



OPEN

DATA

TRANSITION

REPORT



THE CENTER FOR
OPEN DATA ENTERPRISE



**An Action Plan for the
Next Administration**

OCTOBER 2016

OPEN DATA TRANSITION REPORT

An Action Plan for the
Next Administration



THE CENTER FOR
OPEN DATA ENTERPRISE

ABOUT THIS REPORT

The Center for Open Data Enterprise is a grantee of the Laura and John Arnold Foundation, which provided support to prepare this nonpartisan informational report. Drawing on numerous interviews, expert meetings, and other research, the report examines the current open government data ecosystem and makes recommendations to improve open data policies and programs moving forward.

ABOUT THE LAURA AND JOHN ARNOLD FOUNDATION (LJAF)

LJAF is a private foundation that is working to address our nation's most pressing and persistent challenges using evidence-based, multidisciplinary approaches. Its strategic investments are currently focused in sustainable public finance, criminal justice, education, evidence-based policy and innovation, research integrity, and science and technology.

LJAF has offices in Houston, New York City, and Washington, D.C. Learn more at ArnoldFoundation.org.

ABOUT THE CENTER FOR OPEN DATA ENTERPRISE

The Center for Open Data Enterprise is an independent nonprofit organization that develops smarter open data strategies for governments, businesses, and nonprofits by focusing on data users. Our mission is to maximize the value of open data as a public resource.

OPEN DATA TRANSITION REPORT TEAM

Project Lead: **Joel Gurin**
Project Manager: **Katherine Garcia**
Project Writer and Researcher: **Katie Frost**
Research Fellows: **Stephanie Huang, Matt Rumsey**
Project Advisors: **Audrey Ariss, Laura Manley**
Designer: **Zak Bickel**
Illustrator: **Chelsea Beck**

CONTACT US

For general inquiries, contact **Katherine Garcia** at katherine@odenterprise.org.
For partnership opportunities, contact **Laura Manley** at laura@odenterprise.org.

Learn more at OpenDataEnterprise.org.

ACKNOWLEDGEMENTS

The Center for Open Data Enterprise thanks the many open data champions and subject-matter experts whose guidance and input made this report possible.

ADVISORY COMMITTEE

Our advisory committee provided strategic direction and oversight, expert input, and content review from the project's initiation through the final draft.

- **Bryce Pippert**, Booz Allen Hamilton
- **Elizabeth Grossman**, Microsoft
- **Heather Joseph**, SPARC
- **Hudson Hollister**, Data Coalition
- **Mark Doms**, Nomura Securities
- **Michael Stebbins**, The Laura and John Arnold Foundation
- **Theresa A. Pardo**, Ph.D., Center for Technology in Government, University at Albany, State University of New York

WORKING GROUP

Working group members contributed subject-matter expertise in specific policy areas to ensure appropriate context for the recommendations in their domains.

- **Abhi Nemani**, City of Sacramento / EthosLabs
- **Amy Edwards**, U.S. Department of the Treasury
- **Ana Pinheiro Privette**, Climate Data Solutions
- **Bobby Jones**, U.S. Department of Agriculture
- **Dan Morgan**
- **Francine Berman**, Research Data Alliance
- **Jed Kolko**, Indeed
- **Joe Pringle**, Carto
- **Kathryn Pettit**, The Urban Institute
- **Matt Gee**, The Impact Lab and the University of Chicago
- **Rebecca Williams**, The Center for Government Excellence
- **Steve Young**, Innovate! Inc.

ACKNOWLEDGEMENTS

INTERVIEWEES AND REVIEWERS

Interviewees helped brainstorm priority topics and actionable opportunities that led to these recommendations. Reviewers helped refine recommendations through discussion and draft review.

- | | | |
|----------------------------|--------------------|-------------------|
| • Alan Marco | • Henry Kelly | • Miriam Nisbet |
| • Aneesh Chopra | • Hyon Kim | • Neil Thakur |
| • Anne Washington | • Jaime Adams | • Nick Sinai |
| • Archon Fung | • Jed Miller | • Peter Levin |
| • Avi Bender | • Jed Sundwall | • Philip Ashlock |
| • Bryan Sivak | • Jeff Kaplan | • Philip Bourne |
| • Cinthia Schuman Ottinger | • Jenn Gustetic | • Ren Essene |
| • Damon Davis | • Jeremy Roberts | • Robert Grossman |
| • David Portnoy | • John Wilbanks | • Robert Renner |
| • David Yanofsky | • Josh Green | • Sid Burgess |
| • Debbie Brodt-Giles | • Karen Lay-Brew | • Sokwoo Rhee |
| • Doug Laney | • Kevin Merritt | • Sophie Raseman |
| • Emily Shaw | • Kirtan Upadhyaya | • Tim Davies |
| • Eric Handler | • Lisa Abeyta | • Walt Wells |
| • Erie Meyer | • Lourdes German | • Walter Katz |
| • Ethan Gurwitz | • Marc DaCosta | • Will Saunders |
| • Frans Hietbrink | • Michael Pierce | |
| • Gavin Hayman | • Michael Pizzo | |

TABLE OF CONTENTS

1
Executive Summary

4
Introduction

5
DEFINING OPEN DATA

5
THE NATIONAL CONTEXT

6
REPORT ORGANIZATION

6
SUGGESTED OPTIONS
FOR FUNDING THE
RECOMMENDATIONS

7
Methodology

8

**GOAL I: ENHANCE THE
GOVERNMENT OPEN
DATA ECOSYSTEM**

9
RECOMMENDATIONS FOR THE FIRST 100 DAYS

18
RECOMMENDATIONS FOR THE FIRST YEAR

46

**GOAL III: SHARE
SCIENTIFIC
RESEARCH DATA
TO SPUR INNOVATION
AND SCIENTIFIC
DISCOVERY**

47
RECOMMENDATIONS FOR THE FIRST 100 DAYS

56
RECOMMENDATIONS FOR THE FIRST YEAR

26

**GOAL II: DELIVER
DIRECT BENEFITS
TO CITIZENS AND
COMMUNITIES**

27
RECOMMENDATIONS FOR THE FIRST 100 DAYS

34
RECOMMENDATIONS FOR THE FIRST YEAR

60

**GOAL IV: HELP
BUSINESSES AND
ENTREPRENEURS USE
GOVERNMENT DATA AS
A RESOURCE**

61
RECOMMENDATIONS FOR THE FIRST 100 DAYS

68
RECOMMENDATIONS FOR THE FIRST YEAR

EXECUTIVE SUMMARY



Open government data—data that is freely available for use, reuse, and republication—is a major public resource for government, citizens, the scientific research community, and the private sector. Citizens can use open data to improve democracy by interacting directly with the government and holding public officials accountable. The government can apply open data to deliver public services more effectively, foster entrepreneurship, broaden opportunities for businesses, and empower researchers to advance scientific discovery and drive innovation.

The Obama Administration has championed open data as an essential part of open, transparent government since the President's first day in office.¹ Over the past eight years, the federal government has launched a range of programs showing that open data is more than a tool for good government—it is a critical national resource. Federal open data programs are helping students and their parents choose colleges, improving health care, supporting local business, and literally mapping the universe. The White House has also made future commitments to use open data to improve government and to support international data-driven efforts.²

The next administration should build on these accomplishments by prioritizing four goals:

GOAL I: ENHANCE THE GOVERNMENT OPEN DATA ECOSYSTEM

Open data can help federal agencies accomplish their missions more effectively and efficiently. While some agencies have established strong open data programs, most need additional resources and support. The next administration should enhance the open data ecosystem by developing a strong data infrastructure across government, including appropriate personnel, policies, and coordination efforts.

GOAL II: DELIVER DIRECT BENEFITS TO CITIZENS AND COMMUNITIES

In the United States and internationally, open data has helped citizens and communities improve their health care, nutrition, education, public safety, and more. The next administration should identify the major challenges impacting American communities and leverage open data to address them.

GOAL III: SHARE SCIENTIFIC RESEARCH DATA TO SPUR INNOVATION AND SCIENTIFIC DISCOVERY

When scientists freely share their data, the entire research enterprise benefits. Working with the research community, the next administration should develop policy and technology solutions to make open, shared research data the norm.

GOAL IV: HELP BUSINESSES AND ENTREPRENEURS USE GOVERNMENT DATA AS A RESOURCE

American businesses are the heart of the economy and open data can fuel their growth. The next administration should help businesses by making it easier to access valuable government data and simpler to report data to regulatory agencies.

This report includes 27 actionable recommendations to support the next administration in pursuing these four goals. These recommendations are designed to be of value to the next President, his or her transition teams, and to government agencies and departments in the next administration. The administration can launch many of the recommendations in the first 100 days. All recommendations in this report will produce meaningful results in the administration's first year.

RECOMMENDATION	LEAD ACTOR	TIMELINE	PAGE
GOAL I: ENHANCE THE GOVERNMENT OPEN DATA ECOSYSTEM			
1. Define, fund, and appoint Chief Data Officers for each federal agency to establish data as a key element of each agency's mission.	OFFICE OF MANAGEMENT AND BUDGET	100 DAYS	Page 9
2. Ensure government data is born digital to help agencies analyze and open the data efficiently.	OFFICE OF MANAGEMENT AND BUDGET	100 DAYS	Page 12
3. Conduct a National Data Infrastructure Review to identify high-value opportunities for federal investments in tribal, city, and state data infrastructure.	COUNCIL ON FINANCIAL ASSISTANCE REFORM	100 DAYS	Page 16
4. Develop a centralized public database of FOIA requests and information released under FOIA, acting on the principle of "release to one, release to all."	OFFICE OF MANAGEMENT AND BUDGET & THE FOIA OFFICERS' COUNCIL	FIRST YEAR	Page 18
5. Fully implement the DATA Act to provide financial transparency and accountability through open data.	ALL FEDERAL AGENCIES	FIRST YEAR	Page 20
6. Standardize reporting data for federal grants to help make that data more accessible and useful.	U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES & OFFICE OF MANAGEMENT AND BUDGET	FIRST YEAR	Page 22
7. Enable data users to voluntarily provide direct feedback that will improve data quality.	THE WHITE HOUSE OFFICE OF SCIENCE AND TECHNOLOGY POLICY	FIRST YEAR	Page 24
GOAL II: DELIVER DIRECT BENEFITS TO CITIZENS AND COMMUNITIES			
8. Create a National Hunger Heat Map to help distribute food supplies to areas of greatest need and decrease hunger.	U.S. DEPARTMENT OF AGRICULTURE	100 DAYS	Page 27
9. Continue to support and expand local data-driven climate resilience projects to help communities prepare for severe weather incidents.	THE WHITE HOUSE	100 DAYS	Page 30
10. Launch a Police Data Initiative 2.0 to collect data on police violent encounters in a standardized, open format for transparency and research.	THE WHITE HOUSE AND FEDERAL BUREAU OF INVESTIGATION	100 DAYS	Page 32
11. Standardize and update the government's public data on occupations and required skills to help Americans find jobs.	U.S. DEPARTMENT OF LABOR & NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY	FIRST YEAR	Page 34
12. Help communities address opioid addiction by opening up data on drug treatment facilities.	U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES	FIRST YEAR	Page 36
13. Establish a National Data EnviroCorps in which volunteers collect air, water, and soil quality data and help communities identify and manage environmental risks.	U.S. ENVIRONMENTAL PROTECTION AGENCY	FIRST YEAR	Page 38

RECOMMENDATION	LEAD ACTOR	TIMELINE	PAGE
14. Develop guidelines for all consumer-facing regulatory agencies to make consumer complaints available as open data to improve products and services.	OFFICE OF INFORMATION AND REGULATORY AFFAIRS	FIRST YEAR	Page 40
15. Expand open data summer camps to inspire the next generation of data-savvy students and help improve and apply open government data.	U.S. DEPARTMENT OF AGRICULTURE	FIRST YEAR	Page 42
16. Open up data on housing choice voucher wait lists to support low-income families in finding housing.	U.S. DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT	FIRST YEAR	Page 44

GOAL III: SHARE SCIENTIFIC RESEARCH DATA TO SPUR INNOVATION AND SCIENTIFIC DISCOVERY

17. Establish a Federal Research Data Council to coordinate the government's commitment to open research.	NATIONAL SCIENCE AND TECHNOLOGY COUNCIL	100 DAYS	Page 47
18. Develop an Annual Research Data Census to increase awareness of federally funded research and improve access to data.	NATIONAL SCIENCE FOUNDATION	100 DAYS	Page 50
19. Coordinate agency efforts to transition to cloud storage and analytics, enabling more research with government data.	GENERAL SERVICES ADMINISTRATION	100 DAYS	Page 52
20. Work with major research organizations, scientific publications, and professional associations to move toward requiring researchers to publish their data in open, reusable formats.	THE WHITE HOUSE OFFICE OF SCIENCE AND TECHNOLOGY POLICY	100 DAYS	Page 54
21. Establish international standards for collaboration and data sharing, beginning with a pilot in Arctic research.	THE WHITE HOUSE	FIRST YEAR	Page 56
22. Identify and publish large, high-quality datasets across all fields for use in machine learning to support advances in artificial intelligence.	NATIONAL SCIENCE AND TECHNOLOGY COUNCIL & INTELLIGENCE ADVANCED RESEARCH PROJECTS ACTIVITY	FIRST YEAR	Page 58

GOAL IV: HELP BUSINESSES AND ENTREPRENEURS USE GOVERNMENT DATA AS A RESOURCE

23. Identify cost barriers to accessing government data and eliminate fees to level the playing field for data users.	U.S. GOVERNMENT ACCOUNTABILITY OFFICE	100 DAYS	Page 61
24. Launch a Standard Business Reporting Program that will ultimately help businesses lower reporting burdens and costs.	THE PRESIDENT, THE OFFICE OF INFORMATION AND REGULATORY AFFAIRS, & NATIONAL ECONOMIC COUNCIL	100 DAYS	Page 64
25. Partner with the automated vehicle industry to give open data a central role in creating national safety standards.	U.S. DEPARTMENT OF TRANSPORTATION	100 DAYS	Page 66
26. Provide complete, accurate, and timely data on energy use in buildings to help companies save money by increasing energy efficiency.	U.S. DEPARTMENT OF ENERGY	FIRST YEAR	Page 68
27. Make it easier to discover and access government owned intellectual property to help entrepreneurs build on this free resource.	U.S. PATENT AND TRADEMARK OFFICE	FIRST YEAR	Page 70

INTRODUCTION

The last several years have seen broad bipartisan support for opening government data to make it freely available for use, reuse, and republication. Open data about government spending and operations is essential to government transparency and accountability. Beyond that, government leaders in both major parties, as well as stakeholders outside of government, see open data's potential to fuel innovation across all sectors.

This *Open Data Transition Report* is designed to help institutionalize the federal commitment to open data and build on it through the next administration. The report integrates current institutional knowledge about open government data, and provides the new administration's transition teams with an action plan for continuity and further improvements.

In preparing this report, the Center for Open Data Enterprise developed a portfolio of recommendations that will help institutionalize open data in complementary ways. Some recommendations are designed to provide leadership and organizational support at the highest level to firmly establish or improve open data policy across government. Other, more operational recommendations focus on specific actions that individual agencies can take to leverage the value of open data in their own domains. Implementing the recommendations in this report will support broad government-wide advances and will help open data take root in individual federal agencies.



**Openness will strengthen
our democracy and
promote effectiveness
and efficiency in
government.**

- President Barack Obama



**By making this data
more readily accessible,
we take down barriers,
enhance transparency,
and put the people in
charge.**

- House Speaker Paul Ryan

DEFINING OPEN DATA

In May 2013, the Office of Management and Budget (OMB) issued *Open Data Policy—Managing Information as an Asset*, a memorandum emphasizing the importance of open data for the federal government.³ The memo defined open data as “publicly available data structured in a way that enables the data to be fully discoverable and usable by end users.” More specifically, the memo laid out seven principles, saying that open data must be:

1. **Public.** All agencies must “adopt a presumption in favor of openness” within all proper legal, privacy, and security restrictions.
2. **Accessible.** Open data must be convenient and modifiable, in machine-readable formats “that can be retrieved, downloaded, indexed, and searched.”
3. **Described.** The data must include a description of its strengths, weaknesses, analytical limitations, security requirements, and processing requirements.
4. **Reusable.** Truly open data must include an open license with no restrictions on its use.
5. **Complete.** Agencies must publish data in its primary form with as much granularity as possible.
6. **Timely.** Open data must be published as “quickly as necessary to preserve the value of the data.”
7. **Managed Post-Release.** Agencies must designate a point of contact to assist with data use, complaints, and requirements.

This report uses this definition of open data in framing its recommendations.

THE NATIONAL CONTEXT

The Obama Administration has made strong commitments to publishing and improving open government data as well as promoting its use. The administration developed a vision for what could be called Open Data 2.0—the use of open data as a public resource as well as a critical tool for government accountability and transparency.

The administration launched a number of programs that demonstrate both the government’s commitment to open data and open data’s potential to support economic growth and societal benefit.⁴ These were outlined in the *White House Fact Sheet: Data by the People, for the People—Eight Years of Progress Opening Government Data to Spur Innovation, Opportunity, & Economic Growth*. Highlights include:

- [Data.gov](#), designed to be “the home of all federal government data for the open data community” as well as other open government portals for specific agencies;
- [The College Scorecard](#), which helps students and families evaluate educational options;
- Programs using open data to tailor medical treatments to individuals, develop new cancer treatments, and improve health care;
- Initiatives to help local communities through better data on police activities, housing, schools, jobs, and transportation;
- Open science programs to share data on climate, the environment, and other areas of scientific research; and
- A range of data programs to make the federal government more efficient and effective.

Beyond individual programs and projects, the administration has made government-wide commitments to open data. Since the *Open Data Policy* was established in May 2013, the White House has helped guide federal work on open data through initiatives such as a Data Cabinet comprised of top federal data professionals, and an Open Source Policy that directs all federal agencies to make new custom-developed code freely and broadly available for reuse across the federal government. The White House also made advancing open data a Cross-Agency Priority Goal, one of the most important areas for government-wide commitment and activity.⁵

The *Third Open Government National Action Plan (NAP)*, published in October 2015, lays out a number of commitments through 2017.⁶ The NAP includes commitments to use open data for law enforcement, provide climate data, contribute to the Global Open Data on Agriculture and Nutrition program (GODAN), and support the Extractive Industries Transparency Initiative (EITI). A September 2016 update to the NAP added government contracting, foreign aid, and sustainable development commitments as well.⁷

The administration also developed new resources that agencies across government can use to solve technical challenges, including

challenges related to open data. The U.S. Digital Service, a startup at the White House, brings in experts from leading technology firms to build new websites, portals, and tools for a wide range of federal departments. 18F, an office inside the Technology Transformation Service at the General Services Administration (GSA), works with teams in federal agencies to build, buy, and share new digital resources. The Presidential Innovation Fellows program, administered by the White House and GSA, attracts technologists and innovators for one-year stints with the federal government to work on initiatives involving technology and policy. The Office of the U.S. Chief Technology Officer, a position established during President Obama's first term, provides high-level leadership for federal open data programs, together with GSA and OMB.

The current administration has done a great deal to establish open data as a valuable national resource, but these successes are not fully institutionalized. The Open, Public, Electronic and Necessary (OPEN) Government Data Act, now making its way through Congress, would give much of the current Open Data Policy[†] the force of law.⁸ At this writing, however, the OPEN Act's passage is not certain. Moreover, the scope of open government data is so broad that it will need additional organizational and political support to fully realize its potential. There is a critical need to reaffirm and build upon the government's commitment to open data with sustainable infrastructure, policies, practices, and resources that are not dependent on any one administration.

REPORT ORGANIZATION

This report offers 27 recommendations to expand on open data programs successes and advance open data policy.

The report's recommendations are organized around four priority open data goals for the next administration, one for each of the key beneficiaries of open data policy: government, citizens and communities, research, and businesses.

- **Goal I:** Enhance the government open data ecosystem
- **Goal II:** Deliver direct benefits to citizens and communities
- **Goal III:** Share scientific research data to spur innovation

and scientific discovery

- **Goal IV:** Help businesses and entrepreneurs use government data as a resource

Under each goal, the report highlights several actionable recommendations for the new administration's first 100 days and first year. Recommendations include action plans that identify the lead actor and a process for moving forward, as well as a detailed description and rationale.

SUGGESTED OPTIONS FOR FUNDING THE RECOMMENDATIONS

Although some recommendations include specific funding mechanisms, most focus on the actions required to achieve the goal, rather than on the support instruments. The pending Modernizing Government Technology Act would provide funding that agencies could leverage to implement these recommendations. At this time, the Act's passage is uncertain. Regardless of its success, Congress should authorize some form of government-wide investment in modernizing information technology to update legacy systems and take advantage of modern processes and technologies.

Many recommendations are designed to result in net savings from efficiencies and improved resource allocation. For example, by ensuring that government data is born digital (Recommendation #2), the government can eliminate costly processes that are now required to manually re-format data for analysis or public release. Savings from these recommendations can be repurposed to support other recommendations that require direct investments for long-term improvements.

[†] Open Data Policy refers to the Office of Management and Budget Memorandum M-13-13

METHODOLOGY

The Center for Open Data Enterprise designed the methodology for this report to ensure a portfolio of strong recommendations that are:

1. **Feasible**, that is, possible to accomplish given the current state of affairs and relevant constraints;
2. **Actionable**, meaning it is clear how to put the recommendation into operation; and
3. **High-Impact**, as demonstrated through the success of a previous similar program, or otherwise likely to have a large, positive effect.

With these criteria in mind, the Center for Open Data Enterprise approached this report in four overlapping stages over a five month period:

1. Initial Research and Framing
 - Met with the project's Advisory Committee and identified broad policy areas and four high-level goals for additional research and investigation
2. Information Gathering
 - Drew on findings from a series of Open Data Roundtables, led by the Center for Open Data Enterprise, that brought together government agencies and key stakeholders

- Conducted desk research to review over 340 publicly available sources, including reports, articles, and research papers on current open data programs and opportunities for future initiatives
- Used snowball sampling to identify, interview, and consult with 57 experts across a spectrum of policy areas working in academia, government, the private sector, and nonprofit organizations
- Held a public event to gather input from the open data community and federal leadership

3. Content Development

- Analyzed research findings to create 78 draft recommendations
- Worked with the Advisory Committee, Working Group, and Interviewees to choose and shape draft recommendations

4. Review and Refinement

- Used an iterative process to review and refine draft recommendations, including peer review of each recommendation by one to four subject-matter experts
- Worked with the Advisory Committee and interviewees to select the 27 final recommendations, based on the above criteria and a holistic assessment of the entire portfolio

OVER
340
PUBLICLY AVAILABLE REPORTS,
ARTICLES, AND RESEARCH PAPERS

57
interviewees
and reviewers



GOAL I: ENHANCE THE GOVERNMENT OPEN DATA ECOSYSTEM

While the federal government has made significant strides in open data policy, open data also relies on a data-driven culture and strong organizational support and leadership. The government should continue to improve the full ecosystem to bolster greater consistency in open data projects across government and establish the foundation necessary for long-term success.

RECOMMENDATION 1:

Define, fund, and appoint Chief Data Officers for each federal agency to establish data as a key element of each agency's mission.

