

Cross Agency Priority Goal

Quarterly Progress Update

Open Data

Goal leaders:

Tony Scott, Chief Information Officer, Office of Management and Budget

Megan Smith, Chief Technology Officer, Office of Science and
Technology Policy



FY2015 Quarter 2

Goal Statement

- Fuel entrepreneurship and innovation, and improve government efficiency and effectiveness by unlocking the value of government data and adopting management approaches that promote interoperability and openness of these data.

Urgency

- Freely available data from the U.S. Government is an important national resource, serving as fuel for entrepreneurship, innovation, scientific discovery, and economic growth. Making information about government operations more readily available and useful is also core to the promise of a more efficient and transparent government.
- There is more than \$1 trillion dollars of annual potential economic benefit from open data in the U.S., according to McKinsey*.

Vision

- Make open and machine-readable the new default for all government information and strengthen the open data ecosystem to fuel innovation and economic growth.

Progress Update

Policy

- **New Requirement to Include Non-Public Dataset Information in Public Data Listings.** As discussed on [GitHub](#), as of May 31, 2015, agencies are required to include information about all “non-public” data assets in their Public Data Listing (PDL) metadata. If necessary, agencies will work with their Freedom of Information Act (FOIA) offices to redact specific metadata if the text is subject to a FOIA exemption. Guidance for redactions is on [Project Open Data](#). Essentially, this reframes the PDL to be a public version of the full Enterprise Data Inventory (EDI). This policy requirement is a follow-up to the [Sunlight Foundation](#) FOIA request that was reported in last quarter’s update. The White House and Data.gov team appreciate agencies’ cooperation in this new guidance, which we believe further makes the U.S. the most transparent government in the world.
- **New Requirement to Provide License Information.** As part of its quarterly Integrated Data Collection (IDC) , the White House instructed agencies to ensure their PDL contains license or non-license (i.e. Public Domain) information for each dataset. [Guidance](#) and [examples](#) for agencies are provided on Project Open Data. This helps to ensure the public is aware there are no restrictions on copying, publishing, distributing, transmitting, adapting, or otherwise using the information for non-commercial or for commercial purposes. To ensure that it is clear that datasets are available for unrestricted access both domestically and internationally, we recommend that agencies use the “Creative Commons Zero” license discussed on [Project Open Data](#).
- **Project Open Data Dashboard Reviews.** As is done each quarter, the White House Office of Management and Budget (OMB) has evaluated each agency on their Open Data Policy progress on the [Project Open Data Dashboard](#). The White House applauds the Departments of Agriculture (USDA), Interior (DOI), and Veterans Affairs (VA), and the Social Security Administration (SSA), for achieving “green” (highest level possible) scores for all performance indicators for the December through February milestone period.

Tools & Collaboration

- On March 17th, the White House hosted an [Open Government Workshop](#) during [Sunshine Week](#). The workshop was the kickoff event for planning the third U.S. Open Government National Action Plan (second National Action Plan available [here](#)), which contains commitments that increase public integrity, enhance public access to information, improve management of public resources, and give the public a more active voice in the U.S. Government’s policymaking process. The event served as a brainstorming workshop for government and civil society groups to discuss possible commitments for the new plan on topics including open data, public participation, and open innovation.
- The Department of Commerce (DOC), one of the key producers of government data, recently formed the [Commerce Data Advisory Council](#) (CDAC). CDAC’s goals are to provide guidance to agency leaders on how to best use DOC data to enable businesses to be more competitive, governments smarter, and citizens more informed – ultimately increasing the value of the DOC’s data. CDAC is comprised of 19 of the industry’s biggest [data leaders](#) , including the co-founder of LinkedIn, the CEOs of ESRI and Socrata, and the Executive Director of Code for America, among other talented members. The White House applauds this innovative approach to data management.

