US Open Data Policy: Advances and Recommendations

Luis F. Luna-Reyes
University at Albany
135 Western Avenue
USA
lluna-reyes@albany.edu

ABSTRACT

The Federal Government in the United States has a long tradition in promoting citizen access to information. In the last years, and promoted by strong support from the Executive, Federal Government Agencies have engaged in understanding, cataloguing and publishing their data assets. Beyond the exercise of cataloguing data assets, Federal Agencies have also worked on specific projects in their own domain, and sometimes across domains or national borders. The main purpose of this paper is to assess current progress of the Open Government Data (OGD) policy in the US Federal Government, identify enablers and current challenges, also providing some recommendations to move forward with the vision of OGD.

CCS CONCEPTS

• **Information systems~Digital libraries and archives** • *Information systems~Electronic data interchange*

KEYWORDS

Open government data, information policy, case studies

ACM Reference format:

L. F. Luna-Reyes. 2018. US Open Data Policy: Advances and Recommendations. In *Proceedings of the 19th International Conference on Digital Government Research, Delft, The Netherlands, June 2018 (dg.o '18),* 10 pages. DOI: 10.1145/3209281.3209308

1 INTRODUCTION

Since President Obama issued his executive order for *Making Open and Machine Readable the New Default for Government Information* on May 2013, Federal Agencies in the US have made available to the public about 200,000 datasets through the data.gov portal. The main rationale behind the executive order was –in the same spirit of the Memorandum on Open Government—to promote and environment of transparency, collaboration and participation. In the specific case of open data, collaboration and participation would facilitate innovation among entrepreneurs in developing services and solutions to problems by new uses of data. In this way, government agencies, working together with entrepreneurs, developers and final users would constitute an innovation ecosystem that would produce new products and services, promote economic development, and contribute to the solution of important problems affecting US Citizens.[4, 10, 11, 13] Many actions followed the executive order since 2013, resulting in some important successes and key policies, but also facing many different challenges.

This paper has the purpose of offering an assessment of the current progress in the Open Data Policy in the US, as well as a reflection on the main challenges faced by key stakeholders in producing such results. In this way, this paper offers an assessment of the current success of the Open Data Policy, and provides a set of recommendations for the coming years. The questions guiding the research are the following: what are the current results of the open data policy in the US? and what are the main recommendations to continue and strengthen this policy?

To respond to these questions, this research in progress considers the case of the US Department of the Treasury as it relates to the efforts of the Office of Management and Budget. The paper is organized in 6 sections including this introduction. Section 2 includes a brief introduction to important concepts and principles that will be used as a framework to organize the analysis and recommendations. Section 3 of the paper includes the research approach. Section 4 introduces the main findings of the research. The fifth section includes a comparison and a discussion of the cases presented in the paper. The last section includes conclusions and recommendations.

2 SETTING THE STAGE FOR AN OPEN DATA POLICY

Opening Government Data (OGD) is an important component of government's information policy. In order to set the stage for

dg.o '18, May 30-June 1, 2018, Delft, Netherlands

ACM ISBN 978-1-4503-6526-0/18/05…$15.00
https://doi.org/10.1145/3209281.3209308

the analysis of the current state and recommendations for the future of OGD in the US, I would like to start by discussing both the nature of information policy and the nature of OGD. Concepts briefly presented in this section are intended to establish a framework to analyze experiences in the US Federal Government.

2.1 The Nature of Information Policy

Information Policy, in general, refers to plans, strategies, practices, laws and regulations to control the creation, processing, transportation, distribution, use and destruction of information both in the private and the public sectors.[1] Information policy is very important for most government activity, and it also has an important impact on the Economy and Democracy. Information policies may take the form of laws and regulations. However, there are also many less visible components of information policy that include executive orders, administrative memos or practices that de-facto define ways in which information is created, stored, distributed or used. Government information policy can be categorized in three ideal types: Value-oriented policies (operationalize fundamental principles of information flows in society, e.g. freedom of information), instrumental policies (employ information as means to achieve other policy goal, e.g. environmental policy), and managerial policies (specify rules and procedures for managing information, e.g. information security) [3]. Moreover, different information policies interact among each other making interpretation and application harder. For example, national security policies are usually in conflict with open government and open data policies, promoting secrecy instead of openness. Data management practices have a large impact in OGD policies because of their impact on data quality and availability. Finally, information policy has important impacts on both the structure of society, as well as in the ways in which society changes and innovates. Information policies such as OGD affect and are affected by current technical and social structures.