FIRST 100 DAYS


OFFICE OF MANAGEMENT AND BUDGET

ACTION PLAN:


- Within the first 100 days of the new administration, the Office of Management and Budget (OMB), in consultation with the Director of the White House Office of Science and Technology Policy (OSTP) and the U.S. Chief Technology Officer, should provide guidance encouraging all agencies to establish a Chief Data Officer (CDO) and include best practices guidance on the CDO role and authorities. The CDO should be empowered to work across his or her agency, to ensure the organization is treating data as a strategic asset, including identifying where open data can help contribute to each mission priority area.⁹ In each agency:
 - The CDO should serve as senior advisor to the Chief Information Officer (CIO) and the department head. The CIO should delegate his or her data and statistical oversight authorities[†] to the CDO.¹⁰
 - The CDO's office should have a line item in the agency's budget, either as part of the CIO budget or as another senior advisory office.
 - The CDO should direct, manage, and provide policy guidance and oversight of agency data personnel, activities, and operations, including: development of



[†] Authorities defined under 44 U.S. Code § 3506 (e)(3)–(e)(9).



agency data management budgets; recruitment, selection, and training of personnel to carry out agency data management functions; and approval and management of agency data management systems.

- The CDO should promote the use of agency data, including agency-sponsored research data, by:
 - Working to ensure data interoperability and develop policies requiring data sharing, both within the agency through agency grants and contracts;
 - Educating staff about the data they create and the role it can play in supporting their mission;
 - Working with external partners to gauge demand for government data and support industry needs, as appropriate; and
 - Integrating data streams into decision trees to institutionalize data-driven decision making.¹¹
 - OSTP should invite all agency CDOs to meet periodically to coordinate federal open data initiatives.
- 

The federal government lacks strong, consistent resources across departments and agencies to support leadership in implementing open data programs.

In recent years, the Chief Data Officer (CDO) has emerged as a key leadership role for open data programs at all levels of government. The state of Colorado hired the nation's first government Chief Data Officer in 2009,¹² and many federal agencies have since followed that model. Private sector companies are applying the same idea. Harvard Business Review began recommending that companies hire a CDO in 2012.¹³

Despite growing recognition of the CDO's role, only some federal agencies now have a CDO (or a Chief Data Scientist, a different title that often has similar goals). In August 2016, Project Open Data, the government's centralized resource, listed only eight CDOs out of the 24 departments and agencies^{†,14} subject to the Chief Financial Officers Act of 1990.¹⁵ Some agencies that have a CDO do not empower him or her to serve as an effective senior advisor. In many of those government agencies, the CDO's responsibilities orient toward traditional information technology (IT) more than data strategy. Other agencies give the CDO broad responsibilities in principle but do not provide the budget and staff support to help him or her succeed.¹⁶

Across the federal government, the role, authority, and responsibilities for CDOs are inconsistent. Responsibilities range from taking care of backend software and infrastructure, to resolving technical challenges, to developing the next great digital service. CDOs can and should play a role that complements the Chief Information Officers (CIOs) and Chief Technology Officers (CTOs) within agencies. In general, CTOs focus on innovation and exposing the organization to new ideas, technologies, and people, while CIOs build citizen-focused digital services, move data to the cloud, and protect against cyber threats.¹⁷ A CDO should focus on the quality, collection, management, publication, and use of data assets, including agency-sponsored data (e.g., research data).

Agencies with strong CDOs have successfully leveraged their data to solve complex challenges, engaged their data communities, and

adopted a culture of data-driven decision making in their agencies. For example, the CDO at the Consumer Financial Protection Bureau (CFPB) oversees the Bureau's governance, acquisition, documentation, storage, analysis, and distribution of data, and has focused on centralizing and standardizing data. The CDO at the Department of Transportation made it the first agency to proactively publish its enterprise data inventory, a catalog of all of its datasets, both public and nonpublic, before the government required federal agencies to create and maintain data inventories.

Many agencies will need to balance the goal of making data open with the need to keep some kinds of data highly secure. The U.S. Department of Energy, for example, manages a wide range of data sources: from the datasets of the Energy Information Administration, which releases data of high public value, to data on nuclear weapons that could compromise national security if it were released. In these agencies, the CDO and his or her office will need to have perspective and expertise in both open data and cybersecurity, two areas with extremely different goals.

The White House Office of Science and Technology Policy should invite all agency CDOs to meet periodically to advise each other and coordinate their agencies on such matters as consolidation and modernization of data systems, aligning metadata standards, and internal controls.

Open government data needs centralized, expert leadership at the agency level. A cadre of agency CDOs with well-defined responsibilities, authority, and resources can:

- Ensure that data plays a more central role in executing the mission of government agencies;
- Support high-priority projects across government with relevant data, supporting evidence-based policymaking;
- Empower agencies to use their own data to improve their programs;
- Provide citizens and industry with resources funded by their tax dollars;
- Improve data governance and ensure adherence to best practices; and
- Realize economies of scale related to data management, creating efficiencies and helping to ensure consistency.

† The Chief Financial Officers Act of 1990 established the role of Chief Financial Officers (CFOs) across 24 departments and agencies within the federal government, now often called the CFO Act Agencies. They include: (1) The Department of Agriculture (2) The Department of Commerce (3) The Department of Defense (4) The Department of Education (5) The Department of Energy (6) The Department of Health and Human Services (7) The Department of Homeland Security (8) The Department of Housing and Urban Development (9) The Department of the Interior (10) The Department of Justice (11) The Department of Labor (12) The Department of State (13) The Department of Transportation (14) The Department of the Treasury (15) The Department of Veterans Affairs (16) The Environmental Protection Agency (17) The National Aeronautics and Space Administration (18) The Agency for International Development (19) The General Services Administration (20) The National Science Foundation (21) The Nuclear Regulatory Commission (22) The Office of Personnel Management (23) The Small Business Administration (24) The Social Security Administration.

RECOMMENDATION 2:

Ensure government data is born digital to help agencies analyze and open the data efficiently.

FIRST 100 DAYS

OFFICE OF MANAGEMENT AND BUDGET

ACTION PLAN:

- The Office of Management and Budget (OMB) should update implementation guidance for the Government Paperwork Elimination Act to support electronic filing (e-filing) within the administration's first 100 days. OMB should incentivize end-users to use e-filing for all forms for which an agency anticipates more than 50,000 annual submissions.
 - All federal agencies should take steps to require or nudge end-users to use e-filing options, such as making e-filing the default, putting several forms together online to provide a single portal, or adding disincentives for paper filing.
 - The Office of Information and Regulatory Affairs should ensure that agencies digitally collect and release regulatory datasets of high interest.
 - The President's budget should include requirements for digital data collection and publication, where appropriate, and should provide funds to support the transition to a digital approach.
-

For almost two decades, statutory, judicial, and regulatory actions have required or advised federal agencies to collect information electronically whenever feasible. The 1998 Government Paperwork Elimination Act required federal agencies to begin accepting electronic information by 2003, and further stipulated “if [an] agency anticipates receiving 50,000 or more electronic submissions of a particular form, multiple methods of submitting such forms electronically must be in place.”¹⁸ More recently, the 2013 Open Data Policy stipulated that agencies should default to electronic filing (e-filing) and that they “must use machine-readable and open formats for information as it is collected or created.”¹⁹ By ensuring that data is born digital, e-filing can greatly increase the accuracy and accessibility of government data.

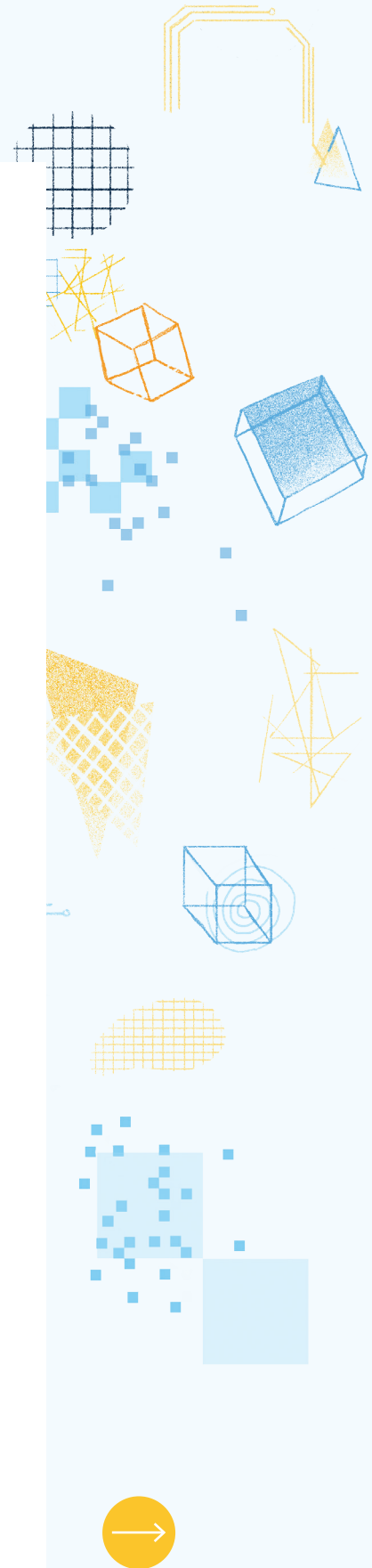
Despite these mandates, the federal government's electronic data collection includes significant gaps. Even when an agency offers the option to submit forms electronically, many users still elect to provide information on paper, especially if the electronic option is cumbersome. For example, although the Internal Revenue Service (IRS) met its long-standing goal of achieving an 80 percent electronic rate²⁰ for individual tax returns,²¹ only 36 percent of users e-filed their employment tax forms in the Form 94x series for 2015.[†] Employers looking to use the electronic version of the 94x series must clear several hurdles, including waiting up to 45 days to receive an electronic filing number and using specialty software.²² The Electronic Tax Administration Advisory Committee has said the IRS is unlikely to meet its 80 percent goal for this series until it implements a new electronic signature process for these forms.²³

When data is not born digital, it becomes

costly and time consuming to transform it into a structured dataset that can be accessed and reused. Accordingly, many key datasets are never digitized. During the second quarter of 2015, 13 percent of open government data downloads were in PDF format. The Small Business Administration and the Department of State had the highest rates of PDF datasets, at 87 percent and 58 percent, respectively.²⁴

The Office of Management and Budget (OMB) should update implementation guidance for the Government Paperwork Elimination Act to incentivize end-users to use e-filing for all forms for which an agency anticipates more than 50,000 annual submissions. Agencies have several tools at their disposal to increase the likelihood of end-users choosing e-filing. Making the e-filing option the default, and presenting it in a way that is easy to use, can drive up its use. The Free Application for Federal Student Aid, for example, has made online submissions the default and now receives more than 95 percent of submissions online.²⁵ Putting several required forms together online can provide a “one-stop-shop” that makes it easy to use electronic filing for all of them. At the same time, disincentives for using paper forms can drive users to e-filing: The Electronic Filing Incentive for the U.S. Patent and Trademark Office, which charges between \$200–\$400 for paper applications, has produced a 93 percent e-filing rate.²⁶

The Office of Information and Regulatory Affairs should ensure that electronic filing facilitates rapid, accurate open data in areas of high public interest. Recent rulemakings have shown the importance of changing from paper forms to making information electronically available in key regulatory areas. Over the past several years, the Federal Communications Commission has required television and radio stations to electronically upload their records of political ad buys online. This replaced the previous practice of keeping paper records, which made it impossible to



† The Form 94x series includes Form 940, Employer's Annual Federal Unemployment (FUTA) Tax Return; Form 941, Employer's Quarterly Federal Tax Return; Form 943, Employer's Annual Federal Tax Return for Agricultural Employees; Form 944, Employer's Annual Federal Tax Return; and Form 945, Annual Return of Withheld Federal Income Tax and related schedules.

survey broad patterns of political ad spending.²⁷ In May 2016, the Occupational Safety and Health Administration similarly ruled that large employers will have to file reports on injuries and illnesses electronically so that the data can easily be made public.²⁸

E-filing allows for electronic data protection and pre-submission validation (e.g., a phone number field can be set to only allow the entry of ten numerals) rather than the unrestricted way information is collected on paper. It also eliminates the substantial risk of introducing inaccuracies when moving data from paper-based to electronic forms.

Ensuring government data is born digital makes data collection faster, reduces errors in the data, and enables agencies to open the data more easily and rapidly. More specifically, uniform electronic filing will save taxpayers money and improve data accuracy. Many agencies have had to print, copy, and store paper forms, incurring costs that can be eliminated by switching to electronic data collection. For example, the U.S. Patent and Trademark Office's Electronic Filing System has saved an estimated \$5,243,440 in paper costs alone in twelve years of electronic data submission.²⁹

The President's budget should specify e-filing where it is appropriate and can make a significant difference, both to realize the potential savings from e-filing and to ensure that funds are available to make the transition from paper to electronic formats. For example, President Obama's 2017 budget requests that all tax-exempt organizations required to file Form 990 series returns must do so electronically.³⁰ That proposal, which would also require the IRS to release Form 990 data electronically, would let organizations apply for a waiver to submit their forms on paper and would allow a transition of up to three years to comply with the new rule. The President's budget should continue to include this initiative and use similar language to support e-filing across the federal government.

Overall, e-filing makes it more efficient for businesses to submit data and for agencies to publish it, saving time and labor costs. Ultimately, adopting e-filing across government will enable businesses to input key information only once and have it applied across all the forms they are required to submit to federal agencies. This approach has the potential to greatly improve efficiency and reduce the regulatory burden on business.

**MAKING SURE
DATA IS
BORN DIGITAL
WILL HELP
GOVERNMENT
COLLECT AND
PUBLISH DATA
MORE EASILY,
QUICKLY, AND
ACCURATELY.**

RECOMMENDATION 3:

Conduct a National Data Infrastructure Review to identify high-value opportunities for federal investments in tribal, city, and state data infrastructure.

FIRST 100 DAYS
COUNCIL ON FINANCIAL ASSISTANCE REFORM

ACTION PLAN:

- The Council on Financial Assistance Reform (COFAR) should conduct a 100-day National Data Infrastructure Review to:
 - Review existing federal grant programs (including components of broader grant programs) that explicitly support state, local, and tribal government efforts to collect, manage, and share data. The review should:
 - Develop and publish a resource list of these grants as a new category on [grants.gov](https://www.grants.gov). While specific departments within state, local, and tribal governments may be aware of the grants that are relevant to them, a comprehensive way to search for these opportunities would be valuable for mayors, governors, and tribal leaders planning multi-year open data initiatives, and to federal agencies considering new grantmaking opportunities.
 - Develop a plan to assess the effectiveness of these grants as vehicles for improving data infrastructure and the conditions for success. The assessment should include the benefits to the granting agencies (e.g., the ability to collect and aggregate better local data) as well as to the grant recipients.
 - Commission a study to identify any gaps in the current grant structure for data initiatives and recommend specific opportunities for other similar grant programs, or opportunities for cross-functional data infrastructure grants that provide comprehensive solutions. The study should recommend target levels and sources of federal funding for local data programs in the context of other infrastructure investments.
-

Many of the most promising opportunities for applying government open data reside at the state, local, and tribal level. Local leaders understand their constituents' needs and can use public data knowledgeably to address their concerns. Despite the potential benefits, however, too few state, local, and tribal governments have the data resources to fuel these kinds of innovative projects. The lack of data infrastructure is particularly pronounced outside of major urban areas.

A 2014–2015 survey by *Governing* magazine demonstrated the data challenges at the state level. *Governing* interviewed over 75 officials in 46 states between October 2014 and March 2015 to assess their view of the quality of government datasets. Every official they interviewed said that he or she encountered data quality issues, with 54 percent reporting facing such challenges frequently.³¹ These data quality issues are mirrored at the local and tribal level and reflect operational, legal, and technological barriers that will take dedicated investment for governments to overcome.

Several government programs are working with cities, tribes, and states to help build their technology capabilities and data resources. For example, the White House's Smart Cities Initiative has invested \$160 million in local governments around the country.³² Additionally, many of the various federal grant programs have data components. For example, Department of Health and Human Services grants provide funding for health data systems, Department of Justice grants provide funding for police data systems, and Department of Education grants fund the collection and publication of performance data.

However, these efforts lack cross-agency coordination. The result is that officials at all levels of government lack a clear vision of the data resources available. Furthermore, city, state, and tribal leaders who want to establish a comprehensive data management plan have agency-specific funding lines that make it difficult to

implement data management best practices. The current grant structure makes developing a multi-year data management strategy more complex, which contributes to broader problems such as siloed databases and a lack of interoperability.

The Council on Financial Assistance Reform should conduct a 100-day National Data Infrastructure Review to identify these unique data grant programs, publish a comprehensive list, and make them easily searchable on [grants.gov](https://www.grants.gov). This list will enable governors, mayors, and tribal leaders to better plan for data infrastructure development and management. It will support local leaders looking to move toward greater data standardization across their jurisdiction. In addition, it will provide models for data stewards across the federal government considering new grant programs to support local data.

State, local, and tribal government data should be a critical consideration in any national effort to improve data infrastructure, for two reasons. First, data collection and management systems themselves should be considered a key part of the infrastructure that modern cities need to run efficiently and effectively. Second, accurate, usable local data is essential to guide other infrastructure investments. Data on traffic flows and commuting patterns, for example, is critical to allocating investments in public transportation; planners must consider data on flood risk when designing construction in flood-prone areas; and engineers need data on energy usage patterns to guide decisions on power generation and energy supply.

Finally, accurate state, local, and tribal data can support evidence-based policymaking by providing information on the impact of different government programs. Better data at this level could strengthen the design and implementation of federal programs across all agencies.

RECOMMENDATION 4:

Develop a centralized public database of FOIA requests and information released under FOIA, acting on the principle of “release to one, release to all.”

FIRST YEAR**OFFICE OF MANAGEMENT AND BUDGET & THE FOIA OFFICERS’ COUNCIL**

ACTION PLAN:

- The Office of Management and Budget (OMB), in consultation with the Freedom of Information Act (FOIA) Officers’ Council, should ensure the development of a single database of FOIA requests with standard data fields that is public, searchable, and sortable for all agencies across the federal government. The plan for this database should be complete by January 2018, with full implementation by December 2018.
 - Entries in this database should connect to online versions of information released in response to those requests, fulfilling the spirit of “release to one, release to all.”
 - OMB should provide all information made available through this database in machine-readable formats.
-

The movement toward open data in the United States began with the Freedom of Information Act (FOIA), first passed 50 years ago to give the public the right to request records from any federal agency. FOIA requires federal agencies to disclose any information requested unless it falls under one of nine exemptions that protect interests such as personal privacy, national security, and law enforcement. But while FOIA ensures a high degree of government transparency in principle, it does not always do so in practice.

The federal government now receives more than 700,000 FOIA requests a year, which are processed by offices throughout federal agencies.³³ FOIA policies and practices have evolved through court decisions and changes in the law and technology over the years. However, there is no real central coordination or standardization for the agencies' FOIA operations. Thus, some agencies release data in PDF, Word documents, or other formats that are not helpful for mining the data or analyzing patterns in FOIA requests. Moreover, many agency FOIA offices have trouble keeping up with demand and may have delays of half a year or more in responding to FOIA requests, as a 2015 study showed.³⁴

The FOIA Improvement Act, which became law in June 2016, is one of the most significant improvements to FOIA since the original law was passed in 1966. Among other changes, the Act requires setting up a single electronic portal for all FOIA requests and establishes a FOIA Officers Council for interagency coordination.³⁵

The Office of Management and Budget (OMB) should build on the FOIA Improvement Act with additional provisions that would rapidly make FOIA a more effective vehicle for opening federal data. To begin, while the FOIA Improvement Act requires agencies to release information electronically, it does not specify how that should be done. OMB should define the "electronic format" required by the Act to refer to a machine-readable format. Specifying release in machine-readable formats will make the information much more useful and bring FOIA in sync with the federal Open Data Policy.³⁶

Beyond that, OMB, in consultation with the FOIA Officers' Council, should ensure the development of a searchable, sortable public database of all FOIA requests—not just the portal for inputting requests that the Act requires. While the law requires agencies to keep records of the requests they receive and how they handle them, they have not been required to publish logs of these requests in an easily usable form. FOIA.gov publishes information on the

number of FOIA requests received by each agency, but not the actual requests themselves.

A single unified database would make it possible for anyone to find out what information has been requested and what has been released under FOIA on topics of interest. The database can be developed using the request portal required by the FOIA Improvement Act as a starting point, with participation of the FOIA Officers' Council to agree on format, data fields, and other details. A possible model is FOIAonline.regulations.gov, which is a website for the public to submit FOIA requests, track their progress, and search information previously made available.³⁷ FOIA Online currently publishes this information for only 12 agencies and offices, but could help provide a basis for a larger effort.³⁸

FOIA should operate on the principle of "release to one, release to all": If an office releases information to one person or organization that requests it, that information should be made available to everyone. The FOIA Improvement Act requires electronic publication of any information that has been requested at least three times. Rather than waiting for multiple requests, agencies should simply use electronic publication proactively for all the information they release under FOIA, making it available to the public without waiting for multiple requests. To provide a framework for agencies to use, OMB should structure the database of FOIA requests so that the requests in the database link to online versions of the actual information released in response. With this structure, anyone who finds a request of interest in the database would be able to access the information that was released with a single click.

Creating a standardized and automated FOIA database will reduce repeated requests, since anyone considering a FOIA request will first be able to see whether the information he or she is looking for has already been released. The database will also make it possible to search by topic, not by agency, reducing the chance that a request will go to the wrong agency to handle it. These improvements will enable agencies to manage FOIA requests more easily. Agencies will also benefit from the ability to analyze interest in FOIA request data by measuring traffic to different parts of the FOIA database, giving them a better understanding of what datasets are high priority to the public.³⁹

Overall, these changes will help to democratize FOIA and the information available through it. A disproportionate number of FOIA requests now come from commercial entities⁴⁰ and from hedge funds,⁴¹ which are able to afford the time and expense required to submit large numbers of requests for business-related information. A FOIA database and electronic publication will make this information available to all, not just those with extensive resources.

RECOMMENDATION 5:

Fully implement the DATA Act to provide financial transparency and accountability through open data.

FIRST YEAR
ALL FEDERAL AGENCIES

ACTION PLAN:

- All federal agencies should continue preparations to report standardized spending data using the Digital Accountability and Transparency Act of 2014 (DATA Act) Information Model Schema. The Department of the Treasury and 18F should support agency initiatives to help ensure they meet the Act's May 2017 deadline to begin this reporting.
 - The Office of Management and Budget (OMB) should continue building on the DATA Act's momentum by announcing a plan to eliminate duplicative legacy reporting systems, such as Treasury systems, the Federal Assistance Award Data System, and the Federal Procurement Data System, once DATA Act reporting is underway.
 - OMB should provide guidance for all federal agency Chief Financial Officers to use the DATA Act data to analyze the performance of their programs.
 - The President should direct OMB and the federal agencies to use the DATA Act Information Model Schema to create, report, and publish all spending-related information, including payment requests. To ensure checkbook-level spending information is captured, the administration should develop its budget using a format consistent with the Schema; and the President should encourage Congress to do the same with appropriations bills.
-

The DATA Act launched the federal government on a path to standardize and publish all federal spending data. Efforts to implement the law are well underway. In April 2016,

the Department of the Treasury issued the DATA Act Information Model Schema, which is the standard format that all agencies will use to report spending information by May 2017.⁴² Treasury will soon display all spending information as open data on a revamped version of the website USASpending.gov. The Department has already launched a beta version of the updated site and is gathering feedback on improvements ahead of the May 2017 deadline.⁴³ Federal agencies should be actively planning now to be prepared for the switch.

The DATA Act was a monumental step forward for open government data, but passing the legislation itself did not ensure the government will accomplish its goals. Realizing the widespread benefits of the DATA Act will require years of effort and broad cooperation across government.⁴⁴ Looking beyond May 2017, OMB should continue to pursue additional gains in opening federal spending data.

