

# 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 3

## 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

- Open data from the U.S. Government is an important national resource, serving as fuel for entrepreneurship, innovation, scientific discovery, and data driven policymaking. Making information about government operations more readily available is also core to a more efficient, transparent, participatory, and collaborative 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.

## Policy

- **New Requirement to Include Explanations for Datasets with Licensing Restrictions.** Building off of FY2015 Quarter 2's new requirement to include license information, agencies were instructed to include explanations for all datasets that do not include a *Public Domain* URL in the license field. [Guidance](#) and [examples](#) for agencies are provided on Project Open Data.
- **New Requirement to Organize Datasets into Collections.** As part of its quarterly Integrated Data Collection (IDC), the White House instructed agencies to enrich their Public Data Listing and Enterprise Data Inventories by ensuring all data assets include the individual datasets. [Guidance and examples](#) for agencies are provided on Project Open Data.
- **New Requirement to Provide 5 Examples of Open Data Maturity Improvements Informed by Public Feedback.** As part of the quarterly IDC requirements agencies were instructed to provide 5 dataset examples of open data improvements, such as: machine-readable formats, real-time delivery, worldwide public domain dedication, bulk downloads, free access to previously charged for data, public access to previously Restricted Public data, and data quality improvements based on public feedback.
- **Project Open Data Dashboard Reviews.** As is done each quarter, the White House Office of Management and Budget has evaluated each agency on their Open Data Policy progress on the [Project Open Data Dashboard](#). The White House applauds the Departments of Agriculture, and Veterans Affairs, the Social Security Administration, and U.S. Agency for International Development for achieving “green” (highest level possible) scores for all performance indicators for the February through May milestone.

## Tools & Collaboration

- On May 18<sup>th</sup>, President Obama launched the [Police Data Initiative](#). Through this effort, 21 local police departments and other participants are responding first to Task Force recommendations within two streams of work: 1) Using open data to increase transparency, build community trust, and support innovation, and 2) Better using technology, such as early warning systems, to identify problems, increase internal accountability, and decrease inappropriate uses of force.
- On May 28<sup>th</sup>, the White House hosted the first [White House Mapathon](#). At the event, participants came together in support of open mapping initiatives. Open mapping, or crowdsourced mapping, describes how millions of people are using a wiki-like approach to contribute to maps of the world and creating shared geo data layers. Recognizing the power that projects like these have to educate, engage, and empower the public to apply their curiosity and contribute their talents to a wide range of scientific and societal problems, President Obama called on Federal agencies in [2013 in the Second Open Government National Action Plan](#) to accelerate and scale the use of citizen science and crowdsourcing projects.
- **Pulse.cio.gov.** In June, 18F, in partnership with the General Services Administration's (GSA) Office of Government-wide Policy, launched [pulse.cio.gov](#), a public dashboard that measures how U.S. government domains are following best practices for federal websites, including the use of HTTPS, and participation in the Federal Government's Digital Analytics Program.

## 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:

- Data Jams: **Brainstorming workshops** with innovators and data owners.
  - OPM hosted a Data Jam and the U.S. Departments of Energy, Transportation and of Veteran Affairs have hosted Open Data Roundtables so far in 2015.
- Datapaloozas: **High-profile summits** to highlight open data progress.
  - For example, U.S. Departments of Energy, Health, and Transportation Datapaloozas have been held so far in 2015.
- Hackathons: 24-72 hour **code-a-thons** to build working prototypes.
  - On May 28<sup>th</sup>, the White House hosted the First [White House Mapathon](#).
  - On June 6<sup>th</sup>, the White House and agencies joined U.S. cities in the [National Day of Civic Hacking](#).
- Incentive prizes: **small dollar prizes to encourage creative use of our 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"><li>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.)</li><li>Fueling the external open data ecosystem and feedback cycle (e.g. Data Jams, Datapaloozas, code-a-thons, incentive prizes, roundtables, etc.)</li><li>Developing central tools and support for agencies and innovators (e.g. Project Open Data, data.gov, API management tools, etc.)</li></ul>	<p><b>Increase in external organizations using open government data, as indicated by:</b></p> <ul style="list-style-type: none"><li>Usage statistics from Data.gov</li><li>Passive identification by agencies (repeat visitors to agency data website)</li><li>Proactive identification by agencies (organizations that cite government data usage)</li><li>Reports from <a href="https://data.gov/impact/">data.gov/impact/</a></li></ul> <p><b>Increased value and outcomes for external entities using government data</b></p> <ul style="list-style-type: none"><li>Survey top 5 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.</li></ul>

# Action Plan Summary (cont.)

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

# 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	6/30/2014	Complete	OMB	.
Launch series of Open Data Roundtables through NYU's GovLab to connect companies from the Open Data 500 with agency data owners	6/30/2014	Complete	OSTP	Several successful roundtables have been hosted with key agencies including the Department of Commerce and USDA.
Launch new metadata validator tool	7/14/2014	Complete	GSA	Completed and available at <a href="http://labs.data.gov/dashboard/validate">http://labs.data.gov/dashboard/validate</a>
GSA and OMB develop integrated dashboard for agency KPIs	7/29/2014	Complete	GSA & OMB	Completed and available at <a href="http://labs.data.gov/dashboard/offices">http://labs.data.gov/dashboard/offices</a>
Release "Hacks, Jams and 'Paloozas" handbook for agencies	8/29/2014	Complete	OSTP	Completed and available at <a href="https://project-open-data.cio.gov/engagement/">https://project-open-data.cio.gov/engagement/</a>
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 <a href="http://labs.data.gov/dashboard/offices">http://labs.data.gov/dashboard/offices</a> (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, Education	See <a href="#">White House Blog</a> for details.
Agencies publish their 2/28/15 Enterprise Data Inventories (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 <a href="#">required</a> 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 <sup>st</sup> , 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 US 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 (cont.)