Progress Update

Fueling the Ecosystem

The White House and the agencies continue to catalyze action, highlight success stories, and get important feedback through a series of events and well-designed prizes. Efforts include:

- **Datajams:** Brainstorming workshops with innovators and data owners.
- **Datapaloozas:** High-profile summits to highlight open data progress.
- **Hackathons:** 24-72 hour code-a-thons to build working prototypes.
 - On March 30th, the White House Office of Science and Technology Policy ([OSTP](#)), the General Services Administration (GSA)'s [18F](#), and several civil society groups co-hosted an [Accessibility Hackathon](#) to help improve the accessibility of federal websites for individuals with disabilities. Accessibility users, developers, policy makers, and technologists from both inside and outside of government came together to develop shareable solutions to digital accessibility problems facing government and users today.
 - On April 11-12, USDA, DOI and the [Recreation.gov](#) team organized the [myAmerica Hackathon](#). Over 70 participants attended the event from a variety of backgrounds. Eleven teams presented apps for recreation, including trails tools that pull in Flickr photos, consolidated permitting information, and games. The event was covered by [FedScoop](#).
 - GSA hosted a [hackathon](#) on May 8th to bring together over 60 policymakers, developers, and data scientists to help build solutions to enhance agency contracting as well as public access to GSA data. At the event, \$12,000 in cash prizes were awarded to two teams who developed working prototypes for the agency.
- **Incentive prizes:** small dollar prizes to encourage creative use of open data.

Action Plan Summary

Sub-Goal	Major Actions to Achieve Impact	Key Indicators
I. Fuel economic growth and Innovation	<ul style="list-style-type: none">Ensuring the continued release and improvement of high priority datasets (e.g. data sets identified as part of the U.S. Open Data Action Plan, high value data assets identified by the public, etc.)Fueling the external open data ecosystem and feedback cycle (e.g. Data Jams, Datapaloozas, code-a-thons, incentive prizes, roundtables, etc.)Developing central tools and support for agencies and innovators (e.g. Project Open Data, data.gov, API management tools, etc.)	<p>Increase in external organizations using open government data, as indicated by:</p> <ul style="list-style-type: none">Usage statistics from Data.govPassive identification by agencies (repeat visitors to agency data website)Proactive identification by agencies (organizations that cite government data usage)Reports from data.gov/impact/ <p>Increased value and outcomes for external entities using government data</p> <ul style="list-style-type: none">Survey top five external data users to understand the value, use, and impact of government data, including job creation, and how agencies can improve the usability of their data.

Action Plan Summary (cont.)

Sub-Goal	Major Actions to Achieve Impact	Key Indicators
II. Make open and machine-readable the new default for all government information	<p>Implementing the Open Data Policy & Executive Order at each agency:</p> <ul style="list-style-type: none"> • Develop and maintain an enterprise data inventory • Make data discoverable to the public • Prioritize and release valuable data through public engagement • Prevent inappropriate disclosure of sensitive information • Assign roles and responsibilities to strengthen the culture of data management 	<p>Develop and maintain an Enterprise Data Inventory</p> <ul style="list-style-type: none"> • % of bureaus and programs in the agency which have data sets represented in the EDI. • Y/N — Inventory includes data sets posted to the agency website, addressed System of Record Notices (SORNs) or Privacy Impact Assessments (PIAs), and any purchased data sets? <p>Make data discoverable to the public</p> <ul style="list-style-type: none"> • % of data sets included in the PDL that have a downloadable URL where the public can access the data. • % of agencies reaching 100% valid metadata. • Y/N – Inventory contains data sets released through FOIA. • % of agencies reaching metadata quality goals (working URLs, file format matches metadata, machine-readable). <p>Prioritize and release valuable data through public engagement</p> <ul style="list-style-type: none"> • Existence of a transparent two-way feedback mechanism. <p>Prevent inappropriate disclosure of sensitive information</p> <ul style="list-style-type: none"> • Agency privacy review processes align with Administration policies. <p>Assign roles and responsibilities to strengthen the culture of data management</p> <ul style="list-style-type: none"> • Is there a Chief Data Officer, Chief Technology Officer, or other senior staff position assigned responsibility for data management and release? • Agencies publicly assign points of Contact for Open Data issues.

