OPEN GOVERNMENT DATA ACT

REPORT

OF THE

COMMITTEE ON HOMELAND SECURITY AND GOVERNMENTAL AFFAIRS
UNITED STATES SENATE

TO ACCOMPANY

S. 760

TO EXPAND THE GOVERNMENT'S USE AND ADMINISTRATION OF DATA TO FACILITATE TRANSPARENCY, EFFECTIVE GOVERNANCE, AND INNOVATION, AND FOR OTHER PURPOSES

JULY 24, 2017.—Ordered to be printed
OPEN GOVERNMENT DATA ACT

JULY 24, 2017.—Ordered to be printed

Mr. JOHNSON, from the Committee on Homeland Security and Governmental Affairs, submitted the following

REPORT

[To accompany S. 760]

[Including cost estimate of the Congressional Budget Office]

The Committee on Homeland Security and Governmental Affairs, to which was referred the bill (S. 760) to expand the Government's use and administration of data to facilitate transparency, effective governance, and innovation, and for other purposes, having considered the same, reports favorably thereon with amendments and recommends that the bill, as amended, do pass.

CONTENTS

I. Purpose and Summary ................................................................. 1
II. Background and Need for the Legislation .................................. 2
III. Legislative History ................................................................. 4
IV. Section-by-Section Analysis ...................................................... 4
V. Evaluation of Regulatory Impact .............................................. 6
VI. Congressional Budget Office Cost Estimate ............................ 6
VII. Changes in Existing Law Made by the Bill, as Reported .......... 7

I. PURPOSE AND SUMMARY

The purpose of S. 760, the Open, Public, Electronic, and Necessary Government Data Act (or the “OPEN Government Data Act”) is to better enable certain Federal Government data to be accessible and useable by the public. It would do so by setting a standard that data assets be maintained in an open format. An open format means the data assets are machine readable and not in a proprietary format or subject to licensing. The bill would require Federal agencies to maintain an inventory of data assets created, owned, or managed by the agency. The bill would also require data assets in the inventory that can be made public be publicly
On May 25, 2016, the Committee approved S. 2852, the Open, Public, Electronic, and Necessary Government Data Act. The bill passed the Senate on December 10, 2016. That bill is substantially similar in purpose to S. 760. Accordingly, this committee report is in large part a reproduction of Chairman Johnson’s committee report for S. 2852, S. Rep. No. 114–396.


See generally, Gov’t Accountability Office, GAO–14–476, Data Transparency Oversight Needed to Address Underreporting and Inconsistencies on Federal Websites (2014).


Id. at 2.


Id. at 51.

available. Finally, the bill strengthens the Federal Chief Information Officer’s (CIO) role in oversight of open data and information resources management policy.1

II. BACKGROUND AND THE NEED FOR LEGISLATION

Open data, or data that is made available to use without restrictions, has led to innovation in both the public and private sectors, supported economic growth, and helped to improve performance and transparency in government programs.2 A prominent example can be found with the creation of “safety.data.gov,” which is a centralized resource for public safety information. In 2012 the Department of Transportation, in partnership with the Department of Justice, Department of Labor, and Consumer Product Safety Commission, launched safety.data.gov which includes information helpful to consumers such as recalls related to cars, household products, recreation products and others.3

According to the Government Accountability Office (GAO), some agencies are not adequately sharing or making their data on Federal spending easy to use.4 This lack of openness has led to problems of quality and completeness of the data presented.5 According to GAO, “a lack of government-wide data standards limits the ability to measure the cost and magnitude of federal investments and hampers efforts to share data across agencies to improve decision-making and oversight.”6

For example, in the GAO’s 2015 report on opportunities to reduce fragmentation, overlap and duplication, and achieve other financial benefits, GAO included examples of agencies potentially unnecessarily duplicating work because they failed to share information.7 In one instance, the Environmental Protection Agency and the Food and Drug Administration both inspected the same laboratories for “good laboratory practices.”8 Because the agencies did not share data, they may have been unaware that they were performing the same oversight function on the same facility. GAO recommended that they improve the quality of their data on inspections and develop a written agreement outlining how they plan to collaborate to remove the duplication and overlap in the process.9

