (1) program source (Treasury Account Symbol

(TAS)) and (2) city.⁴⁴ Our search testing of a nongeneralizable, random sample of awards for data elements required by FFATA successfully found most of the data elements, but we were unable to search for program source (TAS) or city.⁴⁵ TAS and city data can be downloaded and are displayed on award web pages, but we were not able to search for them using either the advanced or keyword search pages. Treasury officials said they did not include functionality to search for these two required data elements on the new website because users searching by these data elements on the Beta version of the website had received confusing results. This is due in part to the fact that agency submissions with these data elements used different standards before and after the DATA Act. Instead, according to our review of the website, users can access TAS and city information using the website's navigation features, which officials said meets the spirit of the FFATA requirement.

However, simply displaying TAS or city information only on award or federal account pages does not meet the FFATA requirement that users be able to search for this information. Users currently have to click on a specific award or federal account page, and scroll through the web page to find the relevant section that shows city or TAS information. If federal agencies and Congress are not able to search for TAS, they cannot easily connect detailed information on financial transactions to federal accounts for management or oversight purposes. In addition, users looking for geographic information related to recipients or federal programs cannot easily search USAspending.gov by city.

⁴⁴TAS is a code assigned by Treasury to individual appropriation, receipt, or other fund accounts. All financial transactions of the federal government are classified by TAS for reporting to Treasury and OMB. TAS includes the account symbol and the period of availability for the funds it contains, among other things. For example, the Department of Education's Higher Education account includes TAS 091-2018/2018-0201-000. Code 091 is the Department of Education's agency code, 2018/2018 indicates that the funds are available for fiscal year 2018, 0201 is the account symbol for the Higher Education account, and 000 is the sub-account. According to OMB Memorandum M-09-19, TAS is considered the official program source for use in USAspending.gov.

⁴⁵The data elements required by FFATA that we tested include the following: Awardee/Recipient Legal Entity Name, Award Type, Funding Agency Name, North American Industry Classification System (NAICS) Code, Catalogue of Financial Domestic Assistance (CFDA) Number, Treasury Account Symbol, Award Description, Primary Place of Performance Address, Primary Place of Performance Congressional District, Legal Entity Address, Legal Entity Congressional District, Awardee/Recipient Legal Entity Identifier, Ultimate Parent Unique Identifier. FFATA, §§ 2(b)(1); 4(b)(3).

Government data catalogues and repositories link to

USAspending.gov. Treasury facilitates discovery of the DATA Act data by linking to USAspending.gov from centralized data repositories and catalogues. Information and links to USAspending.gov can be found on DATA.gov, which is a data catalogue for a variety of U.S. government datasets.⁴⁶

Treasury maintains a web page on GitHub, a public online collaboration website, designed to share information about its process in meeting the requirements of the DATA Act, including information and links to USAspending.gov.⁴⁷ Associated pages on this GitHub site serve as a data repository for the computer code behind the central data submission platform for the DATA Act, called the Data Act Broker, the API for USAspending.gov, and the USAspending.gov website itself.

Conclusions

USAspending.gov is a major open government data program with the potential to be a model for transparently reporting government data—if Treasury takes additional steps to further align it with the five key practices and associated key actions for open data in addition to DATA Act requirements. USAspending.gov has already followed several key actions such as providing the data on the website for free, engaging the public online and in person, providing detailed and disaggregated data for download, and making interactive tools so users can interpret and visualize the data. Treasury has also made progress in disclosing limitations of the data, although it has not fully addressed our prior recommendation to do more to make users fully aware of issues that affect its quality.

However, Treasury has not fully aligned USAspending.gov with some key practices or federal website standards, and has not fully implemented the search functionality required by FFATA, as amended by the DATA Act. As a result, users may not be able to find the information they need, and may not have confidence in the integrity of the data. Treasury updated USAspending.gov in October 2018 so that the web pages for the database download and agency submission files available at that time

⁴⁶According to OMB guidance, individual agencies are required to report their DATA Act data to DATA.gov.

⁴⁷Treasury's "fedspendingtransparency" GitHub page is located at <https://github.com/fedspendingtransparency>.

used the secure HTTPS protocol and a government domain. However, without an effective control process in place, Treasury does not have assurance that any new pages that may be added to USAspending.gov will conform to certain federal information security requirements. Furthermore, without easy access to structured metadata, it may be difficult or time consuming for users of USAspending.gov to find the information they need to determine whether or how to use the data. Similarly, the lack of an explicit open license might discourage some users from using the data to develop innovative commercial products. Users are also not able to easily search award information by program source or city, as required by FFATA, which could limit their ability to find and use these data to inform future decision making.