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 May 2015, 92% of agencies reported qualitative “impact” data, the same as last quarter.	--
	% increase in usage statistics on data views from <a href="http://Data.gov">Data.gov</a>	Data.gov	Q1 FY15 (updated quarter to quarter)	TBD	Quarterly	Views in Q3 FY15: 418,819 Views in Q2 FY15: 371,336 Percent growth: 13%	↑
	% of agencies with Digital Analytics Program (DAP) installed on their <a href="http://agency.gov/data/">agency.gov/data/</a> page to assist in identify 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, the same as last quarter.	--
	# of additional usage reports on <a href="http://Data.gov/impact/">Data.gov/impact/</a>	<a href="http://Data.gov/impact/">Data.gov/impact/</a>	Q1 2015	1 per quarter	Quarterly	5 additions in Q3	--

# Key indicators (cont.)

	Key Implementation Data						
Sub-Goal	Indicator	Source	Baseline	Target?	Frequency	Latest data	Trend
II. A) Develop and maintain an Enterprise Data Inventory	% of bureaus and programs in the agency which have data sets represented in the Enterprise Data Inventory	Agencies /IDC	Q1 FY15	100% (bureaus)  50% (programs)	Quarterly	In May 2015, 80% of bureaus were represented, up from 75% last quarter.  In May 2015, 20% of programs were represented, down from 23% last quarter.	<div> <div>↑</div> <div>(bureaus)</div> </div> <div> <div>↓</div> <div>(programs)</div> </div>
	Y/N—Does inventory include data sets posted to the agency website, addressed in the agency’s System of Record Notices (SORNs) or Privacy Impact Assessments (PIAs), and any purchased data sets?	Agencies/IDC			Quarterly	Under Development  No current way to check for this. Exploring options for metadata tagging.	

# 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	75% of datasets have a downloadable URL in February 2015 (down from 80% last quarter)	↓
	% of agencies with 100% valid metadata	Agency PDL	Q1 2015	100%	Quarterly	79% of agencies have 100% valid metadata in May 2015, the same as last quarter.	--
	% of working download URLs	Agency PDL	Q2 2015	N/A	Quarterly	In May 2015, 72% of download URLs were working, down from 85%.	↓
	% 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 May 2015, 60% matched, down from 64% last quarter.	↓
	% of downloads in HTML format (machine-readability measure)	Agency PDL	Q2 2015	N/A	Quarterly	In May 2015, 19% were in HTML format, an improvement from 27%.	↑
	% of downloads in PDF format (machine-readability measure)	Agency PDL	Q2 2015	N/A	Quarterly	In February 2015, 15% were PDFs, the same as last quarter.	--
	Y/N—Have data sets released via FOIA been included in the Public Data Listing?	Agencies/IDC		Yes	Quarterly	Under Development	

# 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 a transparent 2-way feedback mechanism for data prioritization	Agencies/IDC	Q2 FY15	100%	Quarterly	In May 2015, 29% of agencies have demonstrated 2-way feedback loop, down from 33%.	↓
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	88% have position responsible for data management and release, up from 81%.	↑
	% of agencies who publicly assign points of contact for open data issues	Agencies/IDC	Q1 2015	100%	Quarterly	100% of agencies have identified POCs, up from 96%.	↑

# Resources

Resource	Description
<a href="#">Executive Order 13642—Making Open and Machine Readable the New Default for Government Information</a>	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.
<a href="#">M-13-13—Open Data Policy—Managing Information as an Asset</a>	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.
<a href="#">Supplemental Guidance on the Implementation of M-13-13</a>	Guidance to provide additional clarification and detailed requirements to assist agencies in carrying out the Open Data Policy.
<a href="#">Data.gov</a>	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.
<a href="#">Project Open Data</a>	Free open source tools, case studies, and best practices designed to help agencies improve open data management, release data, and host events.
<a href="#">Project Open Data Dashboard</a>	Dashboard of agency progress in implementation of the Open Data Policy, including evaluations from OMB.
<a href="#">Open Data Biweekly Meeting and Listserv</a>	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 <a href="mailto:listserv@listserv.gsa.gov">listserv@listserv.gsa.gov</a> (the message should have no subject and the body should say "subscribe open-data").
<a href="#">U.S. Digital Services Playbook</a>	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.
<a href="#">StackExchange Open Data Q&amp;A</a>	Open Data Q&A site frequented by federal open data stakeholders.
<a href="#">Open Data 500</a>	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.
<a href="#">U.S. Open Data Institute</a>	Partners with government agencies to help them release data sets and engage users.

# Resources (cont.)

Resource	Description
<a href="#">The Center for Open Data Enterprise</a>	Nonprofit organization with a mission to help develop smarter open data strategies for government, businesses, and other nonprofits by focusing on data users.
<a href="#">Let Me Get That Data For You</a>	A free, open source tool that quickly and automatically creates a machine-readable inventory of all the data files found on a given website.
<a href="#">U.S. Public Participation Playbook</a>	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**