In the case of software management, GAO found that only two
agencies had effective management policies that reduced the number of unused software licenses. GAO recommended that those best practices be shared with all agencies to improve software management Government-wide.10

Ensuring certain Government data assets are available to the public may also lead to more private sector innovation. For example, entrepreneurs used publicly-available Federal Aviation Administration statistics combined with weather data from the National Weather Service to predict flight delays. The resulting application now helps the public make decisions on which flights to take and how much time to leave between connecting flights based on which flights are likely to be on time.11

The Obama Administration issued several directives and memoranda in an effort to encourage agencies to open up their data;12 however, these policies did not result in transparency of all data assets.13 Open Government Data: The Book revealed that most of the data published as a result of the Administration’s policies was already publicly available elsewhere.14 The major concern is that data assets that could be valuable to the public—administrative records and data assets relied on by the Government to make decisions—has not been posted publicly, despite the Office of Management and Budget’s (OMB) direction to do so.15 Additionally, some of the data assets that have been posted publicly are not useful because they are posted in a proprietary format. The need is to both expose the data assets that have not been available and liberate the data assets that have been locked behind licensing and proprietary formats.

Congress has tried to make more data available to the public by passing legislation encouraging agencies to make Federal Government data more transparent. For example, the DATA Act, which was signed by President Obama on May 9, 2014, required Federal agencies to report additional financial and spending information of the Federal Government on a public website.16 Most importantly, the law requires agencies to report the information using a common set of data standards to ensure comparability and usability.17

But while new laws have spurred openness with respect to some Federal data, S. 760 would require, where not otherwise prohibited by law and to the extent practicable, data assets that the Federal Government owns or manages to be open to the public—meaning the data assets are machine-readable and not in a proprietary format or subject to licensing. Importantly, data assets are not in-

---

10 Id. at 181.
14 Id.
15 Id.
17 Id.
tended to include all data, but rather a collection of data elements or data sets that may be grouped together by an agency such as, for example, data charts. Posting these data assets is likely to reduce requests for information through the Freedom of Information Act because it makes the information immediately retrievable.

The bill would also require agencies to develop an inventory that accounts for data assets created or maintained by an agency and to include in the inventory whether each data asset is public or nonpublic. This bill is intended to apply to new Government data collections.

III. LEGISLATIVE HISTORY

S. 760, the OPEN Government Data Act, was introduced on March 29, 2017, by Senators Brian Schatz (D–HI) and Ben Sasse (R–NE). The bill was referred to the Committee on Homeland Security and Governmental Affairs.

The Committee considered S. 760 at a business meeting on May 17, 2017. The legislation was adopted by voice vote en bloc. Senators present for the vote were Johnson, McCain, Portman, Paul, Lankford, Enzi, Hoeven, Daines, McCaskill, Tester, Heitkamp, Peters, Hassan, and Harris. Consistent with Committee Rule 11, the Committee reports the bill with a technical amendment by mutual agreement of the Chairman and Ranking Member.

IV. SECTION-BY-SECTION ANALYSIS OF THE BILL, AS REPORTED

Section 1. Short title

This section establishes the short title of the bill as the “Open, Public, Electronic, and Necessary Government Data Act,” or the “OPEN Government Data Act.”

Sec. 2. Findings

This section highlights the value of Government data as a national resource, and the importance of maximizing the usefulness of the data while sparking innovation and making the Government more efficient. It also stresses that the Government has a responsibility to be transparent and accountable to its citizens. Because of this responsibility, the Government should release its data in an open format that will help agencies identify and reduce waste.

Sec. 3. Definition

This section defines the term “agency.”

Sec. 4. Open Government data

This section adds a new subchapter III to Chapter 35 of title 44, United States Code, to establish a default position for the Federal Government that its data assets shall be published in a machine-readable open format and under an open license to the extent practicable and when not otherwise prohibited by law.

