U.S. OPEN DATA ACTION PLAN

May 9, 2014

Licensing

As a work of the United States Government, this document is in the public domain within the United States. Additionally, the United States Government waives copyright and related rights in this work worldwide through the CC0 1.0 Universal Public Domain Dedication.

Introduction

Throughout his Administration, President Obama has articulated a vision of the U.S. Government managing information as a national asset and opening up its data, where possible, as a public good to advance government efficiency, improve accountability, and fuel private sector innovation, scientific discovery, and economic growth. Putting government data online and making it easy to find and use—while continuing to rigorously protect privacy—can help American families find the right health care provider, identify the college that provides the best value for their money, keep their families safe by knowing which products have been recalled, and much more.

On June 18, 2013, President Obama and other G7 leaders endorsed the Open Data Charter. The Open Data Charter sets out five strategic principles:

- Open Data by Default – foster expectations that government data be published openly while continuing to safeguard privacy;
- Quality and Quantity – release quality, timely and well-described open data;
- Useable by All – release as much data in as many open formats as possible;
- Releasing Data for Improved Governance – share expertise and be transparent about data collection, standards and publishing processes; and
- Releasing Data for Innovation – consult with users and empower future generations of innovators.

The U.S. Government continues to make significant progress ensuring government data is more available and useable to the public. Government data is structured information that is created, collected, processed, disseminated, or disposed of by or for the Federal government. Examples of progress include:

- President Obama's Executive Order in May 2013 to make machine-readable and open the new default for government information;
- An Open Data Policy memorandum to Federal agencies directing management of information as an asset;
- Project Open Data, a White House initiative designed to share best practices, examples, and code to assist agencies with opening data;
- The second U.S. Open Government National Action Plan, published December 2013 which outlines select upcoming open data commitments; and
- Additional, useful data sets being published on Data.gov, as well as the release of open source software code for Data.gov.

Building upon these efforts, and as set forth in the Open Data Charter, the U.S. Government is releasing this U.S. Open Data Action Plan which outlines new commitments as well as plans for enhancements and releases of certain data assets across the categories set forth by the Charter.

Commitments

The United States commits to the following actions in concert with the Open Data Charter and its principles:

1. Publish open data in a discoverable, machine-readable, useful way

Data.gov serves as a national catalog for open data sets—the public can find everything from data about consumer complaints about private student loans from the Federal Consumer Financial Protection Bureau to 911 service-area boundaries from the state of Arkansas. In January 2014, Data.gov was relaunched using open-source software, having been developed in the open with the help of community contributors.

The public can also get information about open government data from each Federal agency. Many Federal agencies are developing websites at [department].gov/data to better explain the data assets they create, manage and publish. All executive branch agencies are now required to develop a machine-readable catalog of their public data at [department].gov/data.json, such as www.treasury.gov/data.json, allowing simple software tools to automatically get updated information about the latest data available and have access to the same information that Data.gov collects and publishes.

Even when government data assets are technically public and available online, there can still be barriers to using and understanding them. The U.S. Government will improve the way it describes and publishes data sets by better focusing on user needs. Specific actions include:

- The General Services Administration (GSA) Data.gov team will conduct regular usability tests to observe whether users are able to easily complete the most important tasks on the site and prioritize enhancements accordingly;
- All Federal Agencies will now have access to free, regular training on user research and usability studies from the GSA; and

- GSA will launch an Application Programming Interface (API) usability training program to support Federal employees in user research to inform when and if an API—used to connect data or services to other digital tools—is appropriate, and how to optimize API documentation and features.

2. Work with the public and civil society organizations to prioritize open data sets for release

New data sets will be prioritized for release based on public feedback. To make that feedback loop possible, each agency will use online feedback mechanisms such as an email address or an online platform to solicit input. The public can give feedback and browse listed data sets that could potentially be made open but are not yet publicly available for download at [department].gov/data. Publishing a list of public data sets that include data that *could* be made public lets civil society, entrepreneurs, researchers, and the general public more easily connect with the Federal government employee responsible for a particular data set and explain why its release should be prioritized.

3. Support innovators and improve open data based on feedback

The U.S. Government will make it simpler for entrepreneurs and innovators to find, understand, and use open government data to develop new products, services, and companies.

There is more than $3 trillion of potential additional economic benefit from open data worldwide, according to a recent McKinsey report. Entrepreneurs and innovators are constantly working on new and interesting ways to use the public, free resource of open government data. New York University's GovLab conducted a study of 500 companies and organizations that are fueled by open data, and is now hosting a series of Open Data Roundtables to bring companies and government data owners together. Specific, actionable feedback from these sessions and others has the potential to improve descriptions, formats, and accessibility of government data.