Technical structures constitute one important type of structure affecting and being affected by information policy. Technical structures include technical platforms, software, hardware and standards that both constrain and enable ways in which information can be made accessible and exchanged among actors and stakeholders of different aspects of our current lives. Basic exchange of data over the Internet, for example, are enabled by many hardware and data exchange protocols and standards to enable hardware interoperability and data exchange. Additionally to these basic protocols, we have also developed higher level standards such as SQL (Standard Query Language), which constitutes a standard way of storing and asking questions to a database, further facilitating information flows and exchange. At an even higher level, we have developed some domain specific standards, such as XBRL (eXtensible Business Reporting Language) [1] or HL7 (Health Level Seven) [2] to further push information exchange for particular business purposes. XBRL, for example, has become the reporting standard for public companies, making the analysis of financial results and portfolio development a much more computer intensive task. HL7, on the other hand, constitutes the backbone for US electronic health records exchange. Government information policy has an impact in this level of structure through offering incentives and regulations to develop and to promote adoption of such standards. Some governments, for example, promote the use of open vs. proprietary standards in the deployment of government services and applications. Technical structures, however, enable (or constrain) government information policy. Many OGD projects, for instance, are enabled (or challenged) by the existence of current technologies and standards (see **Figure** 1).

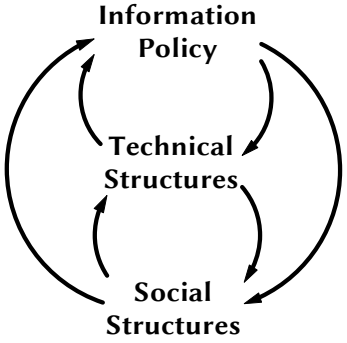

Figure 1. Information Policy as a Socio-Technical System

Standard OGD applications, however, are usually developed in the context of social systems. Main components of these social systems include social actors, the connections among themselves, as well as their activity and practices. Information policy emphasizes in considering stakeholders, as well as their values, interests and power positions. One of the main characteristics of information policy issues is that they many times involve enduring conflict that it is not likely to be fully solved. Think for example on the very same issue of opening government data, there is a continued contention about what are the data that can be fully open, what data represents risks of privacy violations and need to be curated in a special way before making it public or what data represents a risk for national security. These discussions are not likely to be finished soon, and different actors and stakeholders will be continuously pushing the conversation in the direction aligned to their values and interests [5]. Usually, stakeholder networks imply some formal or informal form of governance, understood as basic rules for participation in the network, as well as basic principles and rules for decision making. The last component that I would like to emphasize in the social structure involves the practices and processes of government and other stakeholders. These practices and processes are a key component of the social system that works together with the technical system to produce the outcomes of the policy. In the case of OGD, for example, data management practices in federal agencies cooperated with technical platforms and standards to

[1] XBRL is a widely adopted standard for exchanging business information and data. Public companies in the US and other countries are required to report their results using this standard. See https://www.xbrl.org/.

[2] HL7 is the most widely adopted Health data and information standard in America. It is used both to exchange medical records, and also to coordinate management, evaluation, invoicing and payments. See http://www.hl7.org/.

develop and launch the data.gov portal in the US. The interactions between the technical and social system are influenced by current information policy as it is shown in **Figure** 1. That is to say, the Executive Order from 2013 that asked for "Making Open and Machine Readable the New Default for Government Information" is a good example of a value-oriented information policy that enabled the creation of data.gov, the technical platform that serves as a repository of OGD in the US. The development of such platform, however, was also the result of continuous negotiations among government actors, NGOs, researchers and other OGD users and existing technical possibilities and capabilities.