New section 3561 provides definitions for the terms agency, data, data asset, director, enterprise data inventory, information resources management, information system, information technology, machine-readable, metadata, nonpublic data asset, open Government data asset, open license, and public data asset.
New section 3562 requires that all open government data be machine-readable and, to the extent practicable, available under an open license. It also encourages agencies to collaborate with nongovernmental entities and individuals to examine innovative ways to use public data assets.

New section 3563 requires each agency head, in consultation with the OMB Director, to develop and maintain an enterprise data inventory and lays out the required contents of that inventory. Each agency’s CIO is tasked with making the inventory publicly available in an open format and under an open license on Data.gov, and updated regularly. This section also requires the Director of OMB to establish inventory standards for the agencies to use, and permits data assets that cannot be made public to be maintained in a nonpublic section of the inventory.

New section 3564 outlines general information resource management practices that each agency is required to meet and requires agencies to follow guidelines in disseminating information online.

New section 3565 details the responsibilities of each agency’s CIO with regard to the Enterprise Data Inventory and data asset management.

New section 3566 of subchapter III requires the General Services Administration (GSA) Administrator to maintain a single online public interface to publicly share the Government data assets.

Sec. 5. Evaluation of agency analytical capabilities

Section 5 requires each agency’s CIO to conduct a review that assesses the quality, methods, effectiveness, and independence of the agency’s evaluation research and analysis efforts three years after enactment and submit this report to the Committee on Homeland Security and Governmental Affairs of the Senate, the Committee on Oversight and Government Reform of the House of Representatives, and the OMB Director. Additionally, this section requires the GAO to issue a report to Congress not later than four years after the bill’s enactment that summarizes each agency’s findings, highlights trends across agencies, and recommends further actions.

Sec. 6. Online repository and additional reports

Section 6 requires the OMB to work with the GSA and the Office of Government Information Services to maintain an online repository of tools and best practices to facilitate open government data. This section also requires GAO to complete a report no later than three years after the enactment of the bill to report on the value of information made available and whether it makes sense to expand the information accessible.

Finally, section 6 requires a biennial report by OMB and an annual report by the CIO of each agency regarding compliance with the bill to be submitted to the Committee on Homeland Security and Governmental Affairs of the Senate and the Committee on Homeland Security of the House of Representatives.

Sec. 7. Rule of construction

This section ensures that bill not be read to require the disclosure of information or records that are exempt from disclosure under the “Freedom of Information Act.”
Sec. 8. Effective date

This section establishes an effective starting date of 180 days after the bill is enacted.

V. EVALUATION OF REGULATORY IMPACT

Pursuant to the requirements of paragraph 11(b) of rule XXVI of the Standing Rules of the Senate, the Committee has considered the regulatory impact of this bill and determined that the bill will have no regulatory impact within the meaning of the rules. The Committee agrees with the Congressional Budget Office’s statement that the bill contains no intergovernmental or private-sector mandates as defined in the Unfunded Mandates Reform Act (UMRA) and would impose no costs on state, local, or tribal governments.

VI. CONGRESSIONAL BUDGET OFFICE COST ESTIMATE


Hon. Ron Johnson, Chairman,
Committee on Homeland Security and Governmental Affairs,
U.S. Senate, Washington, DC.

Dear Mr. Chairman: The Congressional Budget Office has prepared the enclosed cost estimate for S. 760, the OPEN Government Data Act.

If you wish further details on this estimate, we will be pleased to provide them. The CBO staff contact is Matthew Pickford.

Sincerely,

Keith Hall.

Enclosure.

S. 760—OPEN Government Data Act

S. 760 would direct federal agencies to publish all data they collect in an open format that can be used by any computer. Under the bill, the Office of Management and Budget (OMB) would establish an inventory of all federal data sets and would direct the General Services Administration (GSA) to maintain an online interface for all such data. The bill also would require the Government Accountability Office (GAO), OMB, and Chief Information Officers at each federal agency to report to the Congress about this effort.