Further engagement will include:

- Holding online and in-person customer feedback and engagement mechanisms, including data jams (workshops) and datapaloozas (high-profile summits);
- Offering shared services and open-source tools to share and improve open data, such as Data.gov's open-source platform that includes WordPress for content management and CKAN for cataloging data, which are made available for improvement and re-use;

- Facilitating sector-specific feedback sessions with open data users and stakeholders; and
- Launching well-designed incentive prizes and challenges to use open government data in pursuit of policy goals.

4. Continue to release and enhance high-priority data sets

In carrying out the President's Executive Order to make open and machine-readable the new default for government information, and following the Open Data Policy memorandum to Federal agencies about managing information as an asset, Federal agencies are regularly releasing and enhancing data sets.

To help with this task, the U.S. Government will recruit a third round of Presidential Innovation Fellows to work on a series of Data Innovation projects. Presidential Innovation Fellows are entrepreneurs, developers, designers, and other innovators typically from the private sector who join government agencies for six to 12-month fellowships and work closely with the office of the U.S. Chief Technology Officer. Presidential Innovation Fellows will further scale the Administration's data innovation initiatives in climate change, health, energy, education, finance, and public safety. These projects include both open data initiatives and "My Data" initiatives. "My Data" initiatives, such as the Blue Button (for healthcare) and Green Button (for energy usage), are designed to give Americans secure electronic access to their own personal data.

Each of the following projects will be led by a Presidential Innovation Fellow:

- **The National Oceanic and Atmospheric Administration** is working to make its vast weather, climate, and Earth observation data holdings more easily available and usable in the cloud to unleash the full potential of these resources, spur economic growth, and help entrepreneurs launch businesses.

- **The Census Bureau** collects and produces a wealth of geospatial, demographic, and economic data resources, and is seeking to make its maps and geospatial information more easy to access and use.

- **The National Aeronautics and Space Administration** works to make its Earth observation data open and machine-readable, and is expanding its climate data to be easily findable, machine-readable, and usable by innovators developing new climate-resilience and climate-preparedness tools.

- **The U.S. Department of the Interior** is working to make a wide variety of newly catalogued government data easy for entrepreneurs and innovators to discover and use, including data about tourism and recreation opportunities on the Nation's public lands and waters.

- **The U.S. Department of Labor,** in support of the President's Skills and Training Data Initiative and Safety Data Initiative, is working to make its skills and safety information data resources more open, machine-readable, and useful for third parties and innovators.

- **The Internal Revenue Service** is introducing many new digital services for taxpayers, including making it easier to securely access their own tax account, make mobile payments, check their refund status, or conduct other transactions. The Agency continues to work with its many third-party stakeholders to deliver better services in support of tax administration.

- **The U.S. Department of Health and Human Services**, in collaboration with **the U.S. Department of Veterans Affairs**, **the U.S. Department of Defense**, and **the Centers for Medicare & Medicaid Services,** is leading the public-private Blue Button Initiative to expand consumer access to electronic health information. Blue Button empowers consumers with secure access to their own healthcare information—such as medical records, prescription information, medical claims, and lab data—so they can choose to share it with health care providers, caregivers, and others they trust.

- **The U.S. Department of Energy** is working to accelerate the commercialization of National Laboratory-generated technologies, in part by making information about those technologies easier to find and use.

Selected Examples of Data Releases

The United States already publishes a number of data sets in the categories set forth by the Open Data Charter (see more on Data.gov). The data sets below are selected examples of expansions, enhancements, and releases.

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Companies	[Consumer Financial Complaints](#)	Public Domain	API, CSV, XML	The data currently includes complaint details about private student loans, mortgages, credit cards, debt collection, bank accounts and services, credit reporting, money transfers, and other consumer loans. The type of data about each complaint includes, among other things company name to length of response time and result. The upcoming enhancement will include new types of complaints and new fields.	August 2014
Companies	[Dynamic Small Business Search](#)	Public Domain	API	The government's small business data base (Dynamic Small Business Search) will be revamped by: • Improving accessibility to the data, and • Making the data available for third parties through an API. The improvements are designed to enable third parties to create, for example, tools to give more context about contractors and suppliers so that companies can more easily locate qualified U.S. suppliers for their manufacturing facilities, ultimately reducing the transaction costs for firms working to source products and manufacture in the United States.	January 2015

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Crime and Justice	Grants Program Plan	Public Domain	API	This action focuses on releasing an API of forecasted fiscal year funding opportunities from the Office of Justice Programs in the Department of Justice. Data available will include funding opportunity title and description, program type/area, awarding office, expected release quarter, actual release date, applicant eligibility type, awarding agency point of contact, amount available, and average award amount.	December 2014
Crime and Justice	Department of Justice Developer Hub		Bulk, API	The Department of Justice will release a developer hub with bulk and API access to data, as appropriate.	December 2014
Earth Observation	International Surface Hourly Data	Public Domain	FTP, .txt, ASCII	This data includes hourly reports on parameters such as wind speed and direction, wind gust, temperature, dew point, cloud data, sea level pressure, altimeter setting, station pressure, present weather, visibility, precipitation amounts for various periods, snow depth as observed by each of 20,000 stations worldwide.	Ongoing