Recommendations for Executive Action

We are making a total of five recommendations to Treasury. Specifically:

The Secretary of the Treasury should establish a process to ensure all pages on the USAspending.gov website use the secure HTTPS protocol, consistent with OMB requirements. (Recommendation 1)

The Secretary of the Treasury should establish a process to ensure all content on USAspending.gov is available from a government domain, consistent with OMB requirements. (Recommendation 2)

The Secretary of the Treasury should fully comply with OMB's requirements by providing metadata in a single location that are easy to find on the USAspending.gov website. (Recommendation 3)

The Secretary of the Treasury should fully comply with OMB's requirements by communicating licensing information on USAspending.gov. (Recommendation 4)

The Secretary of the Treasury should ensure that users can easily search for awards by city and program source (TAS), consistent with FFATA requirements. (Recommendation 5)

Agency Comments, Third-Party Views, and Our Evaluation

We provided a draft of this report to the Secretary of the Treasury, the Director of OMB, and the Administrator of GSA for review and comment. Treasury provided written responses, which are summarized below and reproduced in appendix II. Treasury and OMB also provided technical comments, which we incorporated as appropriate. GSA responded that the agency had no comments on the report.

In its written response, Treasury highlighted areas where USAspending.gov aligned with the key practices that we identified for transparently reporting government data, such as engaging users and providing the data in useful formats. Treasury agreed with our recommendations. Treasury stated that the agency has already taken steps to address our first two recommendations, to establish processes to ensure that all pages on USAspending.gov use the secure HTTPS protocol and to ensure that all content on the website is available from a government domain. Treasury provided us with documentation of a revised process that is intended to address these issues. We revised the report to acknowledge that Treasury has taken these steps. We will continue to monitor Treasury's efforts to develop and implement this new process and update the status of our recommendations accordingly.

We also provided excerpts of the draft report to Connecticut; Kansas City, Missouri; Los Angeles, California; Montgomery County, Maryland; New York City, New York; and Ohio. Los Angeles, Montgomery County, New York City, and Ohio provided technical comments, which we incorporated as appropriate. Connecticut and Kansas City officials responded that they had no comments.

We are sending copies of this report to the Secretary of the Treasury, the Director of OMB, and the Administrator of the General Services Administration, as well as interested congressional committees and other interested parties. This report will be available at no charge on our website at <https://www.gao.gov>.

If you or your staff has any questions about this report, please contact Triana McNeil 202-512-6806 or McNeilT@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of our report. Key contributors to this report are listed in appendix III.

A handwritten signature in black ink, appearing to read "Triana McNeil". The signature is fluid and cursive, with a large initial "T" and "M".

Triana McNeil
Acting Director
Strategic Issues

List of Congressional Addressees

The Honorable Ron Johnson
Chairman
The Honorable Claire McCaskill
Ranking Member
Committee on Homeland Security and Governmental Affairs
United States Senate

The Honorable Rob Portman
Chairman
The Honorable Thomas R. Carper
Ranking Member
Permanent Subcommittee on Investigations
Committee on Homeland Security and Governmental Affairs
United States Senate

The Honorable Trey Gowdy
Chairman
The Honorable Elijah E. Cummings
Ranking Member
Committee on Oversight and Government Reform
House of Representatives

The Honorable Mark Meadows
Chairman
The Honorable Gerald E. Connolly
Ranking Member
Subcommittee on Government Operations
Committee on Oversight and Government Reform
House of Representatives

The Honorable Will Hurd
Chairman
The Honorable Robin Kelly
Ranking Member
Subcommittee on Information Technology
Committee on Oversight and Government Reform
House of Representatives

The Honorable Mark R. Warner
United States Senate

Appendix I: Objectives, Scope, and Methodology

The Digital Accountability and Transparency Act of 2014 (DATA Act) includes a provision for us to review implementation of the act. Over the last 4 years, we have issued 13 reports assessing various aspects of DATA Act implementation.¹ This report builds on our body of work on the DATA Act and (1) identifies key practices for transparently reporting government data on a centralized website, and (2) evaluates the extent to which the new USAspending.gov is consistent with these key practices, as well as existing standards for federal websites.

Identifying Key Practices

To identify key practices for transparently reporting open government data on a centralized website, we conducted a literature review and interviewed experts on open data and representatives of good governance groups. We also identified illustrative examples by interviewing open data practitioners from state and local governments.