Information from the General Services Administration and selected agencies suggest that most of the provisions of the bill would codify Executive Order 13642 and other executive branch policies that set the framework for agencies to promote openness and interoperability in information management. That executive order requires agencies to standardize data sets and to make them publicly available. A website (www.data.gov) has been established to share this government information with the general public. However, CBO expects that implementing S. 760 would cost about $2 million over the 2018–2021 period, for additional administrative and reporting costs for GSA and other agencies and to implement the new reporting requirements for GAO; such spending would be subject to the availability of appropriated funds.

The bill could affect direct spending by agencies not funded through annual appropriations; therefore pay-as-you-go procedures
apply. CBO estimates, however, that any net increase in spending by those agencies would be negligible. Enacting S. 760 would not affect revenues.

CBO estimates that enacting S. 760 would not increase net direct spending or on-budget deficits in any of the four consecutive 10-year periods beginning in 2028.

S. 760 contains no intergovernmental or private-sector mandates as defined in the Unfunded Mandates Reform Act and would not affect the budgets of state, local, or trial governments.

The CBO staff contact for this estimate is Matthew Pickford. The estimate was approved by Theresa Gullo, Assistant Director for Budget Analysis.

VII. CHANGES IN EXISTING LAW MADE BY THE BILL, AS REPORTED

In compliance with paragraph 12 of rule XXVI of the Standing Rules of the Senate, changes in existing law made by the bill, as reported, are shown as follows: (existing law proposed to be omitted is enclosed in brackets, new matter is printed in italic, and existing law in which no change is proposed is shown in roman)

UNITED STATES CODE

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

TITLE 44—PUBLIC PRINTING AND DOCUMENTS

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

CHAPTER 35—COORDINATION OF FEDERAL INFORMATION POLICY

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Subchapter III—OPEN Government Data

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

SUBCHAPTER III—OPEN GOVERNMENT DATA

3561. Definitions.
3562. Requirements for Government data.
3563. Enterprise Data Inventory
3564. Federal agency responsibilities.
3565. Additional Chief Information Officer responsibilities.
3566. Technology portal.

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

SEC. 3561. DEFINITIONS.

As used in this subchapter—

(1) the term “agency”—

(A) has the meaning given the term in section 3502; and

(B) includes the Federal Election Commission;

(2) the term “data” means recorded information, regardless of form or the media on which the data is recorded;

(3) the term “data asset” means a collection of data element or data sets that may be grouped together;

(4) the term “Director” means the Director of the Office of Management and Budget;
(5) the term "Enterprise Data Inventory" means a data inventory developed and maintained pursuant to section 3563;
(6) the terms "information resources management", "information system" and "information technology" have the meaning given those terms in section 3502;
(7) the term "machine-readable" means a format in which information or data can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost;
(8) the term "metadata" means structural or descriptive information about data such as content, format, source, rights, accuracy, provenance, frequency, periodicity, granularity, publisher or responsible party, contract information, method of collection, and other descriptions;
(9) the term "nonpublic data asset"—
   (A) means a data asset that may not be made available to the public for privacy, security, confidentiality, regulation, or other reasons as determined by law; and
   (B) includes data provided by contractors that is protected by contract, license, patent, trademark, copyright, confidentiality, regulation, or other restriction;
(10) the term "open Government data asset" means a data asset maintained by the Federal Government that is—
   (A) machine-readable;
   (B) available in an open format;
   (C) not encumbered by restrictions that would impede use or reuse; and
   (D) based on an underlying open standard that is maintained by a standards organization;
(11) the term "open license" means a legal guarantee applied to a public data asset that the data asset is made available—
   (A) at no cost to the public; and
   (B) with no restrictions on copying, publishing, distributing, transmitting, citing, or adapting; and
(12) the term "public data asset" means a data asset maintained by the Federal Government that—
   (A) may be released to the public;
   (B) has been released to the public in an open format and is discoverable through a search of Data.gov or any successor to Data.gov; or
   (C) is part of the worldwide public domain or, if necessary, published with an open license.
SEC. 3562. REQUIREMENTS FOR GOVERNMENT DATA.
(a) MACHINE-READABLE DATA REQUIRED.—Open Government data assets made available by an agency shall be published as machine-readable data.
(b) OPEN BY DEFAULT.—When not otherwise prohibited by law, and to the extent practicable, public data assets and nonpublic data assets maintained by the Federal Government shall—
   (1) be available in an open format; and
   (2) be available under open licenses.
(c) OPEN LICENSE OR WORLDWIDE PUBLIC DOMAIN DEDICATION REQUIRED.—When not otherwise prohibited by law, and to the extent practicable, open Government data assets published by or for an agency shall be made available under an open license or, if not
made available under an open license and appropriately released, shall be considered to be published as part of the worldwide public domain.