Work plan

Milestone Summary				
Key Milestones	Milestone Due Date	Milestone status	Owner	Anticipated Barriers or other Issues Related to Milestone Completion
Integrate CAP goal KPIs into the E-gov Integrated Data Collection	6/2014	Complete	OMB	.
Launch series of Open Data Roundtables through NYU's GovLab to connect companies from the Open Data 500 with agency data owners	6/2014	Complete	OSTP	Several successful roundtables have been hosted with key agencies including the DOC and USDA.
Launch new metadata validator tool	7/2014	Complete	GSA	Completed and available at http://labs.data.gov/dashboard/validate
GSA and OMB develop integrated dashboard for agency KPIs	7/2014	Complete	GSA & OMB	Completed and available at http://labs.data.gov/dashboard/offices
Release "Hacks, Jams and 'Paloozas" handbook for agencies	8/2014	Complete	OSTP	Completed and available at https://project-open-data.cio.gov/engagement/
Improve automated Data.gov harvesting of agencies' Public Data Listing	8/2014	Complete	GSA	
Update Project Open Data Dashboard to add agency-reported open data use and impact information and include validations for metadata quality and machine readability	12/2014	Partially Complete	OMB & GSA	Use and impact information has been added to the Dashboard, as well as validations for metadata quality. However, validations for machine readability have not yet been added. This addition is anticipated soon.
Issue an updated metadata v1.1 that aligns with international standards	12/2014	Complete	OMB & OSTP	
Host 8 API & data usability testing programs	12/2014	Complete	GSA	
Disseminate best practices through Open Data Dashboard	12/2014	Complete	OMB	Completed and available at http://labs.data.gov/dashboard/offices (see stars)

Work plan (cont.)

Milestone Summary				
Key Milestones	Milestone Due Date	Milestone status	Owner	Anticipated Barriers or other Issues Related to Milestone Completion
Ensure all agencies are using metadata v1.1	2/2015	Complete	OMB, Agencies	
Refresh Open Data CAP Goal with metrics on metadata quality and machine readability	2/2015	Complete	OMB	
Hold accessibility hackathon	3/2015	Complete	OSTP, GSA, ED	See White House Blog for details.
Agencies publish their 2/28/15 EDI on Data.gov; explore options for routine release of EDIs	Spring 2015	Complete	OMB	Agency EDIs are available as downloadable datasets on Data.gov. As of May 31, all agencies are required to include information on non-public data in their PDL.
Incorporate refreshed Open Data CAP Goal metrics into PortfolioStat sessions	Spring 2015	Complete and Ongoing	OMB	The second round of quarterly PortfolioStat sessions is currently underway and incorporating Open Data metrics as appropriate.
Agencies include licensing information in their PDL for all public datasets	Summer 2015	On Track	OMB	As of May 31 st , all agencies are required to include license information for all public datasets in their PDL.

Work plan (cont.)

Milestone Summary				
Key Milestones	Milestone Due Date	Milestone status	Owner	Anticipated Barriers or other Issues Related to Milestone Completion
High level datasets identified through the U.S. Open Data Action Plan released on schedule	8/2015	On Track	OSTP	
Publish the third Open Government National Action Plan by October 31, 2015	10/2015	On Track	OSTP	
Schedule 8 additional roundtables	12/2015	On Track	OSTP	
Manage Open Data Working Group	On-going	On Track	OMB & OSTP	
Host Open Data Co-Working Sessions for agencies	On-going	On Track	GSA	