Although the DATA Act strengthens and standardizes federal spending data, the law does not relieve federal agencies of their legacy reporting requirements, so agencies will have to simultaneously report the same information twice. The Department of the Treasury and OMB should announce their intentions to remove legacy reporting requirements as the DATA Act reporting structure comes to fruition. This announcement will further underscore the importance of all agencies making the switch to the DATA Act Information Model Schema and reduce the future reporting burden for all federal agencies.

The DATA Act does not apply to every stage of the federal spending lifecycle. Notably, the President's annual budget proposals and Congressional appropriations reside outside the requirement.⁴⁵ In support of the DATA Act and open data initiatives, the President should develop the administration's budget using the DATA Act Schema and request that Congress follow suit. Similarly, the DATA Act Information Model Schema does not cover agencies' payment information, the granular, checkbook-level records of each transaction. The President should direct OMB and the federal agencies to extend the Schema to cover payments.

Full implementation of the DATA Act will make it dramatically easier to understand federal government spending.⁴⁶ Federal agencies will have better data to support management decisions and will be able to link to performance data to assess and improve their operations. Researchers and nonprofits will be able to explore trends and identify opportunities to reduce costs. Citizens will have the opportunity to understand the programs their tax dollars support. The DATA Act has the potential to enhance government performance, reduce expenditures, and increase accountability—but only if the federal government continues to strongly support its implementation.

**FULL
IMPLEMENTATION
OF THE DATA ACT
WILL MAKE IT
DRAMATICALLY
EASIER TO UNDER-
STAND FEDERAL
GOVERNMENT
SPENDING.**

RECOMMENDATION 6:

Standardize reporting data for federal grants to help make that data more accessible and useful.

FIRST YEAR

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES & OFFICE OF MANAGEMENT AND BUDGET

ACTION PLAN:

- In August 2017 the Department of Health and Human Services (HHS) expects to finish testing the Common Data Element Repository Library (CDER) to evaluate its effectiveness in reducing the burden of grantee reporting. The Office of Management and Budget (OMB) should build on this foundation to standardize the data structure of all grantee reports, across the federal government.
 - If HHS' testing shows that the CDER Library can reduce grantees' reporting burden through automating their compliance operations, then OMB should exercise its existing authority under the DATA Act to require all agencies to adopt the CDER Library as the official data structure for grant reporting.
 - OMB should convene representatives from grant-awarding agencies and grant recipients to determine how to best implement a national roll-out.
 - OMB should update requirements for the single audit to ensure it aligns with the CDER Library and eliminate the redundant data fields collected during regular reporting requirements.

The federal government supports state, local, tribal, and territorial governments, as well as domestic private and nonprofit organizations, through a suite of nearly 2,300 grants and other financial assistance.⁴⁷ These grants, totaling \$624.4 billion in fiscal year 2015,⁴⁸ help fund nearly every facet of public services, including economic support, research, infrastructure, and many other programs. Yet the federal government has difficulty tracking and accounting for this spending, in part because the more than 28 federal departments and agencies administering grants do not collect information from grantees in a standardized way.

These differences also substantially increase the reporting burden for grantees, primarily state and local governments, who must complete a multitude of forms and reports, each asking for information in its own way. Local governments are under continuous pressure to do more with less, and onerous reporting requirements can be an obstacle in applying for federal grants. The difficulty of complying with reporting requirements and restrictive spend-down timelines has left many cities and towns unable to use the grant money they need.

The Digital Accountability and Transparency Act of 2014 (DATA Act) took a step toward addressing this challenge. The DATA Act directed the Office of Management and Budget (OMB) to develop a standardized electronic version of all federal grant reporting requirements. OMB delegated this task to the Department of Health and Human Services (HHS), which de-

veloped a data dictionary of more than 11,000 data elements used in grant reports, known as the Common Data Element Repository Library.⁴⁹ HHS is currently testing whether this taxonomy will enable grantees to reduce administrative effort and costs by submitting their grant reports electronically. HHS plans to evaluate the taxonomy's effectiveness by August 2017.⁵⁰

If the Common Data Element Repository Library reduces reporting burden, OMB should mandate its use across the federal government for grant reporting beginning in fiscal year 2019.[†] To accomplish this goal, all grant-awarding federal agencies, along with key actors from state, local, and tribal governments, will need to participate in developing a plan for national implementation. They should produce this implementation plan by January 2018. The plan should include full government-wide implementation, including a necessary adjustment period. Additional efforts to reduce burden could include using the same collection mechanism for the single audit program, a largely duplicative review required for larger grantees.⁵²

By replacing document-based grant reporting with standardized, open data, the government can improve the management, transparency, and outcomes of all grants. Implementing the Common Data Element Repository Library across the federal government is the first step toward making this transformation. Additional standardization across grant reporting, particularly with respect to the tracking and reporting of unspent funds, will also be necessary. Even more importantly, streamlined grant procedures across all federal agencies could make the federal grant process much simpler for grantees and can serve as a cornerstone for engendering greater transparency—opening federal grant data for citizens and other stakeholders in expanded and meaningful ways.

**REPORTING ON
FEDERAL GRANTS
THROUGH OPEN
DATA CAN
IMPROVE GRANT
MANAGEMENT,
TRANSPARENCY,
AND OUTCOMES.**

† OMB should confirm that appropriate enabling authority exists to mandate the noted activities.

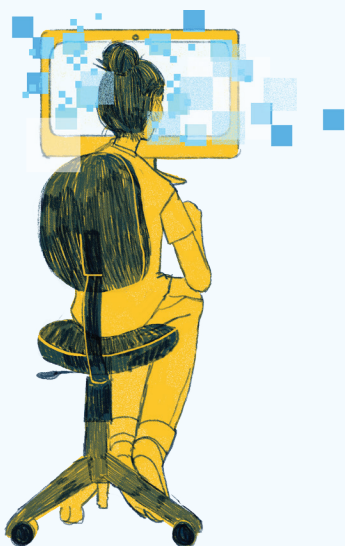
RECOMMENDATION 7:

Enable data users to voluntarily provide direct feedback that will improve data quality.

FIRST YEARTHE WHITE HOUSE OFFICE OF SCIENCE AND TECHNOLOGY POLICY

ACTION PLAN:

- Use federally sponsored challenges and extensive user feedback to improve federal data quality.
 - The Office of Science and Technology Policy (OSTP) should write a memo by January 2018 to provide guidance to all agencies to use direct feedback mechanisms, including crowdsourcing, to improve data quality, and request that agencies engage with their data to identify errors and clean datasets.
 - The memo should describe methods for successfully eliciting and incorporating user feedback, including circumstances in which user participation would be limited to recognized subject-matter experts to ensure data accuracy.
- Agencies and departments providing federal data resources, such as data.gov and other agency-specific portals, should develop effective online feedback channels for users to suggest improvements to data quality on those portals.



Open government data is most useful when it is accurate, timely, and of high quality—but many federal datasets fall short. In 2016, the White House Office of Science and Technology Policy (OSTP) and the Center for Open Data Enterprise hosted an Open Data Roundtable focused on data quality with over 75 attendees.⁵² Several expert participants confirmed that most government data requires considerable cleaning and improvement to make it useful. Additionally, the Government Accountability Office (GAO) has issued hundreds of reports⁵³ with recommendations on improving data quality across dozens of federal agencies, such as the Census Bureau⁵⁴ and Internal Revenue Service.⁵⁵

The federal government can improve data quality by applying a concept that has been demonstrated as an effective practice across the tech community: direct user participation.

Similar to citizen coders addressing bugs in open source software, data users can help identify and fix government data quality problems. OSTP should draft a memo encouraging all agencies administering major open data portals, including data.gov, to provide channels for feedback and proactively invite the public to evaluate and improve data according to basic quality requirements. These channels would go well beyond the “report a bug” option that is currently available on many websites but seldom used. More effective feedback channels can give government data stewards input ranging from simple data corrections (for example, correcting an address or name in a database) to expert insights on addressing deeper quality issues. The memo should include methods for successfully accomplishing this goal. Several government agencies and initiatives have developed effective models for using public input to improve data quality.

One such model is the Department of Health and Human Services’ (HHS) [Demand-Driven Open Data project](#). This project provides users a pathway to tell HHS what data they need and creates a transparent feedback loop that ensures follow-up and follow-through.⁵⁶ This ongoing project has had a range of positive effects on HHS data quality, including improving machine-readability, helping identify and eliminate manual mistakes, and surfacing opportunities for standardization.

[Streetwyze](#), a tool built through the White House’s Opportunity Project, provides another useful engagement model. Streetwyze collects local neighborhood information from residents and pairs that data with existing government datasets to form open maps

and actionable recommendations—allowing citizens to re-engage to make corrections. For example, citizen input can clarify that a building marked as a grocery store is actually a liquor store.⁵⁷

The [USAID Loan Guarantee Map](#) used a third model, convening a crowd of experienced geospatial data volunteers in the Standby Task Force and GIS Corps to review 117,000 loan records and clarify 10,000 difficult-to-identify data points. The volunteers accomplished this task in only 16 hours. Working in partnership with the private sector, USAID customized the event to share data with a community of experts before opening the data to the public.⁵⁸

The private sector has already embraced user feedback initiatives, including crowdsourcing, to improve data quality. Google’s map-maker program allows any Google Maps user to share information about places he or she knows, identifying errors and thus boosting the quality of Google map data. The platform gives increased moderation and editing authority to users who have regularly submitted accurate information.⁵⁹ This two-tiered system allows average users to build expertise, rewards power users, supports engagement, and results in timely, accurate data.⁶⁰

The challenge model, in which government agencies hold competitions inviting the public to solve problems, can also improve data quality.

The U.S. Patent and Trademark Office (USPTO) launched a public competition to provide solutions for data disambiguation, addressing the “ambiguous” instances where the identities of inventors or organizations are not clear. This problem results from repetition or overlap when a single inventor or organization appears in the database under slightly different names, or, conversely, when different inventors share the same name (there are several different inventors registered as “Steve Jobs” and “Steven Jobs,” for example).⁶¹ The winning team created a solution that “uses discriminative hierarchical core reference as a new approach to increase the quality of PatentsView data.”⁶² The USPTO publishes the improved data on [PatentsView](#), a prototype platform to open and visualize U.S. patent data.⁶³

The USPTO has also developed a channel and tools for ongoing feedback on its data resources. The USPTO’s [Developer Hub](#) tool,⁶⁴ which provides access to USPTO’s extensive data collection and APIs to improve accessibility, includes an online community to gain demand-driven requirements from its users, as well as resources for data visualization. Another project, the USPTO [Open Data and Mobility program](#), is advancing how the USPTO provides data, promotes transparency, and empowers data-driven decision making.



GOAL II: DELIVER DIRECT BENEFITS TO CITIZENS AND COMMUNITIES

To be most effective, government policies and programs should start by considering the citizens and communities the government seeks to engage and help. Open data policy should often begin with them too. By identifying the challenges these communities face, the government can strategically select open data projects that will bring the greatest benefit.

RECOMMENDATION 8:

Create a National Hunger Heat Map to help distribute food supplies to areas of greatest need and decrease hunger.

FIRST 100 DAYS**U.S. DEPARTMENT OF AGRICULTURE****ACTION PLAN:**

- The U.S. Department of Agriculture (USDA) should develop an online National Hunger Heat Map to facilitate food distribution to families facing food insecurity within the first 100 days of the new administration. The USDA should combine its own data with data from the Food and Drug Administration (FDA) and the Census Bureau to create the map.
- As a second step, the USDA should invite food banks and food distribution initiatives in the United States to report low or excess supplies of food in real time, and then connect that data to the National Hunger Heat Map.
- In partnership with local governments and organizations, the USDA should use the map to supply food to more Americans in need. The map will enable long-term planning for distributing food more reliably to food deserts, while also enabling distribution of food to meet immediate needs.



Too many Americans are struggling to feed themselves and their families. Poverty and unemployment are common factors contributing to food insecurity. The most recent statistics, from 2014, showed that 42.2 million Americans lived in food insecure households, including 29.1 million adults and 13.1 million children.⁶⁵

The effects of hunger are significant, especially on children. A 2008 study found that “food insecurity, even at the least severe household levels, has emerged as a highly prevalent risk to the growth, health, cognitive, and behavioral potential of America’s poor and near-poor children.”⁶⁶ The Children’s Sentinel Nutrition Assessment Program found that, after adjusting for confounders, food-insecure children had 90 percent greater odds of having “fair/poor” health (versus “excellent/good”), and 31 percent greater odds of having been hospitalized since birth, than similar children in food-secure households.⁶⁷

To address hunger in America, the U.S. Department of Agriculture (USDA) should use data from national surveys such as the Current Population Survey⁶⁸ to identify locations with the greatest needs. As a first step, the USDA should combine its own data with data from the Food and Drug Administration and Census Bureau to create a National Hunger Heat Map—a map to compare areas of food insecurity with food distribution efforts. The Map could draw on the [Feeding America Map](#), which provides insights on how food insecurity impacts each county in the United States.⁶⁹ The National Hunger Heat Map will provide additional detail on food insecurity, perhaps at the census tract level, and include food distribution points, such as food banks and food retailers.

As a second step, USDA should enable food banks and other food distributors to report food shortages or excess supply on a daily basis. By putting this data together with the patterns shown on the National Hunger Heat Map, USDA could run a national data-driven effort to move food quickly from areas of oversupply to areas of need. The map would also be a resource for the hungry, who could log in to see locations with excess food.

This national effort would build on a number of local initiatives that have already proven effective. In Washington, DC, for example, the [Capital Area Food Bank Hunger Heat Map](#) has led to programs that provide healthy food in local elementary schools,⁷⁰ deliver a mobile nutrition program to three hundred children a week in Virginia,⁷¹ and support other neighborhood initiatives in Maryland.⁷²

Additionally, [Waste Not Orange County](#) is a successful initiative in Southern California. It has distributed over 300 tons of surplus food by connecting grocers and restaurants to food recovery agencies.⁷³ With the data gathered from their food insecurity screen tool, the county has created a map of over 230 regularly distributing food facilities to make it easier to identify pantries that donate or accept food. Other county resources include food banks and 2-1-1 Orange County.⁷⁴ Their Community Toolkit includes helpful information for health practitioners screening patients for food insecurity and for food facilities with concerns about the legality of donating food. In just one year, the program helped Orange County reduce the portion of the population facing food insecurity from 12.9 percent to 10.9 percent, helping thousands of residents meet the nutrition needs of their families.⁷⁵

**A NATIONAL
HUNGER HEAT
MAP CAN HELP
MORE THAN
40 MILLION
AMERICANS WHO
LIVE WITH FOOD
INSECURITY.**

RECOMMENDATION 9:

Continue to support and expand local data-driven climate resilience projects to help communities prepare for severe weather incidents.

FIRST 100 DAYS
THE WHITE HOUSE

ACTION PLAN:

- The White House should continue to partner with businesses and nonprofits to further develop the Partnership for Resilience and Preparedness (PREP) and its platform for sharing weather and climate-relevant data within the first 100 days of the administration.
 - The Department of Housing and Urban Development (HUD) should incentivize communities to develop resilience plans based on weather and other data, and share them through PREP so communities can learn from each other. HUD should use these plans to allocate increased funding for resilient housing and infrastructure projects.⁷⁶
-



Extrême weather events are occurring with greater frequency.⁷⁷ Severe weather, such as thunderstorms with damaging winds, heat waves, tornadoes, large hail, flooding, and winter storms, is devastating communities across the country and resulting in casualties, homelessness, and infrastructure destruction. Between 2010 and 2015, the National Oceanic and Atmospheric Administration recorded 59 climate and weather disasters that caused at least one billion dollars in damage in the United States, compared with only 39 the previous five years.⁷⁸ Communities around the country need additional support to prepare.

A growing number of communities, businesses, and nonprofit organizations are looking to assess climate vulnerability and to develop resilience plans. By collecting, maintaining, and sharing weather- and climate-relevant data, the federal government and its partners can help communities prepare for these extreme events. However, efforts to turn data into actionable plans are constrained by several challenges, as robust, actionable data is difficult to find, access, and use.

To overcome these challenges, in September 2016 the White House Office of Science and Technology Policy, World Resources Institute, U.S. Global Change Research Program, and a network of partners launched the Partnership for Resilience and Preparedness (PREP), part of the Global Partnership for Sustainable Development Data. PREP strives to strengthen climate resilience efforts around the world by promoting collaboration among producers and users of information, fostering standards to make data more accessible and interoperable, and developing platforms that improve data accessibility and knowledge sharing.

PREP offers the opportunity to make data-driven resilience efforts more visible and share lessons learned. The PREP platform requests that communities upload and share their own data and information, recognizing that local action requires local data and knowledge.⁷⁹ For example, the city of Sonoma, California created a dashboard on PREP to share their climate resilience projects, their most valuable climate data, and helpful tools.⁸⁰

In addition to continuing to support and promote PREP, the White House should instruct other agencies to work with PREP on areas relevant to their work. The Department of Housing and Urban Development (HUD), in particular, should work with PREP to support its resilience programs. In 2016, HUD and the Rockefeller Foundation awarded \$1 billion across 13 states and communities through the National Disaster Resilience Competition. The funding helps grantees prepare for future natural disasters such as floods and tornados. Sharing those stories through PREP would allow communities to learn from each other and create a collective community of practice.

Finally, the White House and its partners in PREP should actively seek opportunities to work with stakeholders outside of government. Insurance companies, state and local governments, and individuals can use PREP as a valuable tool to leverage the most current climate data. With greater access to sophisticated climate models and open data, the insurance sector can improve its delivery of insurance coverage and buyers will have a better understanding of their own risks. Data-driven solutions can advance climate resilience in an effort to save lives, protect infrastructure, and withstand record-setting weather systems.

RECOMMENDATION 10:

Launch a Police Data Initiative 2.0 to collect data on police violent encounters in a standardized, open format for transparency and research.

FIRST 100 DAYS**THE WHITE HOUSE AND FEDERAL BUREAU OF INVESTIGATION**

ACTION PLAN:

- Within the administration's first 100 days, the White House should announce a Police Data Initiative 2.0 that will support state and local law enforcement agencies to comply with the Federal Bureau of Investigation's (FBI) new system for tracking violent police interactions in an open data format. This will provide greater accountability for participating state and local law enforcement.
 - The White House should ask for commitments from state and local law enforcement agencies to fully comply with the FBI's new initiative and to publicly release the data.
 - The White House should work in partnership with the FBI to provide free, open source resources for tracking, reporting, and publicly releasing data on police interactions to all state and local law enforcement agencies and related agencies, such as the prosecutor's office and public health agencies.
-

In October 2015, James B. Comey, the director of the Federal Bureau of Investigation (FBI), called the government's effort to track deaths during police interactions "unacceptable" and "embarrassing and ridiculous."⁸¹ Since 2003,⁸² the FBI has used the Arrest-Related Deaths Program to track people who died during arrest or in police custody, but a 2015 Bureau of Justice Statistics report found that the program only accounted for about half of the expected number of law enforcement homicides in the United States.⁸³ The finding is consistent with the realization that the FBI figures on police shootings were significantly lower than those in independent databases assembled by the *Washington Post*⁸⁴ and the *Guardian*⁸⁵ newspapers.

This lack of an open, centralized, government-run database on violent police encounters limits police departments' ability to identify challenges and work to correct them and makes it difficult for academic and nonprofit organizations to study critical questions of police-community relations. Moreover, if the public does not have the data it needs to make "fair and informed judgments" about its police, that perceived absence of transparency can undermine trust and legitimacy.⁸⁶

To address this challenge, the FBI announced that it will replace and improve the system for tracking fatal police shootings by 2017.⁸⁷ The new system will cover nearly 20,000 state and local law enforcement agencies and 685 medical examiner or coroner's offices.⁸⁸ It will track all incidents of officer-caused serious injury or death, including shootings, the use of stun guns, and pepper spray.⁸⁹

For this new system to be successful, the federal government must support local law enforcement agencies to track and report this information. Although the FBI currently requires all police departments to report violent encounter data to the Bureau of Justice Statistics,⁹⁰ only three percent of the nation's 18,000 state and local police agencies have done so since 2011.⁹¹ Many police departments around the country are small and lack the resources to track and report the necessary data. As of 2008, about half (49 percent) of all police agencies in the U.S. employed fewer than 10

full-time officers.⁹² An open source solution can help these departments open their data without overtaxing their limited resources.

Prior to the FBI's announcement of a system upgrade, the White House launched the Police Data Initiative (PDI), a community of practice committed to improving the relationship between citizens and police through data and increased transparency.⁹³ As of April 2016, 53 state and local police jurisdictions, covering more than 41 million people, had committed to the PDI, and collectively released over 90 datasets.⁹⁴ While many state and local police departments have used the PDI to jump-start open data and transparency efforts, others that committed to PDI have not taken any action or published any new datasets.

The White House should support and build on the FBI's effort by announcing the Police Data Initiative 2.0 (PDI2.0). Under PDI2.0, state and local law enforcement agencies and their supporting offices would make clear commitments to fully participate in the FBI's initiative, publicly release the data, and incorporate the data into key elements of their operations, including department metrics and individual performance reviews. In exchange, the federal government would support local law enforcement communities by providing resources and technical assistance. The FBI and White House should seek opportunities to develop open source technology and data models, perhaps in partnership with nonprofit organizations, to give local law enforcement a free, user-friendly resource to meet this reporting obligation.

Developers can look to states and municipalities already moving ahead with similar programs in California, New Orleans, and Indianapolis. For example, in September 2016 California launched [URSUS](#), an open source, all-digital police use-of-force data collection system developed by Bayes Impact, in all 800 law enforcement agencies around the state.⁹⁵ Other law enforcement communities could adapt the code with minimal investment to suit their purposes.

Opening data on police violent encounters in this way will enable government, nonprofits, and academics to explore critical questions about police interactions in America, set goals for improvement, and help law enforcement agencies measure themselves against those goals.

RECOMMENDATION 11:

Standardize and update the government's public data on occupations and required skills to help Americans find jobs.

FIRST YEAR**U.S. DEPARTMENT OF LABOR & NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY**

ACTION PLAN:

- The Department of Labor (DOL) and the National Institute of Standards and Technology (NIST) should work together to standardize job skills data and make that data easily available to improve information for job-seekers, job skills trainers, and employers.
 - By January 2018 DOL and NIST should convene key public and private sector stakeholders to develop a Skills Data Standard for the Occupational Information Network (O*NET), the government's centralized resource on occupational and skills data.
 - DOL and NIST should develop a Skills Network Protocol for applying those standards together with key metadata.
 - DOL should then use these standards to incorporate new data into O*NET, including enabling the O*NET database to automatically incorporate jobs posted online using these standards.
 - DOL should revamp the O*NET website and O*NET data services with an improved user interface, a more scalable API, and better developer documentation, developed with 18F, U.S. Digital Service (USDS), or other government resources.
 - The President's budget should include funding for O*NET improvement.
 - DOL, together with the White House, should explore opportunities for public-private collaboration to fund and implement the continued improvement of jobs-related data.
-

Despite improvement in the unemployment rate over the past several years,⁹⁶ too many Americans still struggle to find good jobs to provide for themselves and their families. As of August 2016, 7.8 million Americans were unemployed,⁹⁷ including nearly half a million veterans.⁹⁸ Finding a job is challenging: The Bureau of Labor Statistics reports unemployed Americans spend a median of 11 weeks between jobs,⁹⁹ and job-hunting is especially difficult for the 2 million Americans who have been unemployed for more than 6 months.^{†,101} High-quality, timely open data can help job seekers and reduce the unemployment rate.