(d) INNOVATION.—Each agency may engage with nongovernmental organizations, citizens, nonprofit organizations, colleges and universities, private and public companies, and other agencies to explore opportunities to leverage the public data assets of the agency in a manner that may provide new opportunities for innovation in the public and private sectors in accordance with law and regulation.

SEC. 3563. ENTERPRISE DATA INVENTORY.

(a) AGENCY DATA INVENTORY REQUIRED.—

(1) IN GENERAL.—In order to develop a clear and comprehensive understanding of the data assets in the possession of an agency, the head of each agency, in consultation with the Director, shall develop and maintain an enterprise data inventory that accounts for any data asset created, collected, under the control or direction of, or maintained by the agency after the effective date of this section, with the goal of including all data assets, to the extent practicable.

(2) CONTENTS.—Each Enterprise Data Inventory shall include the following:

(A) Data assets used in agency information systems (including program administration, statistics, and financial activity) generated by applications, devices, networks, facilities, and equipment, categorized by source type.

(B) Data assets shared or maintained across agency programs and bureaus.

(C) Data assets that are shared among agencies or created by more than 1 agency.

(D) A clear indication of all data assets that can be made publicly available under section 552 of title 5 (commonly known as the “Freedom of Information Act”).

(E) A description of whether the agency has determined that an individual data asset may be made publicly available and whether the data asset is available to the public.

(F) Nonpublic data assets.

(G) Open Government data assets.

(b) PUBLIC AVAILABILITY.—The Chief Information Officer of each agency shall use the standards issued by the Director pursuant to subsection (c) to make public data assets included in the Enterprise Data Inventory of the agency publicly available in an open format and under an open license.

(c) STANDARDS FOR ENTERPRISE DATA INVENTORY.—The Director shall issue standards for an Enterprise Data Inventory, including—

(1) a requirement that an Enterprise Data Inventory include a compilation of metadata about agency data assets; and

(2) criteria that the head of each agency shall use in determining whether to make a particular data asset publicly available in a manner that takes into account—

(A) the expectation of confidentiality associated with an individual data asset;

(B) security considerations, including the risk that information in an individual data asset in isolation does not pose a security risk but when combined with other available information may pose such a risk;
(C) the cost and benefits to the public of converting the data into a manner that could be understood and used by the public;

(D) the expectation that all data assets that would otherwise be made available under section 552 of title 5 (commonly known as the “Freedom of Information Act”) be disclosed; and

(E) any other considerations that the Director determines to be relevant.

(d) NonPublic Data Assets.—Nonpublic data assets included in an Enterprise Data Inventory may be maintained in a nonpublic section of the inventory.

(e) Availability of Enterprise Data Inventory.—The Chief Information Officer of each agency
   (1) shall make the Enterprise Data Inventory of the agency available to the public on Data.gov;
   (2) shall ensure that access to the Enterprise Data Inventory of the agency and the data contained therein is consistent with applicable law and regulation; and
   (3) may implement paragraph (1) in a manner that maintains a nonpublic portion of the Enterprise Data Inventory of the agency.

(f) Regular Updates Required.—The Chief Information Officer of each agency shall—
   (1) to the extent practicable, complete the Enterprise Data Inventory for the agency not later than 1 year after the date of enactment of this section; and
   (2) add additional data assets to the Enterprise Data Inventory for the agency not later than 90 days after the date on which the data asset is created or identified.