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Earth Observation	Hurricane Satellite	Public Domain	KML, netCDF	Improvements made to the collection of visible and infrared images collected by U.S. and international satellites. This will be useful for estimation of hurricane intensity and rates of intensification; applications for insurance, and construction industries. Improvements included: • Better data connections to compare and connect data sets, and • More combined data to provide additional context.	January 2014
Earth Observation	Landsat Surface Temperature	Public domain	GeoTiff, XML metadata files	This effort will result in a release of surface temperature data from Landsat 8 operational thermal infrared sensor, which is used in energy models to monitor water use for irrigated lands.	December 2014
Education	National Center for Education Statistics' Common Core of Data and Integrated Postsecondary Education Data Systems	Public Domain	API	This data set will be released via API.	November 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Education	Smithsonian Cooper-Hewitt National Design Museum Collection	CC0	Open	The Smithsonian Cooper-Hewitt National Design Museum Collection plans to make all digitized collections metadata public domain, and digitized collection images without copyright or other restriction publicly available at the highest available resolution for non-commercial, educational use.	December 2014
Education	Smithsonian Freer Gallery of Art and Arthur M. Sackler Gallery Collections	CC0	Open	The Smithsonian Freer Gallery of Art and Arthur M. Sackler Gallery plans to make all digitized collections metadata public domain, and digitized collection images without copyright or other restriction publicly available at the highest available resolution for non-commercial, educational use.	December 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Education	Smithsonian American Art Museum Collection API	CC0	Open	After a successful limited release of an API of the Smithsonian American Art Museum collection and hackathon that resulted in a number of working prototypes, the Smithsonian American Art Museum is planning a staged release, from open metadata, like artist or medium, to an open API of digitized collections images without copyright or other restriction available for non-commercial, educational use.	December 2014
Energy and Environment	Long-term Energy Projections from EIA's Annual Energy Outlook	Public Domain	API, CSV	The U.S. Energy Information Administration released this data set via API and bulk files.	March 2014
Energy and Environment	OpenFEMA initiative		Open	OpenFEMA will regularly release new disaster-related data in a machine-readable format and host workshops to build tools that support first responders, survivors, and impacted communities.	Initiative is Live