Literature review. To conduct the literature review, we first identified relevant publications using a number of bibliographic databases, including ProQuest, the Organisation for Economic Co-Operation and Development's (OECD) iLibrary, the National Technical Information Service, and the Public Affairs Information Service. We reviewed articles that focused on open data programs and practices in OECD countries, including scholarly peer-reviewed articles, working papers, conference papers, and reports by policy research organizations, nonprofit organizations, and associations. We conducted our search in March 2017 and subsequently added relevant articles identified during our background research.² To systematically review these articles, one analyst reviewed each article to identify relevant themes. A second analyst then reviewed the documentation to verify categorization decisions. Then, both analysts met to resolve any discrepancies. We evaluated and synthesized the categorized information to identify commonly-reported key practices for transparently reporting open government data.

Interviews with experts. We selected open data and good governance experts based on recommendations made by other experts, frequent

¹A list of related products, including our prior work on the DATA Act and USAspending.gov, is included at the end of this report.

²We did not limit our review of relevant sources by date because many of the articles we reviewed are undated web publications.

citations in others' work, and recent contributions to the field. We also selected experts that represent a variety of sectors and backgrounds (such as government, academia, and nonprofit organizations). We obtained the views of the following individuals and organizations:

- Andrew Stott, former United Kingdom Director for Transparency and Digital Engagement,
- Center for Open Data Enterprise,
- Code for America,
- Data Coalition,
- Dr. Anneke Zuiderwijk-van Eijk, Delft University of Technology,
- General Services Administration (GSA),
- Global Initiative for Fiscal Transparency,
- Governance Laboratory of New York University,
- IBM Center for the Business of Government,
- International Budget Partnership,
- Johns Hopkins University Center for Government Excellence,
- Project on Government Oversight,
- ProPublica,
- Results for America,
- Sunlight Foundation,
- U.S. Public Interest Research Group,
- What Works Cities, and
- World Bank.

We first had open-ended conversations with experts to obtain their views on what key practices exist for transparently reporting open government data. After developing an initial list of key practices, we then conducted a second round of interviews with experts to finalize the list. We shared a draft of the key practices with the Department of the Treasury (Treasury), the Office of Management and Budget (OMB), and GSA.

Illustrative examples. To obtain illustrative examples showing how those key practices can be implemented, we selected open data practitioners from six selected state and local governments:

-
- Connecticut;
 - Kansas City, Missouri;
 - Los Angeles, California;
 - Montgomery County, Maryland;
 - New York City, New York; and
 - Ohio.

We selected these practitioners because they were identified in our literature search and by the experts we spoke with as having well-regarded open data websites. We also selected practitioners that have websites with both a general open data portal and visualizations showing budget or spending data. We also selected practitioners that represent different locations and levels of government, including cities, counties, and states. We reviewed these practitioners' open data websites and related documentation, and interviewed cognizant state and local government officials.

Evaluating USAspending.gov

To assess the extent to which USAspending.gov is consistent with the key practices and selected standards for federal websites, we reviewed the website, reviewed agency documents, observed Treasury's participation in a hackathon, and interviewed OMB staff and Treasury officials. Specifically, we analyzed the USAspending.gov website to determine how it aligned with the key practices and the extent to which data elements were searchable as required by the Federal Funding Accountability and Transparency Act of 2006 (FFATA). We also assessed USAspending.gov against criteria for federal websites and open data programs, including OMB M-17-06, *Policies for Federal Agency Public Websites and Digital Services*, and OMB M-13-13, *Open Data Policy—Managing Information as an Asset*.

To evaluate the extent to which the USAspending.gov search functionality complies with FFATA requirements, as amended by the DATA Act, we randomly selected a nongeneralizable sample of 30 awards (consisting of 15 contracts and 15 financial assistance awards) downloaded from USAspending.gov for fiscal year 2017. We identified the required FFATA data elements from these awards, searched for these elements on USAspending.gov in July 2018, recorded whether each search successfully resulted in a matching award, and observed any other issues that occurred during testing. Finally, we interviewed Treasury officials to corroborate our observations on search functionality and other aspects of

the website, and discussed any planned improvements to the website. We also interviewed GSA officials and OMB staff to clarify policies and procedures for federal websites.

We conducted this performance audit from February 2017 to December 2018 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Appendix II: Comments from the Department of the Treasury



DEPARTMENT OF THE TREASURY
WASHINGTON, D.C.