(g) Use of Existing Resources.—When practicable, the Chief Information officer of each agency shall use existing procedures and systems to compile and publish the Enterprise Data Inventory for the agency.

SEC. 3564. FEDERAL AGENCY RESPONSIBILITIES.

(a) Information Resources Management.—With respect to general information resources management, each agency shall—
   (1) improve the integrity, quality, and utility of information to all users within and outside the agency by—
      (A) using open format for any new open Government data asset created or obtained on or after the date that is 1 year after the date of enactment of this section; and
      (B) to the extent practicable, encouraging the adoption of open format for all open Government data assets created or obtained before the date described in subparagraph (A); and
   (2) in consultation with the Director, develop an open data plan that, at a minimum and to the extent practicable—
      (A) requires the agency to develop processes and procedures that—
         (i) require each new data collection mechanism to use an open format; and
         (ii) allow the agency to collaborate with non-Government entities, researchers, businesses, and private citi-
zens for the purpose of understanding how data users value and use open Government data assets; 
(B) identifies and implements methods for collecting and analyzing digital information on data asset usage by users within and outside of the agency, including designating a point of contact within the agency to assist the public and to respond to quality issues, usability, recommendations for improvements, and complaints about adherence to open data requirements; 
(C) develops and implements a process to evaluate and improve the timeliness, completeness, accuracy, usefulness, and availability of open Government data assets; 
(D) requires the agency to update the plan at an interval determined by the Director; 
(E) includes requirements for meeting the goals of the agency open data plan including technology, training for employees, and implementing procurement standards, in accordance with existing law, that allow for the acquisition of innovative solutions from the public and private sectors; and 
(F) prohibits the dissemination and accidental disclosure of nonpublic data assets.

(b) INFORMATION DISSEMINATION.—With respect to information dissemination, each agency—
(1) shall provide access to open Government data assets online;
(2) shall take the necessary precautions to ensure that the agency maintains the production and publication of data assets which are directly related to activities that protect the safety of human life or property, as identified by the open data plan of the agency required under subsection (a)(2); and 
(3) may engage the public in using open Government data assets and encourage collaboration by—
(A) publishing information on open Government data assets usage in regular, timely intervals, but not less frequently than annually;
(B) receiving public input regarding priorities for the analysis and disclosure of data assets to be published;
(C) assisting civil society groups and members of the public working to expand the use of open Government data assets; and
(D) hosting challenges, competitions, events, or other initiatives designed to create additional value from open Government data assets.

SEC. 3565. ADDITIONAL CHIEF INFORMATION OFFICER RESPONSIBILITIES.

The Chief Information Officer of each agency, or other appropriate official designated by the head of an agency, is responsible for—
(1) data asset management, format standardization, sharing of data assets, and publication of data assets for the agency; 
(2) the compilation and publication of the Enterprise Data Inventory for the agency required under section 3563; 
(3) ensuring that agency data conforms with open data best practices;
(4) engaging agency employees, the public, and contractors in using open government data assets and encouraging collaborative approaches to improving data use;

(5) supporting the agency Performance Improvement Officer in generating data to support the function of the Performance Improvement Officer described in section 1124(a)(2) of title 31;

(6) reviewing the information technology infrastructure of the agency and the impact of the infrastructure in making data assets accessible to reduce barriers that inhibit data asset accessibility;

(7) ensuring that, to the extent practicable, the agency is maximizing its own use of data, including data assets used in agency information systems (including program administration, statistics, and financial activity) generated by applications, devices, networks, facilities, and equipment, categorized by source type, and such use is not otherwise prohibited, to reduce costs, improve operations, and strengthen security and privacy protections; and

(8) identifying points of contact for roles and responsibilities related to open data use and implementation as required by the Director.

SEC. 3566. TECHNOLOGY PORTAL.

(a) DATA.GOV REQUIRED.—The Administrator of General Services shall maintain a single public interface online as a point of entry dedicated to sharing open government data assets with the public.

(b) COORDINATION WITH AGENCIES.—The Director shall determine, after consultation with the head of each agency and the administrator of General Services, the method to access any open government data assets published through the interface described in subsection (a).