The federal government's Occupational Information Network (O*NET) is a centralized resource for employers, job-seekers, and job skills trainers to help the unemployed find work and plan their careers. O*NET's database includes annually updated information on 974¹⁰¹ occupations,¹⁰² including associated skills, trainings, and experiences. Job seekers can explore the occupational data and then link to one of the 2,500 American Job Centers to identify specific opportunities.¹⁰³

An Open Data Roundtable held by the Department of Labor (DOL) and the Center for Open Data Enterprise in 2015 examined O*NET's current use and opportunities for improvement.¹⁰⁴ While job seekers, academics, and others who interact with O*NET consider it the definitive source for jobs information, they also believe it should incorporate new information from additional sources. DOL updates O*NET through an annual survey that cannot capture rapid changes in job definitions and skills requirements.¹⁰⁵ For this reason, DOL, the White House, and others have been exploring ways to update O*NET regularly with jobs and skills data from the private sector and other sources.

To make it possible to update O*NET with data from diverse sources, the National Institute of Standards and Technology (NIST) should work with DOL, and with key public and private sector stakeholders, to develop a Skills Data Standard for O*NET. NIST should also detail a mechanism for quarterly updates and a set of standard, minimum metadata and language for job postings that link to those skills.¹⁰⁶ The data and metadata standards, and guidelines for using them, can then be applied widely through a Skills Network Protocol developed by NIST and DOL. By clearly defining and standardizing the skills that O*NET uses to describe occupations and providing



consistent metadata with job postings through a Skills Network Protocol, DOL can more easily update the skills associated with occupations.

DOL should also revamp the O*NET website and O*NET data services with an improved user interface, a more scalable API, and better developer documentation. O*NET is currently underutilized: The O*NET site has only 4 million visits per month,¹⁰⁷ compared to 180 million unique visitors to Indeed, a job posting company,¹⁰⁸ and 100 million active monthly users for LinkedIn, a private sector professional networking website.¹⁰⁹ Work to improve O*NET could be supported by 18F, the U.S. Digital Service, or other government resources.

DOL has recognized O*NET's current limitations and requested \$5 million in fiscal year 2017 to modernize the system.¹¹⁰ Congress did not authorize this budget request. Given the potential economic value of improving O*NET, the President's budget should include funding for O*NET improvement going forward.

The White House and DOL should also convene industry and philanthropic leaders to establish a public-private commitment to develop and use the Skills Data Standards and the Skills Network Protocol. As employers, business leaders have a vested interest in improving jobs and skills information that can help the hiring process. The private sector can play several important roles, including:

- Helping to fund the technical work of standards development and O*NET improvement.
- Contributing data using those standards from employment-related websites.
- Collaborating with DOL or federal research agencies to support research on automatically extracting skills data from text in job postings.

† Long-term unemployed includes those who have been jobless for 27 weeks or longer—generalized here to six months.

RECOMMENDATION 12:

Help communities address opioid addiction by opening up data on drug treatment facilities.

FIRST YEAR**U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES**

ACTION PLAN:

- The Substance Abuse and Mental Health Services Administration (SAMHSA) should develop an online portal with information on drug treatment centers, including quality metrics and cost. SAMHSA should release a plan to implement this portal by January 2018.
 - Working with state health agencies and the research community, SAMHSA should establish a list of data elements to collect electronically from drug treatment facilities, including both cost and performance data. SAMHSA should ensure the performance data collected is based on criteria grounded in research and empirical studies.
 - The federal government should collect this data, either directly or through the states, and provide data to the public through an online portal similar to the Department of Health and Human Services' Hospital Compare program.
 - Once the portal is established, SAMHSA should consider adding a confidential, universal patient survey to its data collection.
-

Substance abuse is a longstanding epidemic in America, but the recent surge in opioid addiction and related deaths brings new urgency to addressing the problem. Since 2000, drug overdose deaths have increased 137 percent,¹¹¹ with opioid-overdose related deaths at the forefront. In 2014, overdoses claimed more than 47,000 American lives,¹¹² including nearly 30,000 due to opioid overdose.¹¹³ These alarming statistics demand national action.

Data analysis is now playing an important role in combating opioid addiction. State medical boards and health officials are monitoring prescribing patterns to identify doctors who may be overprescribing opioid drugs.¹¹⁴ The Department of Justice is also developing new methods to share real-time data between public health and public safety officials nationwide.¹¹⁵ In addition to these initiatives, which focus on the use of data by government officials, there is an opportunity to use a different kind of data—open data about treatment facilities—to help people affected by opioid abuse.

Opioid addicts and their families often have difficulty accessing treatment. According to the Substance Abuse and Mental Health Services Administration (SAMHSA), in 2014 only about 12 percent of the 21.2 million Americans who needed treatment for an illegal drug or alcohol problem obtained it.¹¹⁶ While the cost and limited availability of treatment are factors, even people who are committed to getting help have difficulty finding effective treatment: There is no centralized resource they can use to evaluate different options to make an informed choice.

SAMHSA should work with treatment facilities, state health agencies, and the research community to develop a plan to build such a resource. As a first step of the plan, SAMHSA should develop a taxonomy for collecting data from treatment facilities, including performance data and other factors such as treatment availability, services provided, cost, accepted insurance, and any restrictions on treatment (e.g., some facilities do not treat minors or individuals with prior convictions). The nonprofit Treatment Research Institute has developed quality effectiveness ratings for treatment facilities that are based on scientific research. SAMHSA should leverage this effort or similar projects to ensure the performance data is based on empirical principles for what constitutes sound, effective care.

The second step should focus on data collection. SAMHSA should develop a method for reliably and consistently collecting this information from treatment centers around the country. SAMHSA should explore ways to ensure unbiased reporting at regular intervals, including collecting data through state health agencies or partnering with nonprofits, such as the Commission on Accreditation of Rehabilitation Facilities.

The final step of the plan should lead to an open, easy-to-use online portal that displays this information for the public. SAMHSA should consider basing this portal on the Department of Health and Human Services' (HHS) Hospital Compare program. Hospital Compare is a consumer-oriented website that provides information on hospital quality.¹¹⁷ Similarly, the treatment facility comparison portal will provide details on treatment center quality and other basic information that may influence decision making, with similar features such as search by zip code.

OPEN DATA ON TREATMENT OPTIONS CAN BENEFIT OPIOID ABUSERS AND THE FAMILIES AND COMMUNITIES WORKING TO HELP THEM.

In addition to helping opioid abusers and their families, this portal would be valuable for the wide network of individuals working to help addicted people, including police officers, social workers, public defenders, and community leaders. Many communities have embraced a public health solution to the opioid epidemic through medication-assisted treatment and are working to help individuals fight their addiction and avoid incarceration.¹¹⁸

Researchers and policymakers can combine data on treatment facilities with existing datasets from the Centers for Disease Control, the Drug Enforcement Agency, the Federal Bureau of Investigation, and other data from the Department of Justice and HHS, to identify high-priority areas that are in need of clinics, doctors, and treatment centers. They can then use that data analysis to make concrete recommendations to shift federal resources to communities that need them to address the opioid epidemic.

Data on treatment facilities can ultimately serve to improve those facilities themselves. Analysts can review the data collected for this portal to show which kinds of treatment are most effective for specific patient populations, to identify trends in treatment approaches, and to identify facilities that may be providing substandard care. Once the portal is established, SAMHSA should consider developing a confidential, universal patient survey that would support these goals by adding data about patients' experience with these facilities.

RECOMMENDATION 13:

Establish a National Data EnviroCorps in which volunteers collect air, water, and soil quality data and help communities identify and manage environmental risks.

FIRST YEAR**U.S. ENVIRONMENTAL PROTECTION AGENCY**

ACTION PLAN:

- The Environmental Protection Agency (EPA) should launch a National Data EnviroCorps, a community of citizen science volunteers who will collect and share air, water, and soil quality data at a local level by January 2018. Communities can use the data to become more informed about their environment and the research community can use it to evaluate trends over time. The EPA should develop a toolkit for communities to use to collect this data locally.
 - The EPA should develop standards, criteria, feedback, and vetting procedures to ensure data quality, comparability, and usability.
 - The EPA should develop both a centralized national portal for this data and model portals that communities can use for their own data.
 - The EPA should actively support the program through annual prizes and highlighting critical research drawn from its data. The Agency can also use the citizen science data as a method of identifying specific areas that may require additional EPA regulatory attention.
-



The **Flint water** crisis demonstrated that citizens around the country may be unaware of potential environmental threats that can impact their quality of life. For example, although the Environmental Protection Agency (EPA) regulates all community water systems in the United States,¹¹⁹ information on local water testing and results is difficult to locate. Community water suppliers only have to provide this information to the public once a year in an annual report, often called a Consumer Confidence Report.¹²⁰ Information on air and soil contaminants is equally difficult to find. An open data, citizen science initiative can reduce these barriers and empower citizens around the country through easier access to this critical information.

Citizen science programs can serve as major data sources for scientific research, and can be particularly valuable in collecting timely, local data. Citizen scientists make it possible to collect large amounts of data and move large-scale research projects forward cost-effectively. On a local level, these projects engage and inform citizens as well.

The EPA should develop a centralized, coordinated national program, a National Data EnviroCorps, that provides a framework for citizen volunteers to test their water, air, and soil, and then contribute their data to an online EPA platform that hosts environmental quality data from around the country. The [Federal Crowdsourcing and Citizen Science Toolkit](#) can serve as a model for this work, laying out the steps to a successful citizen science project.¹²¹ In addition to providing nationwide data for EPA analysis, the EnviroCorps would enable citizens to assess environmental risks and identify local problems in their neighborhoods. The EPA can also use this data directly to identify areas that require additional EPA attention, including formal sampling and official measurements that may contribute to regulatory action.

The EPA can maximize the benefit of this program by developing standards for data collection, using a feedback/verification method to validate data as other citizen science programs have done, and encouraging students to become EnviroCorps citizen scientists.¹²² EPA can also help citizen scientists find information about available sensors, testing methods, and standards and understand technical issues, calibration/validation parameters, and quality assurance/quality control. EnviroCorps scientists can buy their own water/air/soil testing kits, and local school districts can incorporate testing (including test kit provisions) in middle school science programs.

[Air Quality Egg](#), a Kickstarter-funded initiative that started in 2012, has proven this concept by connecting more than 2,500¹²³ citizen scientists testing air quality around the world.¹²⁴ Similarly, [NoiseTube](#) enables citizens to monitor noise pollution using their smartphones and makes the data centrally available for the public.¹²⁵

Several citizen science initiatives across the country currently play a similar role at a local level. Florida LAKEWATCH, for example, has enlisted thousands of volunteers to collect reliable long-term water quality data for over 1,100 lakes, 175 coastal sites, 120 rivers, and 5 springs across 57 Florida counties; the state uses the data to address lake management issues statewide.¹²⁶ EPA's own volunteer monitoring programs for non-point source water pollution also provide elements that could be useful for helping to build a broader effort.¹²⁷

These citizens science initiatives have demonstrated that citizen monitoring and data gathering can serve as a powerful source for informing communities. The National Data EnviroCorps brings this idea to a national scale and tackles a critical challenge facing American communities.

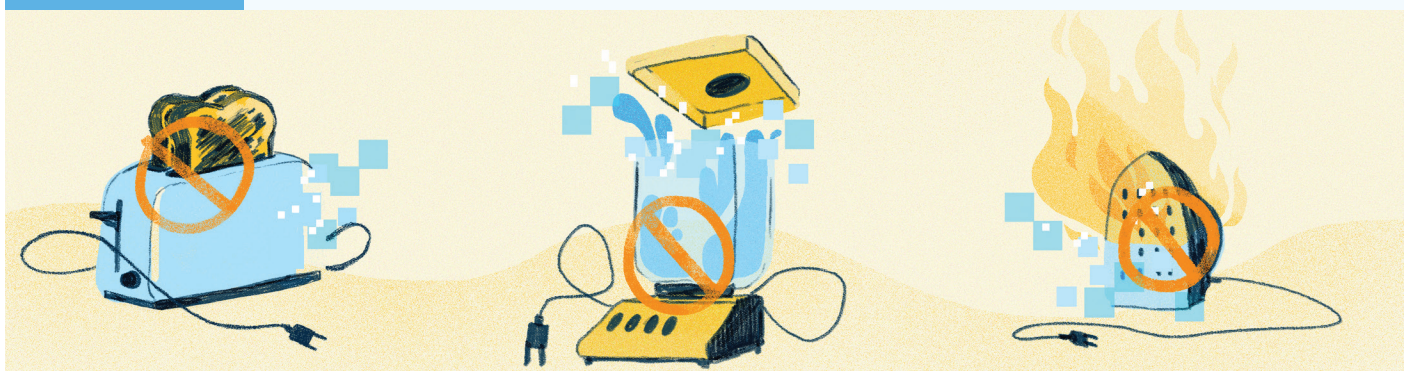
RECOMMENDATION 14:

Develop guidelines for all consumer-facing regulatory agencies to make consumer complaints available as open data to improve products and services.

FIRST YEAR**OFFICE OF INFORMATION AND REGULATORY AFFAIRS**

ACTION PLAN:

- The Office of Information and Regulatory Affairs (OIRA) should publish guidance for regulatory agencies on opening consumer complaint data before January 2018.
 - OIRA should convene representatives from the wide range of regulatory agencies that now collect and analyze consumer complaints. Working with this group, OIRA should identify both challenges and potential benefits in making consumer complaint data publicly available in downloadable, machine-readable form.
 - OIRA should also solicit direct feedback from industry through a request for information.
 - With this input, OIRA should issue guidelines for regulatory agencies to make consumer complaint data available as openly as possible, with appropriate considerations and constraints.
 - The guidelines should help ensure that manufacturers and service providers have an opportunity to respond to complaints or provide additional context.
 - The guidelines should protect consumer privacy and specify whether, when, and how personally identifiable information associated with a complaint may be shared with the company involved or the public at large.
 - The guidelines should weigh any legal and operational issues involved in opening data that may be used for supervision or enforcement.
-



Some of the most important data about consumer products and services comes from consumers themselves, but companies, organizations, and the public often are not aware of consumer complaints. A number of federal agencies now provide consumers the opportunity to “tell their story” about the businesses agencies regulate. Consumer complaints are a tool for regulatory agencies and law enforcement to use in protecting consumer safety and consumer rights. By opening these complaint databases, agencies can help improve consumer markets: They will give buyers the information they need to choose consumer-friendly companies, and will help businesses make proactive changes to address consumers’ concerns.

Agencies with a public-safety mandate use consumer complaints as warning systems for issues to investigate, potentially regulate, and alert the public. The Consumer Product Safety Commission invites consumers to report unsafe products at SaferProducts.gov,¹²⁸ while the Food and Drug Administration’s Adverse Event Reporting System solicits voluntary reports of drug side effects.¹²⁹ The National Highway Traffic and Safety Administration collects vehicle safety complaints as a basis for potential investigations and provides searchable recall information at SaferCar.gov.¹³⁰ The Federal Aviation Administration collects complaints about airplane safety and service, which it analyzes for its monthly Air Travel Consumer Report.¹³¹

Other agencies track consumer problems that, while not life-threatening, can pose the threat of financial loss. The Federal Trade Commission’s (FTC) Consumer Sentinel Network may be the largest: It is “based on the premise that sharing information can make law

enforcement even more effective,” and gives law enforcement members access to millions of complaints submitted to the FTC and other data sources.¹³² In 2015, consumers submitted over 3 million complaints.¹³³ The FTC shares its Sentinel data only with law enforcement officials, not the public. The Federal Communications Commission analyzes complaints on TV, phone, radio, internet, and other communications issues and acts as a mediator between customers and carriers to resolve issues.¹³⁴ The Consumer Financial Protection Bureau (CFPB), with a broad mandate for protecting consumers’ rights, manages consumer complaints about various consumer financial products and services, including student loans, mortgages, and credit cards.¹³⁵

The CFPB is now leading the way in making consumer complaint data available as truly open data, and has demonstrated the value of this open approach.¹³⁶ In July 2013, shortly after the CFPB began releasing its complaints as open data, Yale University published an analysis showing which banks had the best and worst customer service in different areas.¹³⁷ A few months later, *American Banker* took note and recommended that banks “revamp their customer service operations to encourage irate consumers to complain to them instead of turning to the Consumer Financial Protection Bureau, where a complaint could spark added regulatory scrutiny.”¹³⁸

In making its data open, the CFPB has worked to ensure fairness for service providers and confidentiality for consumers. The CFPB obtains consumer consent and redacts all personal information before publishing consumer complaint narratives on its website. Additionally, the CFPB sends each complaint to the company involved in order to confirm a commercial relationship with the customer before publishing the complaint online.¹³⁹ This system gives companies an option to provide a public response to any consumer complaint before or after a narrative goes public.

RECOMMENDATION 15:

Expand open data summer camps to inspire the next generation of data-savvy students and help improve and apply open government data.

FIRST YEAR

U.S. DEPARTMENT OF AGRICULTURE

ACTION PLAN:

- The U.S. Department of Agriculture (USDA) should expand their open data summer camps to include more students and locations in future years.
 - As a first step, USDA should identify another federal agency to partner with to expand USDA's Open Data Summer Camp model by January 2018 with the goal of launching the expanded summer camp program the summer of 2018.
 - For future years, USDA should evaluate the summer camp's effectiveness (for example in influencing college major selection) and then look to duplicate successes across the federal government and diversify city selection around the country.
-

The fields of science, technology, engineering, and math (STEM) account for one in five American jobs¹⁴⁰ and the Bureau of Labor Statistics projects that STEM jobs will grow to 9 million in the United States by 2022.¹⁴¹ Increasingly, an understanding and ability to work with data is core to the development of STEM skills. The workforce has a strong need for the next generation to have proficient data skills, not only for technical areas, but to support analysis and decision making in a wide range of business contexts.¹⁴²

Educators in the United States and other countries have recognized that open government data is a valuable resource for teaching data skills. The company Tuva Labs, for example, has used open data to develop K-12 programs in data literacy. Tuva Labs also has programs around the world to build data analytic capacity for sustainable development, including a demonstration project launched in partnership with the World Bank in Sudan.¹⁴³

Now some government agencies have begun working with teens with a dual purpose: to help teens learn data skills, and to have their help in improving and applying government datasets. The New Orleans Police Department recently took this approach when New Orleans joined the White House's Police Data Initiative. The Department held a three-day summer camp with a local coding academy, giving young developers in training from low-opportunity neighborhoods an opportunity to analyze data on police activities.¹⁴⁴

In 2015, the U.S. Department of Agriculture (USDA) partnered with The GovLab at New York University to develop a larger-scale project: a free, two-week Open Data STEAM (STEM plus agriculture) Summer Camp for middle and high school students. The camp, now held annually in Washington, DC, aims to help American teenagers learn more about data and how data science can improve innovation and security in the nation's food supply. Long-term, USDA hopes

the camp will inspire students to become future data scientists, data analysts, and farmers.¹⁴⁵ The 2015 and 2016 camps were successful,¹⁴⁶ but were only a beginning: They served about 35 students each year and the program has not yet secured long-term funding.¹⁴⁷

To scale this program sustainably, USDA should identify another federal agency, perhaps the Department of Health and Human Services (HHS), as a partner, to expand the open data summer camps. They should develop a plan for expansion by January 2018 and then deliver the joint camp during the summer of 2018. This would enable USDA to create both an "agriculture track" and a second track, broadening the program's appeal to both participating students and potential funding partners. HHS would be a good first partner, as these agencies' shared connection to the biological sciences enables synergies in content and complementary datasets, with overlapping interests in nutrition and public health.

USDA should evaluate the summer camp's effectiveness, and then consider expanding the model across the federal government. By better understanding how the camps can best meet their objectives, such as increasing the student's likelihood of selecting certain college majors, USDA will have firm ground to scale what works across the federal government and add additional cities for camp locations.

Open data camps will directly engage young Americans in open data and government—giving them first-hand experiences with career-advancing skills development. It will help the participating students to better understand the opportunities in open government data and provide a glimpse of working in agriculture, health, and the sciences. Additionally, summer camp students can help improve an agency's own data. The 2015 USDA open data summer camp helped to reveal broken links in USDA's data, which USDA could then fix.¹⁴⁸ The summer camps provide an on-ramp to direct participation in the open government data ecosystem and benefit the sponsoring agencies by helping to improve and apply their data as well.

**DATA-FUELED
SUMMER
PROGRAMS
CAN INSPIRE
STUDENTS TO
BECOME FUTURE
DATA SCIENTISTS
AND ANALYSTS
AND PREPARE
THEM FOR
STEM JOBS.**

RECOMMENDATION 16:

Open up data on housing choice voucher wait lists to support low-income families in finding housing.

FIRST YEAR

U.S. DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT

ACTION PLAN:

- The Office of Housing Choice Vouchers in the Department of Housing and Urban Development (HUD) should partner with a nonprofit organization to standardize local data related to housing voucher wait lists and publish it in a centralized database, launching an initial pilot by January 2018.
 - Working with representative local public housing agencies, HUD should develop standard metadata on the application process and wait list for public housing vouchers, including the maximum rent value for housing voucher, length of the wait list, and how to apply for housing assistance.
 - HUD should make it easy for local public housing agencies to submit their data, and should display it on a single public website for low-income families.
 - HUD should require public housing agencies to update their online information regularly, leveraging outside expertise as needed. The Office of Housing Choice Vouchers should prioritize local communities of interest and provide technical assistance to help them open historical data as well.
-

In many cities, towns, and rural communities across the country, families struggle to afford safe, reliable housing. Although the proportion of renters spending more than 30 percent of their income has decreased in recent years,¹⁴⁹ a 2016 Harvard study found that 11.4 million households spend more than half of their income on rent.¹⁵⁰ The housing choice voucher program assists more than 5 million people, including low-income families, the elderly, and individuals with disabilities, to pay rent for housing on the private market.¹⁵¹ A rigorous evaluation conducted over a four-year period found that vouchers reduced the number of families with children that lived in shelters or on the streets by 75 percent.¹⁵² Reducing relocation for school age children has direct benefits for long-term health, development, and education gains.¹⁵³ Despite these benefits, it is difficult for families facing a housing crisis to obtain basic information on the program and how to take advantage of it.

The housing voucher program lacks a central database on wait lists. The Department of Housing and Urban Development (HUD) funds the program, but it is administered by a network of state and local public housing agencies that determine many of their own policies.¹⁵⁴ The demand for housing vouchers dramatically exceeds supply, so local public

housing agencies set up wait lists. Affordable Housing Online, a website to help low-income Americans find information about housing assistance and other housing opportunities, says that all of the estimated 2,320 public housing agencies have wait lists.¹⁵⁵

HUD regularly receives calls from citizens asking for information on the voucher program that is only available at the local level, such as the amount of housing assistance available in their area, the length of the wait list, and how to join the wait list. The current HUD website refers individuals to the general email and phone number for each public housing authority,¹⁵⁶ which they must then contact one-by-one. Opening this wait list data in a single online platform would reduce the information burden on families seeking housing choice voucher assistance.

In addition to standardizing key wait list fields for each public housing agency, HUD can require local agencies to report data at standard intervals. HUD should explore how to best develop a simple, open source platform for public housing agencies to submit relevant data and then host a public website with the information all in one place. By reducing the difficulty of locating and using public housing vouchers, HUD will remove one barrier for low-income families who seek to move into safer neighborhoods with steady housing.¹⁵⁷

OPENING UP DATA ABOUT HOUSING VOUCHERS ON A SINGLE ONLINE PLATFORM CAN HELP MILLIONS OF FAMILIES.

GOAL III: SHARE SCIENTIFIC RESEARCH DATA TO SPUR INNOVATION AND SCIENTIFIC DISCOVERY

Government-funded research plays a critical role in catalyzing breakthroughs that accelerate discovery, fuel innovation, and drive economic growth. Open data allows scientific researchers, from the laboratory sciences to social sciences, mathematics, and other disciplines, to build directly on one another's previous work in the field, increasing the speed of discovery. Many of the federal government's efforts so far have focused on making scientific publications available. However, as Dr. John Holdren, director of the White House Office of Science and Technology Policy, stated in a 2014 letter to Congress, "... even larger societal benefits will be gained through Federal efforts to make scientific data available for analysis and reuse." The government should take steps to promote increased access to research data in machine readable formats that maximize productive reuse, to facilitate speedy validation of results and enable further breakthroughs.

