

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 1

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

- **Project Open Data Metadata Schema v1.1.** In November 2014, the White House coordinated with the General Services Administration's (GSA) [Data.gov](#) team to launch the [Project Open Data Metadata Schema v1.1](#), which better aligns with international standards for data publishing and will be the new standard for open data metadata across the federal government. These teams conducted extensive public engagement through workshops and [Github](#) and incorporated community feedback into the final product. Nearly every agency has successfully migrated to the new schema ahead of the February 28, 2015 deadline.
- **Enterprise Data Inventories (EDI) to be Released.** As a result of a Freedom of Information Act (FOIA) request, the Office of Management and Budget (OMB) has agreed to release each agency's EDI. EDIs are comprehensive lists of a federal agency's information holdings. The [Sunlight Foundation](#), which made the request, believes the collective EDIs to be the largest index of government information in the world. The White House has instructed agencies to make their EDIs, with redactions as appropriate to protect sensitive information, available to the public on data.gov by the end of Q2 FY 2015. The White House intends to make release of EDIs a regular and routine process.
- **New Requirement to Catalog Application Program Interfaces (API).** As part of its quarterly Integrated Data Collection (IDC), the White House instructed agencies to identify and list each of its APIs in their EDIs and Public Data Listings (PDL), allowing them to [appear publicly](#) on Data.gov. Agencies were required to catalog all APIs by February 28, 2015.

Tools & Collaboration

- On January 20, 2015, the White House and the Environmental Protection Agency (EPA) jointly hosted a government-wide open data best practices sharing session. The session was held at the White House and available via webcast, and it was open to all federal employees and contractors who work on open data issues. U.S. Chief Technology Officer (CTO) Megan Smith gave remarks, and speakers from the Presidential Innovation Fellow (PIF) program, the Environmental Protection Agency (EPA), the General Services Administration (GSA), and the Departments of Agriculture, Commerce, and Veterans Affairs presented on their successes in implementing the [Open Data Policy](#), engaging with data users, and assessing the impact of their data.
- The U.S. Open Data Institute launched a tool called "[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. This helps agencies to automate their data discovery and inventorying process to populate their EDIs and PDLs.
- On February 3, 2015, the White House launched the [U.S. Public Participation Playbook](#) as a resource to help government better build more responsive and efficient public participation programs and measure their effectiveness. Public engagement and participation is a critical part of understanding data users' needs and ascertaining the impact of open data.
- On January 22, 2015, Jean Fox of the Bureau of Labor Statistics led a webinar entitled, "[User Research and The Paperwork Reduction Act: How to Get it Done](#)" to assist agencies in seeking feedback from users for User Experience (UX) work within the Paperwork Reduction Act (PRA) requirements. A video of the session is available on [YouTube](#).

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.
 - [Hack Housing: Empowering Smarter Decisions](#) was a weekend hackathon held on February 6-8, 2015 at the Zillow headquarters in Seattle, WA. Sponsored by Zillow and the University of Washington (UW), the purpose of the hackathon was to bring together coders, designers, and business professionals to come up with creative solutions to use data to make it easier for first-time homebuyers, low-income renters and senior citizens to find a home that meets their needs. [Prizes](#) were awarded at the end of the hackathon, including a \$10,000 prize was awarded to the creators of “[SmartMove](#),” an web app web application that uses data from the Department of Housing and Urban Development’s Enterprise Geographical Information Systems and Google Maps to assist low-income renters and beneficiaries of government programs move closer to work and to the things that matter to them.
- **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 and 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
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	Missed	OMB & GSA	Project Open Data Dashboard is located at Tracking use and impact information has been added to the Dashboard. However, technical challenges have delayed the launch of automated metadata quality and machine readability validations. Launch anticipated before FY15 Q2 CAP Goal update.
Issue an updated metadata v1.1 that aligns with international standards	12/2014	Complete	OMB & OSTP	
Host eight API and 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	All agencies except for the Department of State have migrated. The White House and Data.gov teams are working with State to assist with the transition.
Refresh Open Data CAP Goal with metrics on metadata quality and machine readability	2/2015	Missed	OMB	Placeholder metrics were added to the quarterly update. Technical challenges have delayed the launch of automated quality and machine readability validations. Without this, we cannot establish a baseline or targets. Launch anticipated before FY15 Q2 CAP Goal update.
Hold accessibility hackathon	3/2015	On Track	OSTP, GSA, Education	
Agencies publish their 2/28/15 EDIs on Data.gov; explore options for routine release of EDIs	Spring 2015	On Track	OMB	
Incorporate refreshed Open Data CAP Goal metrics into PortfolioStat sessions	Spring 2015	Complete and Ongoing	OMB	PortfolioStat, a formerly annual process, was recently enhanced by changing it to a quarterly event. The first round of quarterly sessions was just completed, and Open Data metrics we incorporated as appropriate.
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 eight 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	Q4 FY14	N/A	Quarterly	In August 2014, 77% of agencies reported qualitative “impact” data. That number rose to 92% in November 2014. Data will continue to be collected quarterly.	↑
	% increase in usage statistics on data views from Data.gov	Data.gov	Q4 FY14	TBD	Quarterly	Views in Q4 FY14: 203,808 Views in Q1 FY15: 322,569 Percent growth: 58%	↑
	% of agencies with Digital Analytics Program (DAP) installed on their agency.gov/data/ page to assist in identify data users	GSA DAP data	Q1 FY15	100%	Quarterly	As of February 2015, 5 of 24 (21%) of CFO Act agencies have DAP installed on their [agency].gov/data page.	
	# of additional usage reports on Data.gov/impact/	Data.gov/impact/	Q4 2014	1 per quarter	Quarterly	1 addition in Q1	↑

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	Q4 FY14	100% (bureaus) 50% (programs)	Quarterly	In August 2014, 112 of 174 bureaus (64.4%) and 279 of 1,544 programs (18.1%) were represented. By November 30, 122 of 174 bureaus (70.1%) and 268 of 1,553 programs (17.3%) were represented.	<div> <div>↑</div> <div>(bureaus)</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 or Privacy Impact Assessments, and any purchased data sets?	Agencies/IDC	Q2 FY15		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	Q2 FY14	100%	Quarterly	93% of datasets listed in agency PDLs have a downloadable URL in November 2014. (down from 96% in August 2014.)	↓
	% of agencies with 100% valid metadata	Agency PDL	Q2 2014	100%	Quarterly	79% of agencies have 100% valid metadata in November 2014. (up from 65% in August 2014.)	↑
	% of working download URLs	Agency PDL	Q2 2015	N/A	Quarterly	Under Development	
	% of datasets in which the file format matches the file format identified in the data.json metadata	Agency PDL	Q2 2015	N/A	Quarterly	Under Development	
	% of downloads in HTML format (machine-readability measure)	Agency PDL	Q2 2015	N/A	Quarterly	Under Development	
	% of downloads in PDF format (machine-readability measure)	Agency PDL	Q2 2015	N/A	Quarterly	Under Development	
	Y/N—Have data sets released via FOIA been included in the Public Data Listing?	Agencies/IDC	Q2 FY15	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	Q1 FY15 (re-baselined with stricter requirements)	100%	Quarterly	33% of agencies have demonstrated 2-way feedback loop (November 2014)	
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 (November 2014)	--
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	Initiated new data collection for Q2 FY15 to track Chief Data Officers	
	% of agencies who publicly assign points of contact for open data issues	Agencies/IDC	Q4 2014	100%	Quarterly	96% of agencies have identified POCs, up from 92% last quarter (November 2014)	↑

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 (PIF)

General Services Administration

- Data.gov
- 18F