Key indicators

	Key Implementation Data						
Sub-Goal	Indicator	Source	Baseline	Target?	Frequency	Latest data	Trend
I. Fuel economic growth and innovation	Demonstrated value and outcomes for external entities using government data	Agencies Surveys of top data users/IDC Data.gov Center for Open Data Enterprise Open Data 500	Q1 FY15 (updated quarter to quarter)	N/A	Quarterly	In February 2015, 92% of agencies reported qualitative “impact” data, the same as last quarter.	--
	% increase in usage statistics on data views from Data.gov	Data.gov	Q1 FY15 (updated quarter to quarter)	TBD	Quarterly	Views in Q2 FY15: 401,780 Views in Q1 FY15: 322,569 Percent growth: 25%	↑
	% of agencies with Digital Analytics Program (DAP) installed on their agency.gov/data/ page to assist in identifying data users	GSA DAP data	Q1 FY15 (updated quarter to quarter)	100%	Quarterly	As of May 2015, 7 CFO Act agencies (29%) have DAP installed on their [agency].gov/data page, up from 5 (21%) last quarter.	↑
	# of additional usage reports on Data.gov/impact/	Data.gov/impact/	Q1 2015	1 per quarter	Quarterly	9 additions in Q2	↑

Key indicators (cont.)

Key Implementation Data							
Sub-Goal	Indicator	Source	Baseline	Target?	Frequency	Latest data	Trend
II. A) Develop and maintain an EDI	% of bureaus and programs in the agency which have data sets represented in the EDI	Agencies /IDC	Q1 FY15	100% (bureaus) 50% (programs)	Quarterly	<p>In February 2015, 75% of bureaus were represented, up from 64% last quarter.</p> <p>In February 2015, 23% of programs were represented, up from 18% last quarter.</p>	<p>↑ (bureaus)</p> <p>↑ (programs)</p>
	Y/N—Does inventory include data sets posted to the agency website, addressed in the agency’s SORNs or PIAs, and any purchased data sets?	Agencies/IDC			Quarterly	<p>Under Development</p> <p>No current way to check for this. Exploring options for metadata tagging.</p>	

Key indicators (cont.)

	Key Implementation Data						
Sub-Goal	Indicator	Source	Baseline	Target?	Frequency	Latest data	Trend
II. B) Make data discoverable to the public	% of data sets included in the Public Data Listing that have a URL where the public can access the data set	Agency PDL	Q1 FY15	100%	Quarterly	80% of datasets have a downloadable URL in February 2015 (down from 93% last quarter)	↓
	% of agencies with 100% valid metadata	Agency PDL	Q1 2015	100%	Quarterly	79% of agencies have 100% valid metadata in February 2014, the same as last quarter.	--
	% of working download URLs	Agency PDL	Q2 2015	N/A	Quarterly	In February 2015, 78% of were working (e.g., not broken, error, or redirect)	
	% of datasets in which the file format matches the file format identified in the data.json metadata	Agency PDL	Q2 2015	N/A	Quarterly	In February 2015, 66% matched.	
	% of downloads in HTML format (machine-readability measure)	Agency PDL	Q2 2015	N/A	Quarterly	In February 2015, 30% were in HTML format.	
	% of downloads in PDF format (machine-readability measure)	Agency PDL	Q2 2015	N/A	Quarterly	In February 2015, 13% were PDFs.	
	Y/N—Have data sets released via FOIA been included in the PDL?	Agencies/IDC		Yes	Quarterly	Under Development	12

Key indicators (cont.)

	Key Implementation Data						
Sub-Goal	Indicator	Source	Baseline	Target?	Frequency	Latest data	Trend
II. C) Prioritize and release valuable data through public engagement	% of agencies with transparent 2-way feedback mechanism for data prioritization	Agencies/IDC	Q2 FY15	100%	Quarterly	In February 2015, 33% of agencies have demonstrated 2-way feedback loop, the same as last quarter.	--
II. D) Prevent inappropriate disclosure of sensitive information	% of agencies with privacy reviews consistent with the Open Data Policy	Agencies	Q1 FY14	100%	Quarterly	84% of agencies are compliant, the same as last quarter.	--
II. E) Assign roles & responsibilities to strengthen the culture of data management	Y/N—Is there a Chief Data Officer, Chief Technology Officer, or other senior staff position assigned responsibility for data management and release?	Agencies/IDC	Q2 FY15	N/A	Quarterly	81% have position responsible for data management and release.	
	% of agencies who publicly assign points of contact for open data issues	Agencies/IDC	Q1 2015	100%	Quarterly	96% of agencies have identified POCs, the same as last quarter	--