RECOMMENDATION 17:

Establish a Federal Research Data Council to coordinate the government's commitment to open research.

FIRST 100 DAYS**NATIONAL SCIENCE AND TECHNOLOGY COUNCIL****ACTION PLAN:**

- The National Science and Technology Council should establish a Federal Research Data Council under the White House Office of Science and Technology Policy (OSTP) within the first 100 days of the new administration. The Council will be an interagency support structure to expand current cross-government efforts to promote open science, foster new government open research initiatives, and partner with research institutions to support data sharing and collaboration.
- The head of each agency with a research component should designate a lead for open research that reports directly to him or her and who will work with the Federal Research Data Council to implement government-wide initiatives within their agency.
- OSTP and the U.S. Digital Service should coordinate developing a common search tool that will allow researchers to seamlessly discover publications and data sets from federally funded research regardless of which agency funded the research.
- The Federal Research Data Council should lead a review of federal agencies' public access and data management plans, and implementation efforts for both, to identify best practices and support continued progress.



The United States spends over \$475 billion each year supporting research and development,¹⁵⁹ with over \$140 billion coming from the federal government.¹⁶⁰ This investment drives American innovation and discovery across a broad range of disciplines, from catalyzing medical breakthroughs to discovering more about the universe and supporting technological advances to improve the lives of all Americans.

To realize the greatest benefit from this research investment, researchers should make their data publicly available for others to use as well. Initiatives such as the Human Genome Project have demonstrated the power of opening scientific data in a reusable format so that researchers around the world can rapidly build off one another's findings. An MIT study showed that genomic data released under the Project was applied more widely than proprietary genomic data,¹⁶¹ while a Battelle Technology Partnership Practice evaluation demonstrated the federal investment in genomic research provided \$141 in economic benefit for every dollar spent, with an economic impact of \$796 billion.¹⁶² Major open science initiatives span a wide variety of disciplines, including the Open Source Malaria Consortium, which strives to find a cure for malaria, a disease that kills 600,000 people each year;¹⁶³ the Dronecode Foundation, which unites experts across the unmanned aerial vehicle industry to develop an open source platform for drones;¹⁶⁴ and the Sloan Digital Sky Survey, which uses open science to map images of one-third of the sky.¹⁶⁵

For the past several years the federal government has been working to open up research data, especially federally funded research data. In February 2013 the Office of Science and Technology Policy (OSTP) issued a memo directing all federal agencies that expend over \$100 million annually in research grants to require that grantees make the results of their research publicly available. The memo included specific requirements for ensuring fast, free, public access to research articles, as well as a requirement that researchers applying for federal funding develop a data management plan.¹⁶⁶ As of July 2016, sixteen federal departments and agencies had released public access plans and five had plans under development. At that time, nine agencies required their grantees to develop data management plans for new research progress, and seven more were phasing in such a requirement including requiring the sharing of research data.¹⁶⁷

While agencies are actively implementing these policies, opening up research data is more complex than ensuring access to published articles. Strategies for opening data may need to be more closely tailored to specific scientific disciplines. Opening up research data can involve challenges in data privacy, intellectual

property protection, ensuring proper credit to researchers, and other issues that can inhibit data sharing. Research agencies would benefit from a coordinated effort to identify common solutions to these challenges, develop best practices, and share learnings from different approaches to data management.

To provide this coordination, the National Science and Technology Council should establish a Federal Research Data Council. OSTP should oversee the Council, which will focus on coordinating implementation and expansion of open science throughout the federal government with a particular focus on the data underlying scientific publications. The heads of each agency should designate a lead who will work with the Federal Research Open Data Council to represent his or her agency's perspective and help coordinate initiatives within his or her home agency.

While the federal government has formed several interagency data groups over the last decade, the Federal Research Data Council would put a high priority on research data sharing specifically. A possible model is the Federal Privacy Council that President Obama formed in February 2016.¹⁶⁸ Research data sharing, like privacy protection, is a goal that spans a large number of agencies and an area that involves complex technical, legal, and other considerations. A high-level council formed at the direction of the White House would be able to address these issues across government with authority.

Following this model, the Federal Research Data Council should:

- Help officials in research agencies better coordinate and exchange best practices for data sharing;
- Develop recommendations for OSTP on policies for research data sharing;
- Assess and make recommendations on the data science capacity needed in research agencies to share and integrate data, in coordination with the Federal Chief Information Officers Council and other relevant inter-agency groups; and
- Advise OSTP and the U.S. Digital Service on requirements for a search tool to make it possible to search across agencies, which will help researchers find relevant datasets regardless of where or how they were developed.

The Council should also support the final five agencies to develop and release their public access plans as soon as possible. Since these agencies have not kept pace with others in this regard, it should be a priority to have them release and implement their plans. As part of this initiative, the Council should review existing public access plans to identify best practices, and explore opportunities to take the public access requirement to the next step by requiring free access to data underlying published research articles.

**A FEDERAL DATA
RESEARCH COUNCIL
COULD HELP
ENSURE FAST,
FREE, PUBLIC
ACCESS TO THE
SCIENTIFIC DATA
PRODUCED WITH
\$140 BILLION IN
ANNUAL FEDERAL
FUNDING.**

RECOMMENDATION 18:

Develop an Annual Research Data Census to increase awareness of federally funded research and improve access to data.

FIRST 100 DAYS**NATIONAL SCIENCE FOUNDATION**

ACTION PLAN:

- Within the administration's first 100 days, the National Science Foundation (NSF) should develop and implement an Annual Research Data Census—key information about datasets, their characteristics, and stewardship environments developed by its grantees—in connection with NSF's annual reporting requirements.
 - The NSF should develop a data survey through funded research and pilot it with a sample of grantees. After the pilot, the NSF should require the survey as part of grantees' annual report on [research.gov](https://www.research.gov).
 - The NSF should publish the results of the survey as an Annual Research Data Census, available as open data in machine-readable form.
 - If the NSF Annual Research Data Census is demonstrated to be effective, the White House Office of Science and Technology Policy (OSTP) should mandate developing similar, interoperable data collection across all other science agencies.
-

In February 2013, the White House Office of Science and Technology Policy (OSTP) issued guidelines on increasing access to the data and results of federally funded research. Assistant to the President for Science and Technology and Director of OSTP John Holdren instructed federal agencies to ensure “public access” to federally funded research outputs by requiring federal grantees to make their research papers available for free within 12 months of initial publication. Additionally, to ensure the broader aim of “open science,” OSTP directed agencies to require their grantees to develop plans for sharing their research data in a timely manner, and in a reusable format.¹⁶⁹

As a major step in implementing this guidance, the National Science Foundation (NSF) should develop an Annual Research Data Census to provide public, searchable information on the data, its characteristics, and its stewardship environments.[†] This could be accomplished through a survey adding a small number of additional questions to grantees’ annual reporting requirements. Since the NSF already requires its grantees to report annually on their work through the FastLane system on Research.gov, adding these survey questions will not substantially increase reporting burden.¹⁷⁰

The survey can cover the amount and type of data generated; associated publications; basic metadata; whether the data is open, or will be made open; where the data is hosted and who can access it; and other factors. The Census would hold researchers accountable for sharing their data as much as possible while protecting privacy.

The Annual Research Data Census would allow the tracking of data sharing requirements and serve as

a valuable resource to the scientific community. The Census would enable researchers and interested members of the public to search for data collections that may be difficult to find. Scientists would use the Census to identify others who have been collecting data through similar studies, to access the data if it is available, and to contact the investigator to request access to the data if necessary.

The Data Census would also inform the strategic development of data infrastructure needed to ensure that data is discoverable, accessible, usable, and sustainable for current and future innovation. Just as the Population Census informs the development of roads, bridges, hospitals and other elements of “societal infrastructure” that benefit the public, a Data Census will help identify the infrastructure needed to ensure that data can be fully used for future innovation and to advance work by the research community.

Finally, the Annual Research Data Census would make it possible to follow research trends over time with more insight and information. This information would be especially useful to NSF and other grant makers in helping them prioritize funding areas and look for possible gaps or synergies between different investigators and their lines of work. Census analyses could help show what public repositories and standards researchers use to house and manage their data, and areas that need improvement to make sharing of research data easier.

If the NSF Annual Research Data Census is demonstrated to be effective in enabling secondary research or other follow-on research using existing research data, OSTP should mandate that all other science agencies incorporate a similar census into their reviews of grantees. The government should maintain all data census information in a centralized location.

[†] Note that the stewardship environment (the location and hosting of data at repositories, university libraries, in project environments, etc.) is critical for assessing the sustainability of project data and its risk of loss, damage, or inaccessibility for reproducible open science.

RECOMMENDATION 19:

Coordinate agency efforts to transition to cloud storage and analytics, enabling more research with government data.

FIRST 100 DAYS**GENERAL SERVICES ADMINISTRATION**

ACTION PLAN:

- Within the first 100 days of the new administration, the General Services Administration (GSA) and 18F should coordinate across agencies to identify and prioritize open government research datasets that would benefit the research community through enhanced availability in the cloud.
 - GSA should use bulk purchasing power to obtain cloud storage at discount prices and have 18F support other agencies to open up research data on the cloud through cloud.gov.
 - For large datasets, the government should pursue individual contracts for cloud storage and prioritize high-value research datasets.
-

The federal government has been slowly shifting data storage from agency-owned data centers to cloud-based services since 2009,¹⁷¹ but many high-value research government datasets are not yet available on the cloud. Putting data on the cloud makes it more easily accessible for researchers and facilitates a variety of analyses. This is particularly useful for large datasets, which would previously require researchers to buy storage, download the dataset or request it via CDs, and then obtain enough processing power to analyze it—a process that could take weeks or months. With cloud storage, researchers can analyze data and develop new products without downloading and storing the data locally.

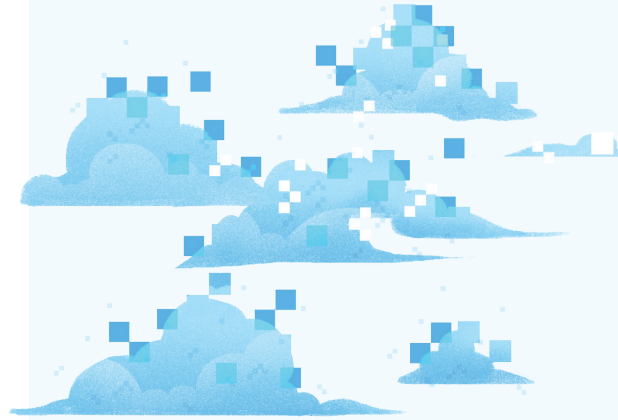
Although both the General Services Administration (GSA) and 18F provide services to support agencies moving data to the cloud, each agency determines its own path toward cloud-based services while complying with requirements of the Federal Risk and Authorization Management Program (FedRAMP) and National Institute of Standards and Technology guidance. The process has led to a fragmented demand for cloud resources, duplicative systems across and within agencies, and management challenges. Agencies reported that they planned to spend \$2 billion on unique cloud computing systems in fiscal year 2016.¹⁷²

By working across the government, GSA and 18F can be more strategic about obtaining cloud storage solutions at discount prices and prioritizing datasets for the cloud. GSA should leverage its purchasing power to bulk purchase cloud storage and 18F should use the cloud.gov platform to help distribute it across the government.

GSA and 18F should also prioritize efforts to open large datasets to support research. Examples of existing datasets that would directly support research initiatives include NEXRAD data, which contains weather radar data from the National Weather Service;¹⁷³ NAIP imagery, which is orthophotography that depicts agriculture growth patterns around the

country;¹⁷⁴ and LIDAR data, a method for surveying that uses laser light to measure distance to a target, as well as many others. Making this data more easily accessible through the cloud, eliminates the lengthy logistical hurdles that currently exist for researchers to use these datasets.

By collectively assessing needs across the federal government and making bulk purchases for cloud storage and processing power, GSA can hasten efforts to move key datasets to the cloud while reducing costs. Cloud storage has already proved extremely valuable for research efforts involving key datasets. For example, the U.S. Geological Survey and NASA partnered with Amazon Web Services (AWS) to host Landsat data, spatial imagery and information on the Earth's composition, in the cloud. The data is available to anyone for free with daily updates, often within hours of production.¹⁷⁵ The National Oceanic and Atmospheric Administration (NOAA) partnered with AWS, Google Cloud Platform, IBM, Microsoft, and the Open Cloud Consortium to put vast amounts of data in the cloud as well.¹⁷⁶ Similarly, the Broad Institute of MIT and Google collaborated to host Broad's massive genomic dataset (Broad DNA sequencers produce more than 20 terabytes of data each day) on the cloud. After implementing optimizations, the Broad-Google collaboration reduced costs while improving processing time eight-fold.¹⁷⁷ These initiatives have demonstrated that the public-private partnership model is effective for cloud storage of large, high-value datasets.



RECOMMENDATION 20:

Work with major research organizations, scientific publications, and professional associations to move toward requiring researchers to publish their data in open, reusable formats.

FIRST 100 DAYS**THE WHITE HOUSE OFFICE OF SCIENCE AND TECHNOLOGY POLICY**

ACTION PLAN:

- The White House Office of Science and Technology Policy (OSTP) should lead a 100-day scoping exercise on requiring researchers to share high-quality, reusable data underlying scientific publications at the time of publication.
 - OSTP should convene representatives of journals, universities, and professional associations to identify potential strategies for helping to shift incentives toward free data sharing.
-

Most research organizations' incentive structures do not encourage (and may even inhibit) the broad sharing of research data. Universities' promotion and tenure standards emphasize the publication of peer-reviewed articles in high-impact journals, and do not address the need for—or benefits of—publicly sharing the data behind those articles. This limits researchers' willingness to share their data. Instead, they are incentivized to restrict access to important research data and maximize the number of publications they can derive from the data.

A 2014 survey of 90,000 recent authors of papers in health, life, physical, and social sciences, and humanities, showed just over half of researchers shared their data once an article was published. Of those who did not, 26 percent reported that they did not share data because they feared it would be scooped or misused. The same survey asked respondents what would motivate them to increase data sharing. Social science, humanities, and physical science researchers sought increased visibility and impact for their work; life science researchers sought guaranteed credit; and health science researchers sought guarantees regarding privacy and other ethical issues.¹⁷⁸

The research community could boost access to open research data by making changes to the current incentive structure, increasing visibility and ensuring credit for publishing data openly. However, it will take time and careful consideration to change the incentive systems that currently drive scientific research. To set a framework for this effort, the White House Office of Science and Technology Policy (OSTP) should work with major scientific journals, universities, and professional associations to undertake a 100-day scoping exercise on changing incentives for data publication. This exercise should identify:

- Promising strategies, including changes to the citation system, the promotion and tenure process, and other approaches;

- Key issues, including differences between scientific domains that impact data publication, economic considerations, and consistency of academic policies;
- Major partners to implement change after the scoping phase, including not only universities and journals but professional associations, such as the Research Data Alliance and CODATA, that address incentive issues through working groups; and
- Near-term actions that can shift incentives for data publication.

This scoping exercise should support changes in the scientific publication process. For example, journals can develop appropriate review systems to ensure that published datasets are rigorously reviewed and useful for additional research efforts, including replication of existing published work. Scientific publishers may also develop citation systems that give researchers visible credit when other scientists use their data.

OSTP could also explore requiring that all federal research funding agencies develop policies that require grantees to share the data underlying scientific publications for free at the time of publication.

In addition, OSTP could work in partnership with major American research universities to support new academic incentive structures for publishing reusable data. The incentives may already be starting to align naturally, as research shows that studies with publicly available datasets receive a higher number of citations than similar studies without available data.¹⁷⁹ By incorporating open publication of datasets into promotion and tenure decisions for faculty, as well as tying open data to advancement or prestige for students and professional research staff, universities can collectively play a significant role in hastening the advance of open research and the broad benefits it delivers for the research community.

THE WHITE HOUSE SHOULD WORK WITH MAJOR SCIENTIFIC JOURNALS, UNIVERSITIES, AND PROFESSIONAL ASSOCIATIONS TO CHANGE INCENTIVES FOR DATA PUBLICATION.

RECOMMENDATION 21:

Establish international standards for collaboration and data sharing, beginning with a pilot in Arctic research.

FIRST YEAR
THE WHITE HOUSE

ACTION PLAN:

- The U.S. Arctic Research Commission should partner with research communities in Canada, Russia, Norway, Denmark, and other relevant countries to establish international standards for data sharing across disciplines and organizations focused on Arctic research by January 2018. This partnership should build on the momentum from the White House Arctic Science Ministerial and serve as a pilot for a larger effort to develop standards for international research.
 - The Commission should review successful efforts to develop international standards and share interoperable data in the public health fields, for example around Ebola and the Zika virus.
 - As the Arctic Research Commission successfully establishes new standards, the White House Office of Science and Technology Policy should study this effort and others to determine how best to facilitate international data standard development in other research areas as well.
- 

Researchers in most disciplines have not yet come to a consensus on international open data standards in their domains. In fact, for many disciplines, scientists and researchers may use vastly different data structures and protocols even within a single country. Although lack of standardization is just one of many barriers to international data sharing, it represents a significant obstacle to collaboration on some of the world's most difficult challenges.

In Arctic science, the international research community has an opportunity to solve this problem and create a model for international standard-setting in the process. In September 2016, science ministers from 25 governments met for the first-ever Arctic Science Ministerial to Advance International Research Efforts.¹⁸⁰ They signed and released a joint statement that charts a "new collective approach in Arctic science" including "Strengthening and Integrating Arctic Observations and Data Sharing."¹⁸¹

This commitment to collaborate primes the community for the next step—ensuring they can share their data internationally and analyze the data in concert.

In a September 2016 addition to the U.S. Open Government National Action Plan, the White House committed to "increase open scientific collaboration on the Arctic."¹⁸² This commit-

ment noted the need for a global approach to Arctic science and cited the Arctic Science Ministerial as a first step to "create a context for increased international and open scientific collaboration on the Arctic over the longer term." A focus on data standards would be a logical next step in this collaboration.

Building on momentum from the Ministerial, the Arctic Research Commission should partner with stakeholders in key Arctic research countries and look to previous successful efforts to standardize and share international data, such as the World Health Organization norms for sharing data in public health emergencies.¹⁸³ They should also consider new data sharing practices for projects that involved direct collaboration, such as the Centers for Disease Control's use of GitHub to collect and share data related to the Zika virus. This effort opens Zika virus data from 13 countries, Puerto Rico, and the U.S. Virgin Islands.¹⁸⁴

By January 2018, once this collaborative effort on Arctic research has developed international data standards, the White House Office of Science and Technology Policy (OSTP) should study this effort and others to identify successful methods and processes for international data sharing. At that time, OSTP should identify another research domain to begin the process of setting international standards. Although sharing research data involves unique considerations for each discipline, the lessons learned can contribute to future efforts across domains.

THE U.S. ARCTIC RESEARCH COMMISSION SHOULD PARTNER WITH RESEARCHERS IN OTHER COUNTRIES ON INTERNATIONAL STANDARDS FOR DATA SHARING.

RECOMMENDATION 22:

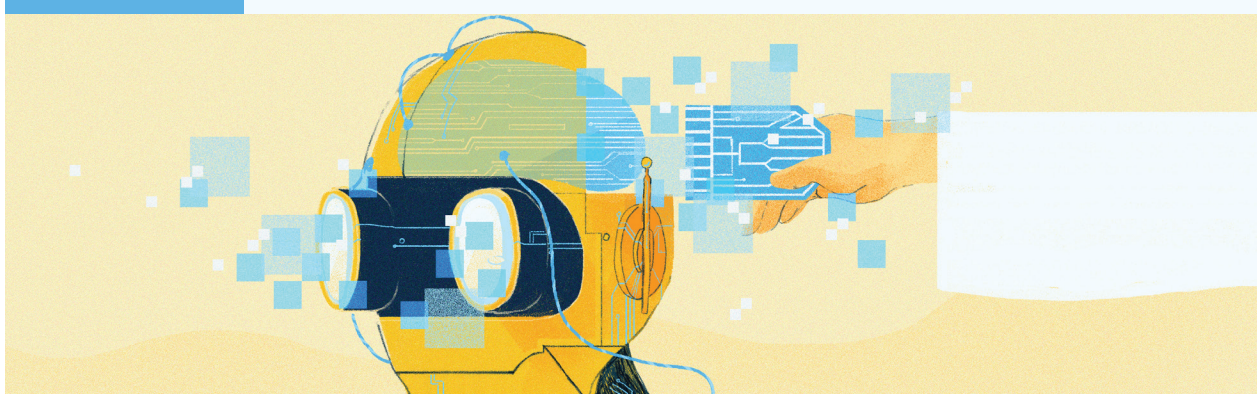
Identify and publish large, high-quality datasets across all fields for use in machine learning to support advances in artificial intelligence.

FIRST YEAR

NATIONAL SCIENCE AND TECHNOLOGY COUNCIL & INTELLIGENCE ADVANCED RESEARCH PROJECTS ACTIVITY

ACTION PLAN:

- The National Science and Technology Council (NSTC), in partnership with Intelligence Advanced Research Projects Activity (IARPA), should review feedback from the recent IARPA solicitation on artificial intelligence (AI) to identify key government datasets that can be opened up for use in machine learning to support AI development across all fields by January 2018.
 - The NSTC should work with IARPA to develop a prioritized list of datasets for release.
 - The NSTC should then convene members of the AI community to provide input on these priorities, finalize a plan for releasing datasets, and develop a timeline and resources to do so.
-



Artificial intelligence (AI) is already playing a significant role in the daily lives of Americans, including voice activated personal assistants in smartphones, website translation, and automated driving features. Additional AI advances are around the corner, and the technology industry is planning for the next wave.¹⁸⁵

Researchers believe AI will have a great impact on the future economy, including public benefits in the medical field, transportation, public safety, and more.¹⁸⁶ Eventually AI could help doctors diagnose patients and suggest treatments tailored to the individual or serve as a tool for teachers to customize lesson plans for each student's personal needs.¹⁸⁷ AI could also help to efficiently allocate government funds.¹⁸⁸

Machine learning—a process in which computers continually improve analytic capacity as they work with more and more data—drives AI. Machine learning requires large, unbiased datasets to create accurate models within the domain of interest. For example, developers used ImageNet, an annotated database of over 14 million pictures,¹⁸⁹ to “train” computers to properly classify images. Other datasets may be valuable for geospatial analysis, language analysis, or other aspects of machine learning.

The National Science and Technology Council (NSTC), in partnership with Intelligence Advanced Research Projects Activity (IARPA) and other interested agencies, should launch a project to ensure that relevant, government owned “training datasets,” as identified by AI researchers and other stakeholders, are made readily available.

This would address a current bottleneck in AI development: the relatively small amount of public data available to train AI systems and enable them to reach their full potential.

As a first step, the Machine Learning and Artificial Intelligence subcommittee of the NSTC should work with IARPA, which should have strong indicators of demand from responses to a recent solicitation¹⁹⁰ on overarching questions in AI. The NSTC subcommittee and IARPA should draft recommendations for prioritizing datasets for use in AI training. The NSTC should then convene a group of government and industry AI specialists to determine the datasets in greatest demand, analyze the barriers to opening those datasets, and develop plans to open them.

If privacy concerns restrict data opening, the NSTC should develop a data enclave similar to the Centers for Medicare & Medicaid Services data enclave within the Department of Health and Human Services or the National Renewable Energy Laboratory's Secure Transportation Data Center. Research and industry partners would have to submit an application detailing their use case before being granted access to the enclave and would take legal liability for protecting the sensitive information.