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Energy and Environment	Recreation on Public Lands Information			This effort will result in a release of data about campsites and other recreational opportunities on public land.	March 2014
Energy and Environment	Watershed Assessment, Tracking, & Environmental Results System	Public Domain	Open	This data set will be expanded to include information about national hydrography.	January 2014
Energy and Environment	Buildings Performance Database	Public Domain	Open	Enhancements include a data table tool that allows users to generate and export statistical data about groups of buildings by type and region, a financial forecasting tool that forecasts cash flows for energy efficiency projects and an API that allows external software to conduct analyses of the data.	January 2014
Finance and Contacts	Natural Resource Revenues			Improve usability and documentation of data on revenues from natural resource extraction from public lands.	April 2014
Geospatial	Wind Turbine Details and Mapping	Public Domain	Open	This data will be released as map with an API.	March 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Geospatial	CropScape	Public Domain	GeoTIFF, htm and .xml, wms, wfs, wcs, GET, POST, and SOAP.	Released new map layers including river, lake, and road layers at national and regional level to the geo-referenced, crop-specific land cover data.	January 2014
Geospatial	Hurricane Katrina Sampling and Monitoring	Public	Open	Released geospatial data about the impact of natural disaster Hurricane Katrina.	January 2014
Global Development	Global Food Security	Public Domain	XLS	The underlying data model will be released, as part of the Global Open Data on Agriculture and Nutrition, an initiative established to increase the quality, quantity, and timeliness of available data to support agriculture and nutrition efforts in food security. By releasing the model, researchers will be able to evaluate the assumptions made in analyzing the data set and to also draw their own conclusions through changes in these assumptions.	Summer 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Global Development	Famine Early Warning Systems Network	CC0	Open	Will be released for bulk download.	January 2015
Global Development	Foreign Assistance Dashboard		Open	U.S. Government agencies that provide Foreign Assistance funds (Development Aid) regularly add data about those activities.	Quarterly Updates
Government Accountability and Democracy	Federal Employee Human Resources Data	Public Domain	Text	Will securely share data about employment status through web services to verify employees and review dynamic HR data to determine data sets that can made publicly available consistent with privacy protections.	September 2014
Government Accountability and Democracy	Federal Employee Payroll Data	Public Domain	Text	Will implement architecture, data quality, and usability improvements.	December 2014
Government Accountability and Democracy	Federal Employee Training Data	Public Domain	Text	Will implement architecture, data quality, and usability improvements and add acquisition training data.	September 2014
Government Accountability and Democracy	Federal Employee Retirement Data Repository	Public Domain	XML	Will begin production data feeds from Federal payroll centers, and integrate data with other personnel systems.	December 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Health	Adverse Event Reporting System	CC0	RESTful API, XML Download	Previously only available through limited quarterly reports and FOIA requests, these 3.9 million individual reports of drug adverse events and medication error reports are submitted anonymously by healthcare professionals and consumers will be made available in their entirety via API and bulk download. They are a critical tool to support safety surveillance.	June 2014
Health	Product Recalls	Public Domain	RESTful API, XML	The Food and Drug Administration will release a RESTful API and XML of the public product recall data and add data not currently accessible from existing resources, such as product identifiers and geographic details.	August 2014
Health	National Notifiable Disease Surveillance System	Public Domain	API, CSV, XML, JSON, REST, SOAP	Notifiable disease data will incorporate Tuberculosis data as well as combining weekly data into more accessible data sets with static URLs.	April 2014
Health	Structured Product Labels	Public Domain	RESTful API, XML Download	Will be released for the first time and made available via API. With Structured Product Label data, for example, you could make tools to more easily identify which specific products have been affected by a recall.	Summer 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Health	The National Health Interview Survey	Public Domain	ASCII	This data set, which covers health status and conditions, disability, access to and use of health services, health insurance coverage, immunizations, risk factors, and health-related behaviors will be expanded to include a representative sample of Native Hawaiian and Pacific Islanders (NHPI). Having a representative sample makes it possible to understand NHPI health problems, strengths, and needs	Summer 2015
Health	Medicare Reimbursements	Public Domain	CSV, XLS, JSON	Historically, a permanent injunction, issued in 1979, prohibited the government from disclosing annual Medicare reimbursements paid to individual physicians under Medicare. In May 2013 this injunction was lifted. As a result, this data set was the first time CMS publicly released information about how physicians and other health care professionals in the Medicare program are paid.	April 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Health	National Sample Survey of Nurse Practitioners	Public Domain	SAS and ASCII	For the first time, this anonymous survey data will be available on nurse practitioner practice patterns including: specialties, services offered, billing arrangements, relationships with physicians, impact of legal restrictions on practice, and job satisfaction.	Spring 2014
Science and Research	Interactive Chemical Safety for Sustainability Dashboard	Public Domain	Open	This dashboard was released in beta.	December 2013

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Science and Research	Lab Data Clearinghouses			Several Federal labs have adopted common data standards and host a central repository for scientific data in a given field, taking into account the need to separate proprietary data from public data. Clearinghouses such as the Secure Transportation Data Project and the Buildings Performance Database address these challenges by aggregating proprietary data into non-proprietary data clusters and stripping personally identifiable information and proprietary information from data points to make some of this data available to the public. By the end of 2015, additional clearinghouses will be formed and will help enable research and business use while protecting personal information and non-public proprietary commercial data.	March 2015
Statistics	Energy Production and Consumption for the United States	CC0	Open	These historical (50 years to present) U.S. energy production and consumption statistics will be released via API and bulk files.	June 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Statistics	The Longitudinal Employer-Household Dynamics	Public Domain	Open	This longitudinal database of linked employer-employee microdata covers more than 95 percent of U.S. private-sector jobs. Today, access to certain microdata is only granted to qualified researchers on approved projects and at restricted sites. This database will expand segments of data to be available online for public use while continuing to protect privacy.	September 2014
Statistics	U.S. Coal Statistics	CC0	Open	This data about coal production, mining, imports, exports, and shipments to electricity plants will be released via API and bulk files.	March 2014
Social Mobility and Welfare	Courses and Curriculum Developed Through the Trade Adjustment Assistance Community College and Career Training Grant Program	CC BY	API	This data about education and career training programs that can be completed in two years or less, and prepare program participants for employment in high-wage, high-skill occupations, will be available to the public on an open, online open educational resources repository.	June 2014

Data Category	Example Data Set	License	Format	Further Activity or Enhancement	To Be Completed By:
Social Mobility and Welfare	Agricultural Census	Public Domain	QuickStats API, XLS	The full agricultural census data set will be released, which can help the public evaluate changes to the American agricultural sector and food production over time.	May 2014
Transport and Infrastructure	Alternative Fuel Station Locator and Downloadable Data	CC0	Open	Will be updated based on feedback received at data summits.	September 2014
Transport and Infrastructure	Travel Warnings and Travel Alerts	Public Domain	API	U.S. Department of State's Bureau of Consular Affairs will release a new API to share travel warnings and alerts.	Fall 2014