Resources

Resource	Description
Executive Order 13642—Making Open and Machine Readable the New Default for Government Information	Established an important new principle in the government’s stewardship of data: going forward, agencies must consider openness and machine-readability as the default when they collect or create data.
M-13-13—Open Data Policy—Managing Information as an Asset	Policy jointly released by OMB and OSTP to provide a framework for agencies to manage information as an asset throughout its lifecycle, which includes requirements to protect personal and confidential data.
Supplemental Guidance on the Implementation of M-13-13	Guidance to provide additional clarification and detailed requirements to assist agencies in carrying out the Open Data Policy.
Data.gov	The home of the U.S. Government’s open data, which contains data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and more.
Project Open Data	Free open source tools, case studies, and best practices designed to help agencies improve open data management, release data, and host events.
Project Open Data Dashboard	Dashboard of agency progress in implementation of the Open Data Policy, including evaluations from OMB.
Open Data Biweekly Meeting and Listserv	The CTO and CIO teams host an open data meeting, currently every other Tuesday from 11 AM to 12 PM EST at GSA. For more information, join the Open Data Listserv. Hosted by GSA, this listserv is designed to connect people who are tackling data innovation, no matter their home agency. Anyone with a .gov or .mil email address can join by emailing listserv@listserv.gsa.gov (the message should have no subject and the body should say "subscribe open-data").
U.S. Digital Services Playbook	This playbook of 13 key “plays” is drawn from successful best practices from private sector and government that will help government build effective digital services. Play 13, default to open, includes actions for opening data.
StackExchange Open Data Q&A	Open Data Q&A site frequented by federal open data stakeholders.
Open Data 500	Study of companies that use open federal data to generate new business and develop new products and services. The roundtables bring data-centric agencies together with the heaviest users of their data.
U.S. Open Data Institute	Partners with government agencies to help them release data sets and engage users.

Resources (cont.)

Resource	Description
The Center for Open Data Enterprise	Nonprofit organization with a mission to help develop smarter open data strategies for government, businesses, and other nonprofits by focusing on data users.
Let Me Get That Data For You	A free, open source tool that quickly and automatically creates a machine-readable inventory of all the data files found on a given website.
U.S. Public Participation Playbook	A resource for government managers to effectively evaluate and build better services through public participation using best practices and performance metrics.

Contributing programs

The White House

- [Office of Management and Budget \(OMB\)](#)
- [Office of Science and Technology Policy \(OSTP\)](#)
- [Presidential Innovation Fellows](#)

General Services Administration

- [Data.gov](#)
- [18F](#)

Acronyms

- **API – Application Program Interface**
- **CAP – Cross Agency Priority**
- **CDAC – Commerce Data Advisory Council**
- **CEO – Chief Executive Officer**
- **CFO Act – Chief Financial Officers Act**
- **CIO – Chief Information Officer**
- **CTO – Chief Technology Officer**
- **DAP – Digital Analytics Program**
- **DOC – Department of Commerce**
- **DOI – Department of the Interior**
- **ED – Department of Education**
- **EDI - Enterprise Data Inventory**
- **FOIA – Freedom of Information Act**
- **GSA – General Services Administration**
- **IDC – Integrated Data Collection**
- **KPI – Key Performance Indicator**
- **N/A – Not Available**
- **OMB – Office of Management and Budget**
- **OSTP – Office of Science and Technology Policy**
- **PIA – Privacy Impact Statement**
- **PDL – Public Data Listing**
- **SSA – Social Security Administration**
- **SORN – System of Record Notice**
- **USDA – United States Department of Agriculture**
- **VA – Department of Veterans Affairs**