Opening “training datasets” will be an important step toward encouraging scientists to open the data generated through AI research. The AI community should consider this topic as the field advances. This discussion should address developing open AI benchmarks, open learned representations, open learned parameters, and even open code when the research is publicly funded.



GOAL IV: HELP BUSINESSES AND ENTREPRENEURS USE GOVERNMENT DATA AS A RESOURCE

Many open government programs have bolstered growth in private industry, even spurring the development of entirely new companies, products, and services. The government should embrace opportunities to support businesses and entrepreneurs by reducing barriers to using data as a resource and directly engaging with industry to decrease the burden of submitting regulatory data.

RECOMMENDATION 23:

Identify cost barriers to accessing government data and eliminate fees to level the playing field for data users.

FIRST 100 DAYS**U.S. GOVERNMENT ACCOUNTABILITY OFFICE****ACTION PLAN:**

- The U.S. Government Accountability Office (GAO), with appropriate Congressional guidance, should conduct a 100-day review to identify current cost barriers to accessing structured government data and issue recommendations for reducing costs to zero or near-zero, beginning with the highest priority datasets.
 - GAO should seek input from private industry, journalists, nonprofits, and academics by distributing a common survey to identify current cost barriers and evaluate the benefits of making data available for free.
- Informed by the 100-day review, each agency should prioritize key datasets to open and commit to reducing costs for making datasets available in priority order. The White House Office of Science and Technology Policy should include those commitments in the Fourth Open Government National Action Plan in 2017.



Data should be available at no charge to be truly open. Since businesses' and citizens' taxes fund all government activities, including data gathering and management, it is fair to argue that the data has already been paid for by U.S. taxpayers. The proposed Open, Permanent, Electronic, and Necessary (OPEN) Government Data Act would legislate that principle.¹⁹¹ But currently, federal agencies can and often do charge for data, limiting its value to industry and inhibiting its use.

Individuals and organizations both in and outside of government often have to pay for essential government data, and the government sometimes charges a prohibitive rate. Panjiva, a company that provides data and analytics on international trade, pays \$100 per day for daily updates on U.S. Customs data, which it receives on CDs. Recently a reporter for Quartz, an online news site, requested U.S. immigration data and was told it would cost \$173,775.¹⁹² The government agency involved said that it had legal authority to charge for the data, according to the reporter.

Making government data open and accessible free of charge would reduce the barriers for entrepreneurs to start businesses and make small businesses more competitive. It would also enable transformative uses of government data that in turn generate business growth and tax revenues.¹⁹³ There are hundreds of documented American businesses whose work is directly supported by open data.¹⁹⁴ As an additional benefit, federal government agencies could end the counterproductive practice of buying data from each other, which incurs the administrative costs of transferring funds to the Treasury Department with no benefit to the agencies or taxpayers.

The Government Accountability Office (GAO) should conduct a 100-day review to identify current government datasets that carry significant access fees;

determine which of these datasets have the greatest value to businesses, journalists, and the research community; and develop recommendations for eliminating cost barriers for high-priority datasets. GAO should seek input from external stakeholders through a common survey, distributed to key groups and available on data.gov and other open government data websites. Informed by this review, each agency should develop a prioritized list of cost barriers and publicly commit to eliminating these charges in the Fourth Open Government National Action Plan in 2017.

The loss of government revenue from data fees should be greatly outweighed by the benefit of opening data for widespread use. Evidence shows that removing cost barriers can lead to explosive growth in the use of government data. When the U.S. Geological Survey in the Department of the Interior stopped charging for Landsat data, which provides unique spatial imagery and information on the Earth's composition, use of the data soared from 38 to over 5,700 scenes per day.¹⁹⁵ Users downloaded over 8 million scenes from 2008–2012 after the government made the data available for free.¹⁹⁶ A 2010 survey of over 2,500 people showed that approximately half of all Landsat data users were from the private sector (18 percent) or academia (33 percent).¹⁹⁷ Researchers have used Landsat data to monitor water quality, glacier recession, sea ice movement, invasive species encroachment, coral reef health, land use change, deforestation rates, damage from natural disasters, and population growth.¹⁹⁸

Some international evidence also supports the immediate benefit of removing cost barriers. In 2002, the Danish Government opened official Danish address data and made it available for free online. Eight years later, the government commissioned a study on the impact of the open address data and found direct financial benefits of EUR 62 million (\$69.2 million). Forgoing fees for the data cost the government only EUR 2 million (\$2.2 million).¹⁹⁹

**EVIDENCE
SHOWS THAT
REMOVING COST
BARRIERS
CAN LEAD TO
EXPLOSIVE
GROWTH IN
THE USE OF
GOVERNMENT
DATA.**

RECOMMENDATION 24:

Launch a Standard Business Reporting Program that will ultimately help businesses lower reporting burdens and costs.

FIRST 100 DAYS

THE PRESIDENT, THE OFFICE OF INFORMATION AND REGULATORY AFFAIRS,
& NATIONAL ECONOMIC COUNCIL

ACTION PLAN:

- Within the first 100 days of the new administration, the President should issue an Executive Order setting a goal of Standard Business Reporting (SBR) across the federal government as a means of streamlining regulatory reporting requirements to reduce the burden on industry. SBR will also ease administrative and compliance oversight by the government while increasing transparency. The President should direct the Office of Information and Regulatory Affairs (OIRA) to lead the SBR program, in partnership with the National Economic Council (NEC). The Executive Order should:
 - Direct OIRA and the NEC to develop an SBR roadmap for the federal government by mid-2017 that will lay out a plan for implementation and adoption, including conducting a cost-benefit analysis that evaluates the potential savings to government and industry; developing an initial taxonomy; testing the taxonomy; piloting the program; and setting a date to begin voluntary submissions through SBR.
 - Industry and software companies should play a leading role in developing the taxonomy and refining the program through all stages of implementation.
 - Task the NEC with convening an SBR Council composed of federal agencies, industry representatives, and the software industry that will support development of the roadmap and implementation of the broader SBR program. The Council should work directly with industry leaders in auditing, accounting, and human resources.
 - Request OIRA to develop biannual progress reports to keep the public informed of SBR's progress.

The federal government has a responsibility to balance the need for appropriate business regulation with the potential burden on businesses to comply. While the cost of regulatory reporting is difficult to gauge, it is significant by any estimate. A 2014 Office of Management and Budget Congressional Report calculated the annual cost of federal rules and paperwork at \$70.2 to \$104.7 billion,²⁰⁰ while the Small Business Administration (SBA) determined the cost at close to \$2 trillion in 2010.²⁰¹ The SBA also found that regulations have a disproportionately large impact on small businesses, which spend 36 percent more per employee than larger firms.²⁰²

Concern about regulatory burden led both Presidents Clinton and Obama to issue executive orders placing reasonable constraints on developing new government regulations. President Clinton's Executive Order 12866, signed in 1993, requires that agencies submit regulatory actions that have an impact of more than \$100 million to the Office of Information and Regulatory Affairs (OIRA) for review.²⁰³ In 2011 President Obama signed Executive Order 13563, which directs agencies to analyze and "modify, streamline, expand, or repeal" existing rules "that may be outmoded, ineffective, insufficient, or excessively burdensome."²⁰⁴

While finding the right level of regulation for any industry is a complex calculation, the effort required for regulatory reporting compliance is to some extent a data problem. Using shared, standardized, and open data, the federal government can make it easier for businesses to collect and submit regulatory data across the board.

Standard Business Reporting (SBR) is a proven approach to reducing businesses' reporting burden. When fully implemented, SBR will enable companies to realize business efficiencies by automatically compiling reporting requirements through regular business functions. For example, the software used for payroll will automatically calculate tax obligations and package the data in a format that can be submitted directly to the government. SBR creates a standard approach across government that can lead to additional efficiencies, such as enabling companies to provide information only once and allowing multiple government agencies to access the same data. This allows all companies, regardless of size, to streamline their work processes and records management.

On the government side, as business data becomes more comparable, it becomes easier for government regulators to ensure

transparency and compliance. Regulatory agencies can compare results across offices, helping keep pace with the complexity of the modern economy and predict risks with greater accuracy. Some legislators are already promoting the benefits of this kind of approach: The proposed Financial Transparency Act²⁰⁵ would follow many of these same principles for the financial industry.

The President should issue an Executive Order to launch SBR and set an ambitious agenda requesting a roadmap and implementation plan by mid-2017. The roadmap should clearly lay out the quantified benefits to the American economy, major milestones for developing and implementing an SBR program, and key roles and responsibilities inside and outside of government. It must allow flexibility to support an agile approach to prioritization, decision making, and exploring alternative solutions. The implementation plan should include an SBR Council, composed of public and private sector leaders, that will examine best practices from SBR implementation around the world and lean on the software industry to detail lessons learned from previous efforts abroad. The Executive Order should require OIRA to submit a public progress report every six months to ensure transparency and that the government stays on track to roll out SBR standards for businesses to use voluntarily in the near future.

Several countries have already demonstrated the value of SBR. The Netherlands Government pioneered SBR with the adoption of the Nederlandse Taxonomie Project (or Dutch Taxonomy Project) in 2004. The Netherlands has since expanded the project with full SBR as the exclusive channel for corporate income tax filings and continued expansion into other reporting requirements, including pilot programs for educational institutions and housing corporations.²⁰⁶ Australia has implemented a major transition to SBR in recent years. Brazil and New Zealand are drawing on Australia's SBR program to develop a project for intra-government reporting, and Singapore is in the process of developing a business case for SBR.

The Australian experience also shows how to implement SBR efficiently and effectively. In Australia's [Standard Business Reporting program](#), the government is partnering with the software industry to automatically collect information during regular business activities, so companies can simply review the data and submit it to the government without additional data entry or calculations.²⁰⁷ Australia is now asking businesses to participate voluntarily, and will make SBR mandatory for many government forms beginning in 2018.²⁰⁸ The Australian government estimates that businesses saved over \$1 billion in the 2015–2016 fiscal year and projects nearly \$5 billion in cumulative savings by 2017–18.²⁰⁹

RECOMMENDATION 25:

Partner with the automated vehicle industry to give open data a central role in creating national safety standards.

FIRST 100 DAYS**U.S. DEPARTMENT OF TRANSPORTATION**

ACTION PLAN:

- The Department of Transportation National Highway Traffic Safety Administration (NHTSA) should support industry efforts to share automated vehicle data that could lead to safer vehicle development, such as testing data or challenging scenarios that have led to crashes or near misses.
 - In releasing templates and additional guidance for the Safety Assessments described in the Federal Automated Vehicles Policy over the next several months, the NHTSA should ensure these assessments are born digital. Rather than collecting PDFs or documents, NHTSA should collect key data fields in a format that enables automated analysis.
 - As the technology moves toward large-scale deployment, the NHTSA should partner with industry and the research community to work toward developing an open set of criteria for an “automated driver’s test” that will eventually enable automated vehicles to self-qualify for operating on public roads.
-

In early 2016, Secretary of Transportation Anthony Foxx declared that “we are on the cusp of a new era in automotive technology” as fully automated vehicles transform from a distant vision to a near-term reality.²¹⁰ At least 33 companies are researching and/or developing automated vehicle technologies,²¹¹ with some claiming they will deploy automated fleets broadly in the next three to five years.²¹² A small number of automated vehicles are already on the road (with a human at the wheel to seize control if necessary) in California, Texas, Washington, Arizona,²¹³ Pennsylvania,²¹⁴ and several large urban areas abroad. Automated vehicles can usher in a new era of transportation efficiency and safety while providing mobility to millions of additional Americans. Open data can help ensure that both the technology and the public are ready for the road ahead.

The Department of Transportation is already preparing for this shift. The National Highway Traffic Safety Administration (NHTSA) released the Federal Automated Vehicles Policy,²¹⁵ which provides safety guidance for the development of automated vehicles, in September 2016.²¹⁶ The policy offers vehicle performance guidance that gives industry a flexible approach to pursue the technology, summarizes the need for a consistent national framework for operating automated vehicles, outlines the current regulatory tools available to the NHTSA, and examines potential future tools and authorities. NHTSA is continuing to solicit feedback on the policy and plans to release additional guidance or updates in the future.²¹⁷

As NHTSA continues its research to examine the unique opportunities and risks provided by automated vehicles, such as cybersecurity and performance metric development, the administration should lay the foundation for a robust open data ecosystem to support the automated vehicle industry in the future. They can support this vision by:

- Fostering data sharing across the autonomous vehicle industry and government. The NHTSA Policy outlines a plan to establish a mechanism to “facilitate autonomous data sharing,” which is a critical next step to support safer vehicle development. The shared data should include testing environments and scenarios as well as data on accidents

or near-misses, which would help the field of developers to address known weaknesses in the technology. Several industry leaders are already voluntarily sharing information with the public, such as monthly reports, that include much of this information but not the datasets themselves.

- Ensuring automated vehicle system safety assessments are born digital. The NHTSA Policy already requests that manufacturers provide 15-point safety assessments as well as collect and share data on incidents, crashes, and near-misses. NHTSA plans to release templates for these assessments in the near future and eventually make them mandatory.²¹⁸ NHTSA should ensure these templates collect data in a structured and reusable format that is readily available for analysis and potential future sharing in machine-readable formats. Rather than collecting information in PDF or separate documents, NHTSA should embrace best practices in electronic data collection and replace forms and reports with data fields.

As the technology continues to develop, NHTSA should work with industry to develop a standard automated driver’s test based on open data that will enable future vehicles to self-qualify for market entry. Developing this standard in an open format will support a level playing field for industry, enable the research community to participate, and ensure the public has a voice in this epic shift in modern transportation. This role of public participation aligns with the Department of Transportation’s view that “larger questions [concerning automated vehicles] will require longer and more thorough dialogue with government, industry, academia and, most importantly, the public.”²¹⁹

NHTSA should then work with companies to open their autonomous driver’s test data, including videos and quantitative data. It would function similarly to open crash data from the Insurance Institute of Highway Safety.²²⁰ Opening this data would simultaneously make manufacturers accountable for ensuring vehicle safety, and would give the public confidence in the vehicles. Several surveys show a majority of Americans have safety concerns with automated vehicles.²²¹ By developing an automated driver test and publishing the results, NHTSA and the automotive industry could help make the public more receptive to automated vehicles’ safety benefits.

RECOMMENDATION 26:

Provide complete, accurate, and timely data on energy use in buildings to help companies save money by increasing energy efficiency.

FIRST YEAR**U.S. DEPARTMENT OF ENERGY**

ACTION PLAN:

- The Department of Energy's Energy Information Administration (EIA) and the Office of Energy Efficiency and Renewable Energy (EERE) should lead an effort to tap modern building energy management systems and emerging Internet of Things (IoT) technologies as new sources of data on energy usage and efficiency. EIA and EERE should complete a plan by January 2018 and implementation by December 2018.
 - The EERE should work with equipment suppliers and utilities to develop standard protocols for gathering information from building control systems. This public-private process should also set standards that ensure data from different building control systems is interoperable.
 - The EIA should convene industry leaders, including utilities, manufacturers of building control systems, IoT and data sensing leaders, and building owners' associations, to agree on opportunities and goals for collecting and using these new data sources. This might include incentives for building owners to gather and communicate detailed end use data in a standard format at standard intervals instead of funding conventional surveying methods.
 - When possible, EIA should publish this new data on building energy use as open data, together with the other information it makes available to the public.
-

Today, residential and commercial buildings' heating, cooling, and operational systems consume 40 percent of all domestic energy²²² and more than 75 percent of electricity.²²³ The primary sources of information about building energy use in the United States are the Commercial and Residential Building Energy Consumption Surveys.²²⁴ The Department of Energy (DOE) conducts these surveys every four years. The survey data does not directly include many types of energy consumption, data on lighting and plug loads, which require a labor-intensive collection and cleaning process.

The DOE Building Performance Database (BPD), the country's largest dataset on energy characteristics of individual commercial and residential buildings, includes actual utility use data and key characteristics from over 740,000 buildings.²²⁵ BPD allows users to identify and forecast the benefits of energy-saving opportunities. However, DOE currently relies on industry and governments to voluntarily provide energy use data to populate the database.²²⁶

New and emerging technologies provide an alternative that could present better, more timely data. Building owners are installing increasingly sophisticated control systems that provide advanced monitoring and can control the building's energy requirements and use.²²⁷ These systems, which may belong to the building or an energy utility, are a rapidly growing tool of the Internet of Things (IoT). They collect large amounts of data on how the building uses energy that building managers use to optimize energy performance and anticipate maintenance problems. However, the lack of interoperability standards for sensors, controls, and controllers has been a major barrier. This data would be of enormous value to help utilities, researchers, and policymakers could analyze to improve energy efficiency strategies nationwide.

The Energy Information Administration (EIA) and the Office of Energy Efficiency and Renewable Energy (EERE) should work with the private sector to tap building energy management systems as data sources for BPD. DOE should consider providing incentives for building owners to standardize and share this data. Through agreements with utilities, DOE could automate data collection from these

systems to support more sophisticated and timely analysis. The manufacturers of these systems could provide technical expertise, along with government centers, such as Lawrence Berkeley National Laboratory, which cleans the BPD data,²²⁸ and the Pacific Northwest National Laboratory, which has been researching the potential for using data from energy management systems.

Data from building energy management systems will become more valuable as they are more widely deployed. Many large buildings now have sophisticated energy management systems, and smaller buildings are expected to adopt similar but simpler control systems in the near future. But while the growth of this technology can provide a wealth of data, the data from different systems made by different manufacturers could be incompatible and not interoperable. Today, these manufacturers operate independently without coordinating their data structure or collection.

To address that problem, the EERE should develop standards for buildings' energy usage data. The EERE could use a public-private approach similar to the process that DOE used to develop data standards for Green Button, the program that provides data on energy usage back to individual consumers.

In parallel, emerging technologies around the IoT are helping to improve building energy efficiency and could dramatically increase available data, including commercial building utility data and residential utility data. Opportunities include smarter systems and better sensing of conditions that can inform more efficient energy management, such as localized weather information. EERE should partner with industry and other key players, such as the National Institute of Standards and Technology, to insure that energy-relevant IoT technologies can interoperate with building energy management systems and provide, where appropriate, real-time open data.

Developing new sources of energy data would have significant benefits. Drawing on data from building energy management systems would provide data in close to real time, with enough granularity to correlate changes in energy use with weather and other factors. By publishing this data as open data, the EIA would enable private industry, nonprofits, and academics to analyze the data and find new opportunities to increase energy efficiency.

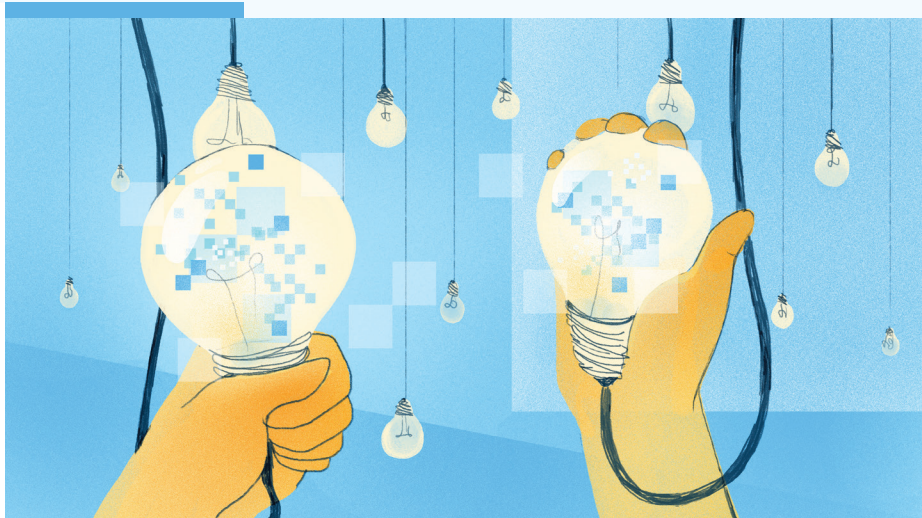
RECOMMENDATION 27:

Make it easier to discover and access government-owned intellectual property to help entrepreneurs build on this free resource.

FIRST YEAR**U.S. PATENT AND TRADEMARK OFFICE**

ACTION PLAN:

- The U.S. Patent and Trademark Office (USPTO) should develop a direct, open source online search tool for government-owned patents and any other patents available for use at no cost (e.g., expired patents) to drive entrepreneurs and businesses to take advantage of the resources that are available, but often not easily discoverable. Federal science agencies should assist in the identification of intellectual property of the federal government.
 - The USPTO should complete a plan for this search tool by June 2017, in consultation with other federal agencies and users of patent data, and implement it by January 2018.
 - The USPTO should continue to implement the roadmap for its Open Data Initiative, which includes several tools that could support this search functionality.
-



The federal government holds title to over 45,000 patents,^{229,230} which any business can use directly or adapt for a new invention. Except in limited circumstances, licensing government-owned patents is royalty free, revocable, and nonexclusive.²³¹ However, most entrepreneurs do not know about these resources and those who do commonly have trouble locating them.

Since U.S. patents are already accessible through an online database, the U.S. Patent and Trademark Office (USPTO) should now make it easier to search that database for government-owned patents in areas of interest to the user. In the United Kingdom, for example, the Intellectual Property Office makes it possible to sort patents by those that are Endorsed License of Right or Not in Force, categories that can be licensed at little or no cost.²³² A similar approach with sophisticated search functionality could greatly facilitate the use of government patents available free of charge. Additionally, providing easy access to information about expired patents would facilitate follow-on innovation.

The USPTO should continue to implement its Open Data Initiative roadmap and select one of their open data platforms to add this functionality. One option is [USPTO's PAIR Bulk Data \(beta\)](#), which allows users to download data in bulk form with an application programming interface (API), making it possible to load the data into databases or other analytical tools for research and analysis. APIs power a majority of mobile applications, many IT programs,

and create a market for the private sector to develop data-driven products and services. USPTO's Open Data Initiative includes plans to continue improving the discoverability, accessibility, and usability of public patent and trademark information through APIs. Another potential platform to provide this functionality is [PatentsView.org](#), a data visualization platform displaying over 40 years of patent data.²³³ PatentsView represented a significant step forward in opening and improving patent data, and could be improved further.

Although PAIR Bulk Data (beta) and PatentsView both include government patents, they do not yet provide a way to search for all government-owned patents or easily show patents that are available for free license, such as those that have expired. Improved search capabilities on one or both of these platforms would make it possible to easily find free government-owned intellectual property and to search by fields such as agriculture, renewable energy, or other areas.

USPTO has been experimenting with new, sophisticated approaches to improving search functionality that could be applied to help users find government-owned patents. The USPTO is a partner in the White House Cancer Moonshot, a billion-dollar initiative to develop new approaches to treating cancer.²³⁴ The Cancer Moonshot includes a public challenge to use data from the USPTO, the National Institutes of Health, and the U.S. Food and Drug Administration, to identify trends, build visualizations, and analyze factors such as cancer types, diagnosis methods, survival rates, clinical trials, and more.²³⁵ In order to facilitate this challenge, USPTO has developed new ways to search patents by key words and concepts that could help lead to strategies for identifying government-owned patents as well.

By reducing the barriers to locating government-owned intellectual property, USPTO will empower entrepreneurs, businesses, and inventors to capitalize on these resources.

CONCLUSION

Over the past ten years, the federal government has made tremendous strides in open data policy. Yet persistent challenges and a changing world present opportunities for open data to impact all sectors, including government, citizens, the research community, and businesses.