ASSISTANT SECRETARY

November 26, 2018

Triana McNeil
Acting Director, Strategic Issues
General Accountability Office
441 G Street, NW
Washington, DC 20548

Dear Ms. McNeil:

We appreciate the opportunity to review the Government Accountability Office (GAO) draft report GAO-19-72 (the Draft Report) addressing the alignment of USASpending.gov with key practices identified by GAO for transparently reporting government data. The Draft Report (1) identifies key practices for transparently reporting government data on a centralized website, and (2) evaluates the extent to which the new USASpending.gov is consistent with these key practices, as well as existing standards for federal websites and applicable laws.

The Department of the Treasury (Treasury) and the Office of Management and Budget (OMB) have been leading the implementation of the Digital Accountability and Transparency Act of 2014 (DATA Act) since its passage, with the goal of displaying more accessible, searchable, and reliable spending information for the purposes of promoting transparency, facilitating better decision-making, and improving operational efficiency. From the beginning, we established an open, transparent and interactive process to ensure that federal agencies and external stakeholders could actively participate in the development of the new USASpending.gov website following the agile development methodology and user-centered design principles. Following these principles, we launched the beta version of new USASpending.gov website in May 2017, which linked spending data from across the government for the first time. Treasury has continued to enhance the website based on user feedback since that time and recently released new download tools to allow users to explore federal spending by budget function or federal account.

We are pleased that USASpending.gov successfully meets many aspects of the five key practices identified by GAO for transparently reporting government data. For example:

1. *Provide Free and Unrestricted Data:* Treasury has made all of the data available to users at no cost in six formats including the agency submission files, awards archives, Application Program Interface (API), and download functionality by awards, account and the full USASpending.gov database.
2. *Engage with Users:* Treasury identifies data user profiles and evaluates their needs through user research, and also obtains and responds to user feedback. Treasury reaches

- out to potential users to encourage data use by conducting interviews with the general public. Treasury educates the public about the data available through various forums.
3. *Provide Data in Useful Formats:* Treasury provides data through data visualizations and multiple machine-readable formats including Comma-Separated Values (CSV) files, JavaScript Object Notation (JSON) defined by a specific JSON schema, and eXtensible Business Reporting Language (XBRL).
 4. *Fully Describe the Data:* The website discloses data sources and includes an interactive glossary that provides plain language definitions of terms that describe the spending data, data dictionaries, the Data Model (DATA Act Information Model Schema (DAIMS)), and publicly available code that explains the Broker validation rules. Moreover, the site is developed with open-source code which enables users to explore and reuse the underlying website code.
 5. *Facilitate Data Discovery for All Users:* The USAspending.gov website features a variety of search tools and interactive visualizations that enable and encourage user exploration and discovery. The Spending Explorer provides an easy-to-navigate feature for exploring spending by Budget Function/Sub-Function, Agency and Object Class. The Profiles pages (Agencies, Federal Accounts, States, and Recipients) also provide different ways to visualize the data. In addition, the Keyword Search provides users the ability to explore award data in broad strokes, and our Advanced Search offers filters that let users customize their search. The Data Lab site also facilitates data discovery through visualizations such as the Federal Account Explorer, Contract Explorer, and Contract Spending Analysis.

While Treasury is pleased with the success of USAspending.gov thus far, we recognize the need for continuous improvement. The Draft Report contains five recommendations for Treasury regarding improvements that could be made to better align USAspending.gov with the identified key practices, applicable guidance and data element search requirements. We accept GAO's recommendations, and note that we have already taken steps to address the issues noted in the first two recommendations and have provided GAO with documentation of our revised process.

We appreciate your helpful suggestions for improvements to the USAspending.gov website. We remain committed to working with GAO and the agencies to meet the DATA Act's requirements and objectives. Thank you again for the opportunity to review and comment on the Draft Report.

Sincerely,



David A. Lebryk
Fiscal Assistant Secretary

Appendix III: GAO Contact and Staff Acknowledgments

GAO Contact

Triana McNeil, (202) 512-6806 or mcneilt@gao.gov

Staff Acknowledgements

In addition to the contact named above, Thomas J. McCabe, Assistant Director, and Laurel Plume, Analyst-in-Charge, supervised the development of this report. Colenn Berracasa, Samuel Gaffigan, and Parke Nicholson made major contributions to this report. Also contributing to this report in their areas of expertise were Michael Bechetti, Steven Campbell, Mark Canter, Jenny Chanley, Jacqueline Chapin, Peter Del Toro, Nancy Donovan, Kathleen Drennan, Sarah Gilliland, Sarah Kaczmarek, Michael LaForge, Paula M. Rascona, Andrew J. Stephens, and James Sweetman, Jr.

