

Cross Agency Priority Goal

Quarterly Progress Update

Open Data

Goal leaders: Steven VanRoekel, Federal Chief Information Officer, OMB; Todd Park, Chief Technology Officer, OSTP



FY2014 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 this 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.

*http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information

Progress Update

Policy

- **U.S. Open Data Action Plan.** May 9, 2014—In fulfillment of the G7 Open Data Charter, the Administration published the US Open Data Action Plan, a roadmap for releasing and improving federal government data for the public, including a list of planned data releases and improvements.
- An updated **Integrated Data Collection (IDC)** has been developed and is to be utilized for the August 2014 IDC activities, and specific impact assessment questions were added to have agencies identify top users of their published data sets as well as the perceived value and quality of that data by the end users.

Tools & Collaboration

- The White House maintains a collaborative site called **Project Open Data** that includes free tools, case studies, and software code which make it easier for agencies to liberate data. **Project Open Data Dashboard** was launched and was utilized to review Milestone 3 agency submitted data. This tool also crawls agency open data repositories to verify inventories and their viability, including metadata validity and accessibility via published URLs. It will continue to undergo refinements with feedback from the user community.
- OSTP & OMB host a bi-weekly, inter-agency meeting with open data leads across government; GSA hosts weekly co-working sessions where teams can collaborate on common problems, and OSTP launched a 300+ person open data email group to share information across the US Government. GSA also hosts **user research sessions** with developers and data scientists to review and improve their experience in using open government data.

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.
 - Incentive prizes: **small dollar prizes to encourage creative use of our data.**
- Industry, academia and the non-profits have been key partners:
 - **StackExchange** has set up an open Q&A site about Open Data
 - **McKinsey** published a report outlining more than \$3 trillion in additional economic value to be gained from open data globally.
 - NYU's GovLab created an **Open Data 500**, a study of companies that use open federal data to generate new business, and develop new products and services.
 - The GovLab is hosting a series of **Open Data Roundtables** with a number of agencies, bringing data-centric agencies together with the heaviest users of their data – starting with Department of Commerce in June 2014.
 - The non-profit U.S. **Open Data Institute** was founded in 2014, to partner with government agencies to help them release data sets and engage users.

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 continued release and improvements 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:</p> <ul style="list-style-type: none">• Usage statistics from Data.gov• Passive identification by agencies (repeat visitors to agency data website)• Proactive Identification by agencies (organizations that cite gov data usage) <p>Increased value and outcomes for external entities using government data</p> <ul style="list-style-type: none">• Survey top 5 external data users to understand value and job creation

Action Plan Summary

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"> A. Develop and maintain an enterprise data inventory B. Make data discoverable to the public C. Prioritize and release valuable data through public engagement D. Prevent inappropriate disclosure of sensitive information E. 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, divisions or sections of the agency which have data sets represented in the Enterprise Data Inventory • Y/N—Does inventory include all data sets posted to the agency website, addressed in the agency’s System of Record Notices (SORNs), 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 Public Data Listing that have a URL where the public can access the data set • Y/N—Have all data sets released via FOIA been included in the Public Data Listing? <p>Prioritize and release valuable data through public engagement</p> <ul style="list-style-type: none"> • How has feedback and input led to change in data management and release practices? <p>Prevent inappropriate disclosure of sensitive information</p> <ul style="list-style-type: none"> • Average number of days it takes to complete the privacy review process and release a data set? • Number of known privacy breaches or other incidents where privacy has been compromised? <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?

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	Complete	OMB E-Gov	Agencies may (or may not) have a good two-way relationship with their data users, and this may result in differing levels of quality and completeness of IDC.
Launch series of Open Data Roundtables through GovLab to connect companies from the Open Data 500 with agency data owners	6/30	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	Complete	GSA	
GSA and OMB develop integrated dashboard for agency KPIs	7/29	Complete	GSA & OMB	Completed and available at http://labs.data.gov/dashboard/offices
Release “Hacks, Jams and ‘Paloozas” handbook for agencies	8/29	Complete	OSTP	Completed and available at http://project-open-data.github.io/engagement/
Improve automated Data.gov harvesting of agencies Public Data Listing	8/29	On Track	GSA	
Issue an updated metadata v1.1	12/15	On Track	OMB & OSTP	
High level datasets identified through the US Open Data Action Plan released on schedule	12/15	On Track	OSTP	
Schedule 8 additional roundtables	12/15	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	
Host an additional API & Data usability testing programs (6 held to date)	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	# Increase in external organizations using Open Government Data	Agencies Data.gov Open Data500	Q1 FY15	N/A	Quarterly	Under Development	N/A
	\$ Increased value and outcomes for external entities using government data	Agencies Surveys of top data users/IDC	Q1 FY15	N/A	Bi-Annual	Agencies reported survey results via August IDC activity. 77% have reported qualitative "impact" data.	↑
II. A) Develop and Maintain an Enterprise Data Inventory	% of bureaus, divisions or sections of the agency which have data sets represented in the Enterprise Data Inventory	Agencies /IDC	Q1 FY15	100% by 11/30/14	Quarterly	The average agency has 72% of its subcomponents represented in the EDI.	↑
	Y/N—Does inventory include all data sets posted to the agency website, addressed in the agency's System of Record Notices (SORNs), Privacy Impact Assessments (PIAs), and any purchased data sets?	Agencies/IDC	Q1 FY15	100% by 11/30/14	Quarterly	41% of agencies are compliant with this today.	↑
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	Agencies PDL	Q2 FY14	100%	Quarterly	The average agency has 78% of datasets in the PDL which include an access URL.	↑
	Y/N—Have all data sets released via FOIA been included in the Public Data Listing?	Agencies/IDC	Q1 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	How has feedback and input led to change in data management and release practices?	Agencies/IDC	Q2 FY14	N/A	Quarterly	100% of agencies have public engagement mechanisms, and 50% of agencies have demonstrated 2-way feedback look	↑
II. D) Prevent inappropriate disclosure of sensitive information	# Average number of days it takes to complete the privacy review process and release a data set?	Agencies/IDC	Q1 FY15	N/A	Quarterly	Under Development	
	# Number of known privacy breeches or other incidents where privacy has been compromised?	Agencies/IDC	Q1 FY15	0	Quarterly	Under Development	
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 FY14	Yes	Quarterly	DOC and DOT have new CDOs	↑