- In **government**, the Obama Administration has promoted open data and developed data-driven initiatives to make government more responsive and effective. However, small teams of individuals often drove these projects and many of these projects may lack the institutional support to become permanent government initiatives. The next administration must focus on advancing the open data government ecosystem to support and build on the gains already made. The administration should develop more expertise within government, data management structures, and cultural buy-in to continue to support open data initiatives.
- For **citizens**, open data has provided new information and services in health care, education, transportation, energy, finance, and other areas that affect their daily lives. Yet communities around the country continue to face significant challenges that can be improved through open data programs. The next administration should leverage open data as a resource to develop stronger, safer, and more equitable communities.
- Perhaps no group has more to gain from open data than the **research community**. By moving beyond access to publications to ensuring research data is open and re-usable, the community can more easily validate existing research, build off of one another's findings, and drive new discoveries. The next administration should embrace open research data as a valuable, effective means of supporting research activities across all disciplines.
- The **business community** has benefitted greatly from the expansion in available open government data. American entrepreneurs have developed new companies, programs, and organizations using the 185,000+ datasets now available on data.gov and other federal websites.²³⁶ Yet too many of the government's key data resources are not yet as accessible and usable as businesses need them to be. The next administration should continue to work with the business community to identify high-value data, determine how it needs to be improved, and make it widely available through public-private collaborations.

The promise of open data will only be realized if the federal government addresses these challenges and continues to advance and implement open data policy. The next administration has the opportunity to institutionalize open data as a critical component for achieving a wide range of policy and programmatic goals. This report and its recommendations are intended to be a useful resource to the federal government and stakeholders in federal data as these efforts move forward.



APPENDIX: ACRONYMS

AI	Artificial Intelligence
BPD	Building Performance Database
CBECs	Commercial Building Energy Consumption Survey
CDER	Common Data Element Repository
CDO	Chief Data Officer
CFPB	Consumer Financial Protection Bureau
CIO	Chief Information Officer
CMS	Centers for Medicare and Medicaid Services
COFAR	Council on Financial Assistance Reform
CTO	Chief Technology Officer
DATA Act	The Digital Accountability and Transparency Act of 2014
DOE	United States Department of Energy
DOL	United States Department of Labor
E-filing	Electronic filing
EERE	Energy Efficiency and Renewable Energy
EIA	Energy Information Administration
EITI	Extractive Industries Transparency Initiative
EPA	United States Environmental Protection Agency
FBI	Federal Bureau of Investigation
FDA	United States Food and Drug Administration
FOIA	Freedom of Information Act
FTC	Federal Trade Commission
GAO	United States Government Accountability Office
HHS	United States Department of Health and Human Services
HUD	United States Department of Housing and Urban Development
IARPA	Intelligence Advanced Research Projects Activity
IoT	Internet of Things
IRS	Internal Revenue Service
IT	Information Technology
NAP	National Action Plan
NEC	National Economic Council
NHTSA	National Highway Traffic Safety Administration
NIST	National Institute of Standards and Technology
NSF	National Science Foundation
NSTC	National Science and Technology Council
O*NET	Occupational Information Network
OIRA	Office of Information and Regulatory Affairs
OMB	United States Office of Management and Budget
OPEN Act	The Open, Permanent, Electronic, and Necessary Government Data Act
OSTP	White House Office of Science and Technology Policy
PDI	Police Data Initiative
PII	Personally Identifiable Information
PREP	Partnership for Resilience and Preparedness
RECS	Residential Energy Consumption Survey
SAMHSA	Substance Abuse and Mental Health Services Administration
SBA	Small Business Administration
SBR	Standard Business Reporting
STEM	Science, Technology, Engineering, and Math
USAID	United States Agency for International Development
USDA	United States Department of Agriculture
USPTO	United States Patent and Trademark Office

ENDNOTES

1. White House, "President's Memorandum on Transparency and Open 1 Government – Interagency Collaboration," www.whitehouse.gov/sites/default/files/omb/assets/memoranda_fy2009/m09-12.pdf (accessed 10/14/2016).
2. The Open Government Partnership, "Announcing New Open Government Initiatives," www.whitehouse.gov/sites/default/files/docs/new_nap_commitments_final.pdf (accessed 10/04/2016).
3. White House, "M-13-13 Memorandum for the Heads of Executive Departments and Agencies," www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf (accessed 10/04/2016).
4. The White House, "FACT SHEET: Data by the People, for the People – Eight Years of Progress Opening Government Data to Spur Innovation, Opportunity, & Economic Growth," www.whitehouse.gov/the-press-office/2016/09/28/fact-sheet-data-people-people-eight-years-progress-opening-government (accessed 10/11/2016).
5. Performance.gov, "Open Data," www.performance.gov/content/open-data (accessed 10/04/2016).
6. The Open Government Partnership, "Third Open Government National Action Plan for the United States of America," www.whitehouse.gov/sites/default/files/microsites/ostp/final_us_open_government_national_action_plan_3_0.pdf (accessed 10/04/2016).
7. The Open Government Partnership, "Announcing New Open Government Initiatives," www.whitehouse.gov/sites/default/files/docs/new_nap_commitments_final.pdf (accessed 10/04/2016).
8. Data Coalition, "The OPEN Government Data Act: A Sweeping Open Data Mandate for All Federal Information," www.datacoalition.org/the-open-government-data-act-a-sweeping-open-data-mandate-for-all-federal-information (accessed 10/04/2016).
9. Nick Sinai, "2015: Year of the Chief Data Officer," <https://medium.com/@ShorensteinCtr/2015-year-of-the-chief-data-officer-f4c4d5a4370f#bwsxjs115> (accessed 10/04/2016).
10. Cornell University Law School, "44 U.S. Code § 3504 – Authority and functions of Director," www.law.cornell.edu/uscode/text/44/3504 (accessed 10/04/2016).
11. IBM Center for the Business of Government, "The Emerging Role and Merits for Chief Data Officers," www.businessofgovernment.org/blog/business-government/emerging-role-and-merits-chief-data-officers (accessed 10/04/2016).
12. FedScoop, "Meet Federal Reserve Board's new chief data officer," fedscoop.com/meet-federal-reserveboards-new-chief-data-officer (accessed 10/04/2016).
13. Harvard Business Review, "Your C-Suite Needs a Chief Data Officer," hbr.org/2012/10/your-c-suite-needs-a-chief (accessed 10/04/2016).
14. U.S. General Accounting Office, "The Chief Financial Officers Act of 1990," www.gao.gov/special.pubs/af12194.pdf (accessed 10/04/2016).
15. Center for Open Data Enterprise, "Briefing Paper on Open Data and Privacy," www.opendataenterprise.org/reports/BriefingPaperonOpenDataandPrivacy.pdf (accessed 10/04/2016).
16. Linda Powell, "Building the Chief Data Officer Toolkit," www.slideshare.net/Chief_15/Data_Officer_Forum/linda-powell-presentation-at-the-chief-data-officer-forum-government (accessed 10/04/2016).
17. Nick Sinai, "2015: Year of the Chief Data Officer," <https://medium.com/@ShorensteinCtr/2015-year-of-the-chief-data-officer-f4c4d5a4370f#bwsxjs115> (accessed 10/04/2016).
18. National Archives, "Why Federal Agencies Need to Move Towards Electronic Recordkeeping," (accessed 10/04/2016).
19. White House, "Open Data Policy – Managing Information as an Asset," www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf (accessed 10/04/2016).
20. Bloomberg BNA, "Electronic Filing of Forms 94x Falls Short of IRS Goal, New System May Help Increase Rate," www.bna.com/electronic-filing-forms-n57982076287 (accessed 10/04/2016).
21. Electronic Tax Administration Advisory Committee, www.irs.gov/pub/irs-pdf/p3415.pdf (accessed 10/04/2016).
22. Bloomberg BNA, "Electronic Filing of Forms 94x Falls Short of IRS Goal, New System May Help Increase Rate," www.bna.com/electronic-filing-forms-n57982076287 (accessed 10/04/2016).
23. Ibid.
24. White House Office of Management and Budget, "Open data CAP goal FY 2016 Q2 progress update," www.performance.gov/downloadpdf?file=FY%2016%20Q2%20Open%20Data%20FINAL.pdf (accessed 10/04/2016).
25. California Student Aid Commission, "New Option for Submitting FAFSA," www.csac.ca.gov/pubs/forms/grnt_frm/FAFSAstory.pdf (accessed 10/04/2016).
26. United States Patent and Trademark Office, "File your Application Electronically to Avoid the Surcharge for Paper Filing," www.uspto.gov/custom-page/file-your-application-online-avoid-surcharge-paper-filing (accessed 10/04/2016).
27. Sunlight Foundation, "FCC Votes to Expand Transparency for Political Ads," (accessed 10/04/2016).
28. Politico, "How to Improve Workplace Safety," www.politico.com/agenda/story/2016/07/how-to-improve-workplace-safety-000167; OSHA Record Politico, "Keeping Rule Out Today," www.politico.com/tipsheets/morning-shift/2016/05/osharecord-keeping-rule-out-today-house-holds-overtime-hearing-uber-cuts-deal-with-machinists-214234 (accessed 10/04/2016).
29. U.S. Patent and Trademark Office, "File your Application Electronically to Avoid the Surcharge for Paper Filing," www.uspto.gov/custom-page/file-your-application-online-avoid-surcharge-paper-filing (accessed 10/04/2016).
30. U.S. Department of Treasury, "General Explanations of the Administration's Fiscal Year 2017 Revenue Proposals," www.treasury.gov/resource-center/tax-policy/Documents/General-Explanations-FY2017.pdf (accessed 10/04/2016).
31. Governing, "The Causes, Costs and Consequences of Bad Government Data," www.governing.com/topics/mgmt/gov-bad-data.html (accessed 10/04/2016). also, Governing, "Bad Data Infographic," media.navigatored.com/documents/BadData_infographic.pdf (accessed 10/04/2016).
32. White House, "Fact Sheet: Administration Announces New 'Smart Cities' 32 Initiative to Help Communities Tackle Local Challenges and Improve City Services," (accessed 10/04/2016).
33. U.S. Department of Justice, "Summary of Annual FOIA Reports for Fiscal Year 2015," www.justice.gov/oip/reports/fy_2015_annual_foia_report_summary/download (accessed 10/04/2016).
34. The FOIA Project, "Good News and Bad News on FOIA Responsiveness," foiaproject.org/2016/01/26/good-news-and-bad-news-on-foia-responsiveness (accessed 10/04/2016).
35. U.S. Department of Justice, "OIP Summary of the FOIA Improvement Act of 2016," www.justice.gov/oip/oip-summary-foia-improvement-act-2016 (accessed 10/04/2016).
36. White House, "Open Data Policy – Managing Information as an Asset", (accessed 10/04/2016).
37. FOIA Online, "Frequently Asked Questions," foiaonline.regulations.gov/foia/action/public/home/faqs (accessed 10/07/2016).
38. FOIA Online, "About FOIA Online," foiaonline.regulations.gov/foia/action/public/home/about (accessed 10/07/2016).
39. The federal FOIA Advisory Committee's Final Report for 2014-16 summarizes challenges for agencies in making proactive disclosures in electronic format, notably with documents that were not born digital, and creating useful FOIA logs See pp. 25-33.
40. Kwoka, Margaret B., FOIA, Inc. (November 2, 2015). 65 Duke Law Journal 1361 (2016); U Denver Legal Studies Research Paper No. 15-57. Available at SSRN: papers.ssrn.com/sol3/papers.cfm?abstract_id=2685402 (accessed 10/04/2016).
41. The Wall Street Journal, "Open-Government Laws Fuel Hedge-Fund Profits," www.wsj.com/articles/opengovernment-laws-fuel-hedgefund-profits-1379905298?tesla=y (accessed 10/04/2016).
42. Data Foundation, "The DATA Act: Vision & Value," www.datafoundation.org/data-act-vision-and-valuerreport (accessed 10/04/2016).
43. OpenBeta, "Open Beta USA Spending," (accessed 10/04/2016).
44. 18f, "How DATA Act implementation is opening up federal spending," 18f.gsa.gov/2015/06/09/data-act-data-act-explainer (accessed 10/04/2016).
45. Data Foundation, "The DATA Act: Vision & Value," www.datafoundation.org/data-act-vision-and-valuerreport (accessed 10/04/2016).
46. 18f, "How DATA Act Implementation is opening up Federal Spending," 18f.gsa.gov/2015/06/09/data-act-data-act-explainer (accessed 10/04/2016).
47. Catalog of Federal Domestic Assistance, "Home," www.cfa.gov (accessed 10/04/2016).
48. White House, "Analytical Perspectives Budget of the U.S. Government," www.whitehouse.gov/sites/default/files/omb/budget/fy2017/assets/spec.pdf (accessed 10/05/2016).
49. Data Foundation, "The DATA Act: Vision & Value," www.datafoundation.org/data-act-vision-and-valuerreport (accessed 10/04/2016).
50. Ibid.
51. White House, "OMB Circular A-133," www.whitehouse.gov/sites/default/files/omb/assets/a133/a133_revised_2007.pdf (accessed 10/04/2016).
52. Center for Open Data Enterprise, "Open Data Roundtables," opendataenterprise.org/convene (accessed 10/04/2016).
53. U.S. Government Accountability Office, "GAO search for 'data quality' results in 1,352 reports," www.gao.gov/search?q=%22data+quality%22&now_sort=score+desc&page_name=main&tab=Solr&rows=50 (accessed 10/04/2016).
54. U.S. Government Accountability Office, "Census Bureau Needs to Accelerate Efforts to

- Develop and Implement Data Quality Review Standards," www.gao.gov/products/GAO-05-86 (accessed 10/04/2016).
55. U.S. Government Accountability Office, "IRS' Actions To Improve the Accuracy of Non-Wage Income Data Are Vital," www.gao.gov/products/IMTEC-86-17 (accessed 10/04/2016).
 56. David Portnoy, "Intro to Demand Driven Open Data for Data Users," www.slideshare.net/DavidPortnoy/intro-to-demanddriven-open-data-for-data-users (accessed July 1, 2016).
 57. Wilson Center, "How do Smart Cities Talk to People?," www.wilsoncenter.org/article/how-do-smart-cities-talk-to-people (accessed 10/06/2016).
 58. USAID, "With a Little Help from the Crowd, USAID Increases Government Transparency," blog.usaid.gov/2012/06/with-a-little-help-from-the-crowd-usaid-increases-government-transparency (accessed 10/04/2016).
 59. MapMaker, "About," www.google.com/mapmaker/about/regionalleads (accessed 10/04/2016).
 60. Center for Open data Enterprise, "Briefing Paper on Open Data and Data Quality," www.opendataenterprise.org/reports/BriefingPaperonOpenDataandImprovingDataQuality.pdf (accessed 10/04/2016).
 61. U.S. Patent and Trademark Office, "PatentsView Inventor Disambiguation Technical Workshop," www.patentsview.org/workshop (accessed July 1, 2016).
 62. U.S. Patent and Trademark Office, "PatentsView," www.patentsview.org/web (accessed July 1, 2016).
 63. Ibid.
 64. <https://www.whitehouse.gov/the-press-office/2016/09/28/fact-sheet-data-people-people-eight-yearsprogress-opening-government>
 65. Feeding America, "Hunger and Poverty Facts and Statistics", www.feedingamerica.org/hunger-in-america/impact-of-hunger/hunger-and-poverty/hunger-and-poverty-fact-sheet.html (accessed 10/04/2016).
 66. Wiley Online Library, "Food Security, Poverty, and Human Development in the United States", onlinelibrary.wiley.com/doi/10.1196/annals.1425.001/full (accessed 10/04/2016).
 67. Ibid.
 68. U.S. Census Bureau, "Current Population Survey", www.census.gov/programs-surveys/cps.html (accessed 10/04/2016).
 69. Feeding America, "Food Insecurity in the United States", map.feedingamerica.org/country/2014/overall (accessed 10/04/2016).
 70. Capital Area Food Bank, "Bowser spikes joy into 71 the Moten community", www.capitalareafoodbank.org/2015/03/bowser-spikes-joy-into-the-moten-community (accessed 10/04/2016).
 71. Capital Area Food Bank, "A bus to get food where the need is greatest", www.capitalareafoodbank.org/2015/06/the-wheels-on-the-bus-go-round-and-round (accessed 10/04/2016).
 72. Prince George's County, "Transforming Neighborhoods Initiative", www.princegeorgescountymd.gov/1048/Transforming-Neighborhoods-Initiative-TN (accessed 10/04/2016).
 73. Waste Not OC Coalition, "About us", www.wastenotoc.org/#!about-us/coda (accessed 10/04/2016).
 74. Waste Not OC Coalition, "Standard Practices for Clinics", www.wastenotoc.org/identifying-foodinsecure-individuals (accessed 10/04/2016).
 75. BioCycle, "Coalition 'Feeds The Need' In California County", www.biocycle.net/2016/09/15/coalitionfeeds-need-california-county (accessed 10/04/2016).
 76. U.S. Department of Housing and Urban Department, "HUD Awards \$1 Billion Through National Disaster Resilience Competition," portal.hud.gov/hudportal/HUD?src=/press/press_releases_media_advisories/2016/HUDNo_16-006 (accessed 10/04/2016).
 77. National Oceanic and Atmospheric Administration, "New report finds human-caused climate change increased the severity of many extreme events in 2014," www.noaa.gov/stories/2015/110515-new-report-human-caused-climate-change-increased-the-severity-of-many-extreme-events-in-2014.html (accessed 10/05/2016).
 78. National Oceanic and Atmospheric Administration, "NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2016)," (accessed 10/04/2016).
 79. White House, "Fact Sheet: Launching New Public-Private Partnership and Announcing Joint Declaration on Leveraging Open Data for Climate Resilience," www.whitehouse.gov/the-press-office/2016/09/22/fact-sheet-launching-new-public-private-partnership-and-announcing-joint (accessed 10/04/2016).
 80. The Partnership for Resilience and Preparedness, "Sonoma County's Climate Resilience Dashboard", www.prepdata.org/dashboard/understanding-sonoma-countys-climate-adaptation-plan (accessed 10/04/2016).
 81. Washington Post, "FBI to sharply expand system for tracking fatal police shootings," www.washingtonpost.com/national/fbi-to-sharply-expand-system-for-tracking-fatal-police-shootings/2015/12/08/a60fbc16-9dd4-11e5-bce4-708fe33e3288_story.html (accessed 10/04/2016).
 82. U.S. Department of Justice, "Arrest-Related Deaths Program Assessment," www.bjs.gov/content/pub/pdf/ardpatr.pdf (accessed 10/04/2016).
 83. U.S. Department of Justice, "Assessment of Coverage in the Arrest-Related Deaths Program," www.bjs.gov/content/pub/pdf/acardp.pdf (accessed 10/04/2016).
 84. Washington Post, "People Shot Dead by Police in 2015," www.washingtonpost.com/graphics/national/police-shootings (accessed 10/12/2016).
 85. Guardian, "The Counted: People Killed by Police in the US," www.theguardian.com/us-news/nginteractive/2015/jun/01/the-counted-police-killings-us-database (accessed 10/12/2016).
 86. RAND Corporation, "Respect and Legitimacy: A Two-Way Street," www.rand.org/content/dam/rand/pubs/perspectives/PE100/PE154/RAND_PE154.pdf (accessed 10/04/2016).
 87. Washington Post, "FBI to sharply expand system for tracking fatal 87 police shootings," www.washingtonpost.com/national/fbi-to-sharply-expand-system-for-tracking-fatal-police-shootings/2015/12/08/a60fbc16-9dd4-11e5-bce4-708fe33e3288_story.html (accessed 10/04/2016).
 88. New York Times, "Justice Department to Streamline Tracking of Police Killings," www.nytimes.com/2016/08/10/us/politics/justice-department-to-streamline-tracking-of-police-killings.html (accessed 10/12/2016).
 89. Washington Post, "FBI to sharply expand system for tracking fatal police shootings," www.washingtonpost.com/national/fbi-to-sharply-expand-system-for-tracking-fatal-police-shootings/2015/12/08/a60fbc16-9dd4-11e5-bce4-708fe33e3288_story.html (accessed 10/04/2016).
 90. U.S. Department of Justice, "Requirement born of the Death in Custody Reporting Act of 2000. Assessment of Coverage in the Arrest-Related Deaths Program," www.bjs.gov/content/pub/pdf/acardp.pdf (accessed 10/04/2016).
 91. Washington Post, "FBI to sharply expand system for tracking fatal police shootings," www.washingtonpost.com/national/fbi-to-sharply-expand-system-for-tracking-fatal-police-shootings/2015/12/08/a60fbc16-9dd4-11e5-bce4-708fe33e3288_story.html (accessed 10/04/2016).
 92. U.S. Department of Justice, "Census of State and Local Law Enforcement Agencies, 2008," www.bjs.gov/content/pub/pdf/csllea08.pdf (accessed 10/04/2016).
 93. White House, "Fact Sheet: White House Police Data Initiative Highlights New Commitments," www.whitehouse.gov/the-press-office/2016/04/22/fact-sheet-white-house-police-data-initiative-highlights-new-commitments (accessed 10/04/2016).
 94. White House, "Fact Sheet: White House Police Data Initiative Highlights New Commitments," www.whitehouse.gov/the-press-office/2016/04/22/fact-sheet-white-house-police-data-initiative-highlights-new-commitments (accessed 10/04/2016).
 95. Bayes Impact, "Announcing our first product to bridge the divide between police and communities," www.bayesimpact.org/stories/?name=bridge-uof-launch (accessed 10/04/2016).
 96. U.S. Bureau of Labor Statistics, "Databases, Tables & Calculators by Subject," data.bls.gov/timeseries/LNS14000000 (accessed 10/04/2016).
 97. U.S. Bureau of Labor Statistics, "Employment Situation Summary," www.bls.gov/news.release/empst.nr0.htm (accessed 10/04/2016).
 98. U.S. Bureau of Labor Statistics, "Employment Situation of Veterans Summary," www.bls.gov/news.release/vet.nr0.htm (accessed 10/04/2016).
 99. U.S. Bureau of Labor Statistics, "Unemployed persons by duration of unemployment," www.bls.gov/news.release/empst.t12.htm (accessed 10/04/2016).
 100. U.S. Bureau of Labor Statistics, "Economic Situation Summary". www.bls.gov/news.release/empst.nr0.htm (accessed 10/04/2016).
 101. 18F, "O*NET today and beyond," 18f.gsa.gov/2015/03/11/onet-today-and-beyond (accessed 10/04/2016).
 102. O*NET Resource Center, "About," www.onetcenter.org/overview.html (accessed 10/04/2016).
 103. U.S. Department of Labor, "FY 2017 DOL Budget," www.dol.gov/sites/default/files/documents/general/budget/FY2017BIB_0.pdf (accessed 10/04/2016).
 104. The Center for Open Data Enterprise, "Open Data for the Labor Market," www.opendataenterprise.org/reports/DOLOpenDataRoundtableReport.pdf (accessed 10/04/2016).
 105. 18F, "O*NET today and beyond," 18f.gsa.gov/2015/03/11/onet-today-and-beyond (accessed 10/04/2016).
 106. Ibid.
 107. Tyrone Grandison, "Exploring the Skills Economy," www.tyronegrandison.org/blog/exploring-the-skills-economy (accessed 10/04/2016).