Related GAO Products

DATA Act: Reported Quality of Agencies' Spending Data Reviewed by OIGs Varied Because of Government-wide and Agency Issues. [GAO-18-546](#). Washington, D.C.: July 23, 2018.

DATA Act: OMB, Treasury, and Agencies Need to Improve Completeness and Accuracy of Spending Data and Disclose Limitations. [GAO-18-138](#). Washington, D.C.: November 8, 2017.

Open Innovation: Executive Branch Developed Resources to Support Implementation, but Guidance Could Better Reflect Leading Practices. [GAO-17-507](#). Washington, D.C.: June 8, 2017.

DATA Act: As Reporting Deadline Nears, Challenges Remain That Will Affect Data Quality. [GAO-17-496](#). Washington, D.C.: April 28, 2017.

DATA Act: Office of Inspector General Reports Help Identify Agencies' Implementation Challenges. [GAO-17-460](#). Washington, D.C.: April 26, 2017.

DATA Act: Implementation Progresses but Challenges Remain. [GAO-17-282T](#). Washington, D.C.: December 8, 2016.

DATA Act: OMB and Treasury Have Issued Additional Guidance and Have Improved Pilot Design but Implementation Challenges Remain. [GAO-17-156](#). Washington, D.C.: December 8, 2016.

Open Innovation: Practices to Engage Citizens and Effectively Implement Federal Initiatives. [GAO-17-14](#). Washington, D.C.: October 13, 2016.

DATA Act: Initial Observations on Technical Implementation. [GAO-16-824R](#). Washington, D.C.: August 3, 2016.

DATA Act: Improvements Needed in Reviewing Agency Implementation Plans and Monitoring Progress. [GAO-16-698](#). Washington, D.C.: July 29, 2016.

DATA Act: Section 5 Pilot Design Issues Need to Be Addressed to Meet Goal of Reducing Recipient Reporting Burden. [GAO-16-438](#). Washington, D.C.: April 19, 2016.

DATA Act: Progress Made but Significant Challenges Must Be Addressed to Ensure Full and Effective Implementation. [GAO-16-556T](#). Washington, D.C.: April 19, 2016.

DATA Act: Data Standards Established, but More Complete and Timely Guidance Is Needed to Ensure Effective Implementation. [GAO-16-261](#). Washington, D.C.: January 29, 2016.

DATA Act: Progress Made in Initial Implementation but Challenges Must be Addressed as Efforts Proceed. [GAO-15-752T](#). Washington, D.C.: July 29, 2015.

Federal Data Transparency: Effective Implementation of the DATA Act Would Help Address Government-wide Management Challenges and Improve Oversight. [GAO-15-241T](#). Washington, D.C.: December 3, 2014.

Government Efficiency and Effectiveness: Inconsistent Definitions and Information Limit the Usefulness of Federal Program Inventories. [GAO-15-83](#). Washington, D.C.: October 31, 2014.

Data Transparency: Oversight Needed to Address Underreporting and Inconsistencies on Federal Award Website. [GAO-14-476](#). Washington, D.C.: June 30, 2014.

Federal Data Transparency: Opportunities Remain to Incorporate Lessons Learned as Availability of Spending Data Increases. [GAO-13-758](#). Washington, D.C.: September 12, 2013.

Government Transparency: Efforts to Improve Information on Federal Spending. [GAO-12-913T](#). Washington, D.C.: July 18, 2012.

Electronic Government: Implementation of the Federal Funding Accountability and Transparency Act of 2006. [GAO-10-365](#). Washington, D.C.: March 12, 2010.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (<https://www.gao.gov>). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to <https://www.gao.gov> and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, <https://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on [Facebook](#), [Flickr](#), [Twitter](#), and [YouTube](#).
Subscribe to our [RSS Feeds](#) or [E-mail Updates](#). Listen to our [Podcasts](#).
Visit GAO on the web at <https://www.gao.gov>.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Website: <https://www.gao.gov/fraudnet/fraudnet.htm>

Automated answering system: (800) 424-5454 or (202) 512-7700

Congressional Relations

Orice Williams Brown, Managing Director, WilliamsO@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800 U.S. Government Accountability Office, 441 G Street NW, Room 7149 Washington, DC 20548

Strategic Planning and External Liaison

James-Christian Blockwood, Managing Director, spel@gao.gov, (202) 512-4707 U.S. Government Accountability Office, 441 G Street NW, Room 7814, Washington, DC 20548



Please Print on Recycled Paper.