108. Indeed, "About," www.indeed.com/about (accessed 10/04/2016).
109. Venture Beat, "LinkedIn now has 400M users, but only 25% of them use it monthly," venturebeat.com/2015/10/29/linkedin-now-has-400m-users-but-only-25-of-them-use-it-monthly (accessed 10/04/2016).
110. U.S. Department of Labor, "FY 2017 DOL Budget," www.dol.gov/sites/default/files/documents/general/budget/FY2017BIB_0.pdf (accessed 10/04/2016).
111. Centers for Disease Control and Prevention, "Increases in Drug and Opioid Overdose Deaths —United States, 2000–2014," www.cdc.gov/mmwr/preview/mmwrhtml/mm6450a3.htm (accessed 10/04/2016).
112. Ibid.
113. American Society of Addiction Medicine, "Opioid Addiction," www.asam.org/docs/default-source/advocacy/opioid-addiction-disease-facts-figures.pdf (accessed 10/04/2016).
114. Wall Street Journal, "States Fight Opioid Epidemic With Prescription Databases," www.wsj.com/articles/states-fight-opioid-epidemic-with-prescription-databases-1472854536 (accessed 10/13/2016).
115. U.S. Department of Justice, "Department of Justice Releases Strategy Memo to Address Prescription Opioid and Heroin Epidemic," www.justice.gov/opa/pr/departement-justice-releases-strategy-memo-address-prescription-opioid-and-heroin-epidemic (accessed 10/13/2016).
116. Substance Abuse and Mental Health Services Administration, "Behavioral Health Treatments and Services," www.samhsa.gov/treatment (accessed 10/04/2016).
117. Centers for Medicare & Medicaid Services, "Hospital Compare Star Ratings Fact Sheet," www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2015-Fact-sheets-items/2015-04-16.html (accessed 10/04/2016).
118. PBS, "Chasing Heroin," www.pbs.org/wgbh/frontline/film/chasing-heroin (accessed 10/13/2016).
119. U.S. Environmental Protection Agency, "Drinking Water Contaminants - Standards and Regulations," www.epa.gov/dwstandardsregulations#primary (accessed 10/04/2016).
120. Centers for Disease Control and Prevention, "Public Water System FAQ," www.cdc.gov/healthywater/drinking/public/faq.html (accessed 10/04/2016).
121. Citizen Science, "Federal Crowdsourcing and Citizen Science Toolkit," <https://crowdsourcing-toolkit.sites.usa.gov> (accessed 10/04/2016).
122. U.S. National Library of Medicine National Institutes of Health, "Citizen Science 123 Initiatives: Engaging the Public and Demystifying Science," www.ncbi.nlm.nih.gov/pmc/articles/PMC4798796 (accessed 10/04/2016).
123. Nature, "Environmental science: Pollution patrol," www.nature.com/news/environmental-science/pollution-patrol-1.16654 (accessed 10/04/2016).
124. Air Quality Egg, "Air Quality Egg," airqualityegg.com (accessed 10/04/2016).
125. Noise Tube, "Turn your mobile phone into an environmental sensor and participate in the monitoring of noise pollution," noisetube.net/#&panel1-1 (accessed 10/04/2016).
126. University of Florida, "Florida LAKEWATCH: Citizen Scientists protecting Florida's aquatic systems," www.basinalliance.org/data/files/Hoyer%20et%20al%202014%20Citizen%20Science11.pdf (accessed 10/04/2016).
127. U.S. Environmental Protection Agency, "Polluted Runoff: NonPoint Source Pollution: Volunteer Monitoring," U.S. Environmental Protection Agency, U.S. Environmental Protection Agency, U.S. www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-volunteer-monitoring (accessed 10/04/2016).
128. U.S. Consumer Product Safety Commission, "Saferproducts.gov," www.saferproducts.gov/CPSRMSPublic/Incidents/ReportIncident.aspx (accessed 10/04/2016).
129. The Food and Drug Administration, "Questions and Answers on FDA's Adverse Event Reporting System (FAERS)," www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects (accessed 10/04/2016).
130. Safercar, "Keeping you Safe," www.odi.nhtsa.dot.gov/owners/SearchSafetyIssues (accessed 10/04/2016).
131. U.S. Department of Transportation, "DOT File a Consumer Complaint," www.transportation.gov/airconsumer/file-consumer-complaint (accessed 10/04/2016).
132. U.S. Department of Transportation, "Consumer Sentinel Network," www.ftc.gov/enforcement/consumer-sentinel-network (accessed 10/04/2016).
133. Federal Trade Commission, "FTC Releases Annual Summary of Consumer Complaints," www.ftc.gov/news-events/press-releases/2016/03/ftc-releases-annual-summary-consumer-complaints (accessed 10/04/2016).
134. Federal Communications Commission, "Consumer Help Center," <https://crowdsourcing-toolkit.sites.usa.gov> (accessed 10/04/2016).
135. Federal Communications Commission, "CFPB Submit a Complaint," www.consumerfinance.gov/complaint (accessed 10/04/2016).
136. Consumer Financial Protection Bureau, "Consumer Complaint Database," www.consumerfinance.gov/complaintdatabase (accessed 10/04/2016).
137. Social Science Research Network, "Skeletons in the Database: An Early Analysis of the CFPB's Consumer Complaints," papers.ssrn.com/sol3/papers.cfm?abstract_id=2295157 (accessed 10/04/2016).
138. American Banker, "Customers Are Now Banks' Greatest Regulatory Threat," www.americanbanker.com/issues/178_176/customers-are-now-banks-greatest-regulatory-threat-1061975-1.html (accessed 10/04/2016).
139. Consumer Financial Protection Bureau, "Consumer 140 Complaints Database," www.consumerfinance.gov/complaintdatabase (accessed 10/04/2016).
140. Brookings, "The Hidden STEM Economy," www.brookings.edu/research/the-hidden-stem-economy (accessed 10/04/2016).
141. Occupational Outlook Quarterly, "STEM 101: Intro to tomorrow's jobs," <http://www.stemedcoalition.org/wp-content/uploads/2010/05/BLS-STEM-Jobs-report-spring-2014.pdf> (accessed 10/04/2016).
142. Forbes, "The Supply And Demand Of Data Scientists: What The Surveys Say," www.forbes.com/sites/gilpress/2015/04/30/the-supply-and-demand-of-data-scientists-what-the-surveys-say/#79606d83205e (accessed 10/04/2016).
143. K-12 Business Sustainable Development, tuvalabs.com (accessed 10/04/2016).
144. U.S. Department of Justice, "The President's Taskforce on 21st Century Police Reporting," www.cops.usdoj.gov/pdf/taskforce/TaskForce_Annual_Report.pdf p. 10 (accessed 10/04/2016).
145. U.S. Department of Agriculture, "USDA Participated in a 'First of Its Kind' Camp for DC-Area Teens Focused Specifically on Open Data and Agriculture," blogs.usda.gov/2015/08/18/usda-participated-in-a-first-of-its-kind-camp-for-dc-area-teens-focused-specifically-on-open-data-and-agriculture/#more-60086 (accessed 10/04/2016).
146. New York University, "Open Data Kids Camp," engineering.nyu.edu/news/2016/08/04/open-data-kidscamp (accessed 10/04/2016).
147. FedScoop, "USDA hopes more donors support open data summer camp for teens," fedscoop.com/usda-hopes-more-donors-support-open-data-summer-camp-for-teens (accessed 10/04/2016).
148. LinkedIn, "USDA's Open Data STEAM Summer Camp," www.linkedin.com/pulse/usdas-open-datasteam-summer-camp-joyce-hunter (accessed 10/04/2016).
149. Turner Center, "Housing Highlights from the 2015 American Community Survey," turnercenter.berkeley.edu/blog/housing-highlights-from-the-2015-american-community-survey (accessed 10/04/2016).
150. Harvard University, "The State of the Nation's Housing 2016," jchs.harvard.edu/sites/jchs.harvard.edu/files/jchs_2016_state_of_the_nations_housing_lowres.pdf (accessed 10/04/2016).
151. Center on Budget and Policy Priorities, "Policy Basics: The Housing Choice Voucher Program," www.cbpp.org/research/housing/policy-basics-the-housing-choice-voucher-program (accessed 10/04/2016).
152. Center on Budget and Policy Priorities, "Research Shows Housing Vouchers Reduce Hardship and Provide Platform for Long-Term Gains Among Children," www.cbpp.org/research/housing/researchshows-housing-vouchers-reduce-hardship-and-provide-platform-for-long-term?fa=view&id=4098#_ftn2 (accessed 10/04/2016).
153. Center on Budget and Policy Priorities, "Policy Basics: The Housing Choice Voucher Program," www.cbpp.org/research/housing/policy-basics-the-housing-choice-voucher-program (accessed 10/04/2016).
154. U.S. Department of Housing and Urban Development, "Housing Choice Vouchers Fact Sheet," portal.hud.gov/hudportal/HUD?src=/program_offices/public_indian_housing/programs/hcv/about/fact_sheet (accessed 10/04/2016).
155. Affordable Housing Online, "Section 8 Wait Lists," affordablehousingonline.com/open-section-8-waiting-lists#about (accessed 10/04/2016).
156. U.S. Department of Housing and Urban Development, "PHA Contact Information," portal.hud.gov/hudportal/HUD?src=/program_offices/public_indian_housing/pha/contacts (accessed 10/04/2016).
157. The Center on Budget and Policy Priorities, "Policy Basics: The Housing Choice Voucher Program," www.cbpp.org/research/housing/policy-basics-the-housing-choice-voucher-program (accessed 10/04/2016).
158. Office of Science and Technology Policy, "Letter in fulfillment of the requirement in Title III of the Joint Explanatory Statement for the 2014 Omnibus for the Office of Science and Technology Policy (OSTP) to report to the Committees on progress in developing and implementing policies on increasing public access to the results of federally funded scientific research," https://www.whitehouse.gov/sites/default/files/microsites/ostp/public_access_report_to_congress_ostp_11.13.14.pdf (accessed 10/19/2016).

159. OECD, "Gross Domestic Spending on R&D," data.oecd.org/rd/gross-domestic-spending-on-r-d.htm (accessed 10/11/2016).
160. American Association for the Advancement of Science, "Trends in Federal R&D, FY1976-2017," www.aaas.org/sites/default/files/DefNon%3B.jpg (accessed 10/11/2016).
161. SPARC, "From Ideas to Industries: Human Genome Project," sparcopen.org/impact-story/human-genome-project (accessed 10/11/2016).
162. National Human Genome Research Institute, "Calculating the economic impact of the Human Genome Project," www.genome.gov/27544383/calculating-the-economic-impact-of-the-human-genomeproject (accessed 10/13/2016).
163. SPARC, "Battling Disease with Open: Open Source Malaria Consortium," sparcopen.org/impact-story/open-source-malaria-consortium (accessed 10/12/2016).
164. Dronecode, "About the Project," www.dronecode.org/about (accessed 10/12/2016).
165. Sloan Digital Sky Survey, "Science Results," www.sdss.org/science (accessed 10/12/2016).
166. White House, "Increasing Access to the Results of Federally Funded Scientific Research," www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf (accessed 10/04/2016).
167. Office of Science and Technology, "July 2016 Progress in Developing, Finalizing, and Implementing a Plan to Increase Public Access to the Results of Federally-funded Scientific Research," www.whitehouse.gov/sites/default/files/public_access_-_report_to_congress_-_jul2016_.pdf (accessed 10/11/2016).
168. White House, "Executive Order - Establishment of the Federal Privacy Council," www.whitehouse.gov/the-press-office/2016/02/09/executive-order-establishment-federal-privacy-council (accessed 10/13/2016).
169. White House, "Increasing Access to the Results of Federally Funded Scientific Research," www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf (accessed 10/04/2016).
170. NSF, "Research.gov," www.research.172.gov (accessed 10/04/2016).
171. Congressional Research Service, "Overview and Issues for Implementation of the Federal Cloud Computing Initiative: Implications for Federal Information Technology Reform Management," www.fas.org/sgp/crs/misc/R42887.pdf (accessed 10/12/2016).
172. Government Accountability Office, "Cloud Computing: Agencies Need to Incorporate Key Practices to Ensure Effective Performance," www.gao.gov/assets/680/676395.pdf (accessed 10/12/2016).
173. National Oceanic and Atmospheric Administration, "NEXRAD," www.ncdc.noaa.gov/data-access/radar-data/nexrad (accessed 10/12/2016).
174. U.S. Department of Agriculture, "NAIP," www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery (accessed 10/12/2016).
175. Amazon Web Services, "Landsat on AWS," aws.amazon.com/public-data-sets/landsat (accessed 10/12/2016).
176. U.S. Department of Commerce, "U.S. Secretary of Commerce Penny Pritzker Announces New Collaboration to Unleash the Power of NOAA's Data," www.commerce.gov/news/press-releases/2015/04/us-secretary-commerce-penny-pritzker-announces-new-collaboration-unleash (accessed 10/13/2016).
177. Google, "Genomic Data Processing on Google Cloud Platform," research.googleblog.com/2016/04/genomic-data-processing-on-google-cloud.html (accessed 10/12/2016).
178. The Scholarly Kitchen, "To Share or not to Share? That is the (Research Data) Question...," scholarlykitchen.sspnet.org/2014/11/11/to-share-or-not-to-share-that-is-the-research-data-question (accessed 10/11/2016).
179. PeerJ, "Data reuse and the open data citation advantage," peerj.com/articles/175 (accessed 10/11/2016).
180. White House, "FACT SHEET: United States Hosts First-Ever Arctic Science Ministerial to Advance International Research Efforts," www.whitehouse.gov/the-press-office/2016/09/28/fact-sheet-united-states-hosts-first-ever-arctic-science-ministerial (accessed 10/12/2016).
181. White House, "Joint Statement of Ministers," www.whitehouse.gov/the-press-office/2016/09/28/joint-statement-ministers (accessed 10/12/2016).
182. White House, "The Open Government Partnership: Announcing New Open Government Initiatives," www.whitehouse.gov/sites/default/files/docs/new_nap_commitments_final.pdf (accessed 10/13/2016).
183. World Health Organization, "Developing Global Norms for Sharing Data and Results during Public Health Emergencies," www.who.int/medicines/ebola-treatment/data-sharing_phe/en (accessed 10/12/2016).
184. U.S. Department of Health and Human Services Center for Disease Control, "Zika Data Repository," github.com/cdcepi/zika (accessed 10/12/2016).
185. NPR, "Tech Giants Team Up To Tackle The Ethics Of Artificial Intelligence," www.npr.org/sections/alltechconsidered/2016/09/28/495812849/tech-giants-team-up-to-tackle-the-ethics-of-artificial-intelligence (accessed 10/04/2016).
186. Stanford, "At Stanford, experts explore artificial intelligence's social benefits," news.stanford.edu/2016/06/23/stanford-experts-explore-artificial-intelligences-social-benefits (accessed 10/04/2016).
187. White House, "Preparing for the Future of Artificial Intelligence," www.whitehouse.gov/blog/2016/05/03/preparing-future-artificial-intelligence (accessed 10/04/2016).
188. Economist, "Machine Learning Of prediction and policy," www.economist.com/news/finance-and-economics/21705329-governments-have-much-gain-applying-algorithms-public-policy (accessed 10/04/2016).
189. ImageNet, "Summary and Statistics," image-net.org/about-stats (accessed 10/04/2016).
190. White House, "Request for Information: Preparing for the Future of Artificial Intelligence," www.whitehouse.gov/webform/rfi-preparing-future-artificial-intelligence?cm_mc_uid=99468494561914749857488&cm_mc_s (accessed 10/13/2016).
191. Congress, "S.2852 - OPEN Government Data Act," www.congress.gov/bill/114th-congress/senate-bill/2852/text (accessed 10/04/2016).
192. Quartz, "I'm suing the U.S. government for its data on who's entering the country," qz.com/685956/imsuing-the-us-government-for-its-data-on-whos-entering-the-country (accessed 10/04/2016).
193. Sunlight Foundation, "10 Principles for Opening Up Government Information," sunlightfoundation.com/policy/documents/ten-open-data-principles (accessed 10/04/2016).
194. Center for Open Data Enterprise, "Open Data Impact Map," opendataenterprise.org/map.html (accessed 09/02/2016).
195. National Geospatial Advisory Committee, "Landsat Advisory Group Statement on Landsat Data Use and Charges," www.fgdc.gov/ngac/meetings/september-2012/ngac-landsat-cost-recovery-paper-FINAL.pdf (accessed 10/04/2016).
196. Ibid.
197. USGS, "The Users, Uses, and Value of Landsat and Other Moderate-Resolution Satellite Imagery in the United States—Executive Report," pubs.usgs.gov/of/2011/1031/pdf/OF11-1031.pdf (accessed 10/04/2016).
198. Landsat Science, "Landsat Science Data," landsat.gsfc.nasa.gov/?page_id=9 (accessed 10/04/2016).
199. Danish Enterprise and Construction Authority, "The value of Danish address data," www.adresseinfo.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf (accessed 10/04/2016).
200. Office of Management and Budget, "2015 Draft Report to Congress on the Benefits and Costs of Federal Regulations and Agency Compliance with the Unfunded Mandates Reform Act," www.whitehouse.gov/sites/default/files/omb/inforeg/2015_cb/draft_2015_cost_benefit_report.pdf (accessed 10/04/2016).
201. Small Business Administration Office of Advocacy, "The Impact of Regulatory Costs on Small Firms," [www.sba.gov/sites/default/files/The%20Impact%20of%20Regulatory%20Costs%20on%20Small%20Firms%20\(Full\).pdf](http://www.sba.gov/sites/default/files/The%20Impact%20of%20Regulatory%20Costs%20on%20Small%20Firms%20(Full).pdf) (accessed 10/04/2016).
202. Ibid.
203. U.S. Environmental Protection Agency, "Summary of Executive Order 12866 - Regulatory Planning and Review," www.epa.gov/laws-regulations/summary-executive-order-12866-regulatory-planning-and-review (accessed 10/04/2016).
204. White House, "Executive Order 13563 - Improving Regulation and Regulatory Review," www.whitehouse.gov/the-press-office/2011/01/18/executive-order-13563-improving-regulation-and-regulatory-review (accessed 10/04/2016).
205. Data Coalition, "The Financial Transparency Act," <http://www.datacoalition.org/issues/financial-transparency-act> (accessed 10/14/2016).
206. Banken, "What is the Difference between SBR and XBRL?," sbrbanken.nl/wp-content/uploads/2016/01/SBR-English-presentation.pdf (accessed 10/12/2016).
207. SBR, "Standard Business Reporting," www.sbr.gov.au (accessed 10/04/2016).
208. MYOB, "Cut your payroll processing time," www.myob.com/au/blog/cut-your-payroll-processing-time (accessed 10/04/2016).
209. Deloitte, "ABR Program Savings Review."
210. U.S. Department of Transportation, "Secretary Foxx Unveils President Obama's FY17 Budget Proposal of Nearly \$4 Billion for Automated Vehicles and Announces DOT Initiatives to Accelerate Vehicle Safety Innovations," www.transportation.gov/briefing-room/secretary-foxx-unveils-president-obama%E2%80%99s-fy17-budget-proposal-nearly-4-billion (accessed 10/04/2016).
211. CBS, "33 Corporations Working On Autonomous Vehicles," www.cbinsights.com/blog/autonomous-driverless-vehicles-corporations-list (accessed 10/04/2016).
212. Driverless Car Market Watch, "Autonomous car forecasts," www.driverless-future.com/?page_id=384 (accessed 10/04/2016).
213. Google, "Google Self-Driving Car Project Monthly Report," static.googleusercontent.com/media/www.google.com/en/selfdrivingcar/files/reports/report-0816.pdf (accessed 10/04/2016).

214. Reuters, "Uber debuts self-driving vehicles in landmark Pittsburgh trial," www.reuters.com/article/usuber-autonomous-idUSKCN11K12Y (accessed 10/04/2016).
215. U.S. Department of Transportation, "Federal Automated Vehicles Policy," www.transportation.gov/sites/dot.gov/files/docs/AV%20policy%20guidance%20PDF.pdf (accessed 10/04/2016).
216. White House, "FACT SHEET: Encouraging the Safe and Responsible Deployment of Automated Vehicles," www.whitehouse.gov/the-press-office/2016/09/19/fact-sheet-encouraging-safe-and-responsible-deployment-automated (accessed 10/04/2016).
217. U.S. Department of Transportation, "Federal Automated Vehicles Policy," www.transportation.gov/sites/dot.gov/files/docs/AV%20policy%20guidance%20PDF.pdf (accessed 10/04/2016).
218. Ibid.
219. Ibid.
220. White House, "2015 Traffic Fatalities Data Has Just Been Released: A Call to Action to Download and Analyze," www.whitehouse.gov/blog/2016/08/29/2015-traffic-fatalities-data-has-just-been-released-call-action-download-and-analyze (accessed 10/13/2016).
221. Morning Consult, "Voters Aren't Ready for Driverless Cars, Poll Shows," morningconsult.com/2016/02/08/voters-arent-ready-for-driverless-cars-poll-shows (accessed 10/04/2016). and Product, "Survey Shows Amount of Trust in Autonomous Vehicles," www.pddnet.com/news/2015/10/survey-shows-amount-trust-autonomous-vehicles (accessed 10/04/2016).
222. U.S. Energy Information Administration, "Frequently Asked Questions," www.eia.gov/tools/faqs/faq.cfm?id=86&t=1 (accessed 10/04/2016).
223. Department of Energy, "Chapter 5 – Increasing Efficiency of Building Systems and Technologies," energy.gov/under-secretary-science-and-energy/downloads/chapter-5-increasing-efficiency-buildings-systems-and (accessed 10/13/2016).
224. Department of Energy, "About CBECS," www.eia.gov/consumption/commercial/about.cfm (accessed 10/13/2016). and Department of Energy, "About the RECS," www.eia.gov/consumption/residential/about.cfm (accessed 10/13/2016).
225. U.S. Department of Energy, "BPD Overview," energy.gov/sites/prod/files/2014/06/f17/bpd_overview_2014.pdf (accessed 10/04/2016).
226. U.S. Department of Energy, "Frequently Asked Questions About the Building Performance Database," energy.gov/eere/buildings/frequently-asked-questions-about-building-performance-database#sources (accessed 10/04/2016).
227. Climate Tech Wiki, "Building Energy Management Systems (BEMS)," www.climatechwiki.org/technology/jiqweb-bems (accessed 10/04/2016).
228. U.S. Department of Energy, "Contributing Data," energy.gov/eere/buildings/contributing-data
229. U.S. Patent and Trademark Office, "US Government Patents," www.uspto.gov/web/offices/ac/ido/oeip/taf/govt/total_counts/govt_ct_list.htm (accessed 10/04/2016).
230. The Minerals, Metals & Materials Society, "The Federal Funding of R&D: Who Gets the Patent Rights?," www.tms.org/pubs/journals/JOM/matters/matters-9004.html (accessed 10/04/2016).
231. Cornell University Law School, "US Code," www.law.cornell.edu/cfr/text/34/6.3 (accessed 10/04/2016).
232. Intellectual Property Office, "Patents Endorsed License of Right (LOR) and Patents Not in Force (NIF)," www.ipo.gov.uk/p-di-notinforce.htm?filter=&sort=LOR+Start+Date&perPage=10 (accessed 10/04/2016).
233. U.S. Patent and Trademark Office, "New USPTO Tool Allows Exploration of 40 Years of Patent Data," www.uspto.gov/about-us/news-updates/new-uspto-tool-allows-exploration-40-years-patent-data (accessed 10/04/2016).
234. The White House, "FACT SHEET: Investing in the National Cancer Moonshot," <https://www.whitehouse.gov/the-press-office/2016/02/01/fact-sheet-investing-national-cancer-moonshot> (accessed 10/19/2016).
235. Cancer Moonshot Medium, "Unlocking Patent Data to Spur Cancer Breakthroughs," <https://medium.com/cancer-moonshot/unlocking-patent-data-to-spur-cancer-breakthroughs-26325501e9c2#.325ce3s73>.
236. Data.gov, "The Home of the U.S. Government's Open Data," www.data.gov (accessed 10/07/2016).



THE CENTER FOR
OPEN DATA ENTERPRISE