2.2 The Nature of Open Government Data

Open Government Data is a particular case of government value-oriented information policy that has the purposes of improving government services, promoting innovation and scientific development, as well as promoting economic growth. Data is conceptualized as a key resource or asset that will create value through their use and re-use. Open government data is commonly defined as "data produced with public resources and made publicly available with a license that allows for re-use and re-packaging in innovative applications" [8]. One widely accepted definition of OGD included initially 8 principles.[3] That is to say, OGD needs to be complete (bulk datasets), primary (collected at the source with no aggregation), timely (as quickly as to preserve its value), accessible (through the Internet), machine processable (reasonably structured), non-discriminatory (available to anyone), non-proprietary (using non-proprietary formats), and license-free (no intellectual property attached). Compliance to all these principles should be verifiable by an independent third-party. OGD practitioners have also identified a "five star" open government data rating to account for the level of openness of data in terms of being available online, in structured formats, usable in open software packages, with web addresses or other uniform resource identifiers to enable users to locate data, and linked to other data to develop applications [9].

It is important to note that "openness" is a characteristic of technological and socio-technical systems, and it is common to read and hear about open standards, open platforms, open source, open innovation as well as open data. A main characteristic of these open systems is that they are usually built around communities. The most successful communities are those that had not only a good project that creates value, but also those that have clear governance structures, which define rules for participation, authority, hierarchy, incentive systems as well the coordination of tasks [2]. In the context of open government data, the term ecosystem has been used to describe such communities. The metaphor of the ecosystem had been used before for information-intensive communities. For example, the Internet ecosystem was described as a "social, technical and material formation shaped through its interactions with technical and management organizations, but also with end users, governments, business, civil society organizations and technical experts," [6] Visions of innovation in these open ecosystems involve what Pollock called data cycles, implying the processing and use of OGD by some members in the ecosystem, and then making the new versions of the data resources available to the community as open data again to be improved and re-used by other members in the system. **Figure** 2 constitutes a simplified view of the main actors and interactions in an Open Government Data ecosystem. At the center of the figure, government, innovators and users get together in a community to take advantage of OGD through applications that create many different types of value, including economic, social, environmental or democratic values among others. The interactions between government and innovators constitute the place where practice innovations occur, developing more democratic, effective or efficient ways of delivering services or creating value for the society. The interactions between innovators and users yield interests and expectations about social problems or opportunities for innovation. Finally, interactions between government and users of OGD are events of public engagement that offer opportunities to obtain direct feedback on programs and policies.

Although every member on the community –or every species in the ecosystem—has an importance on its own, it is also known that there exist keystone species, which are crucial for the sustainability of the ecosystem. In the case of OGD, government is one of these keystone species that play an important role on developing the intentionality of the community, works as a steward of value creation and seeks for mechanisms and models to promote sustainability [6]. Nonetheless, citizen and innovators participation are also important for value creation. Unfortunately, citizen participation and user engagement has been observed to be limited across countries [11].

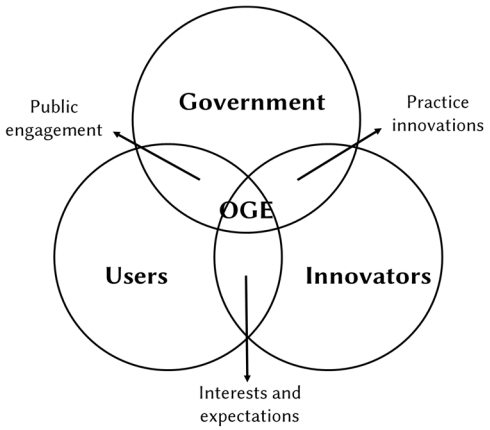

Figure 2. A simplified view of the OGD ecosystem.[Adapted from 6]

3 RESEARCH APPROACH

The research approach involves qualitative analysis associated with case studies [12]. The case studies in this research involve

[3] See https://opengovdata.org/

the main effort lead by the Office of Management and Budget, as well as the work inside the US Department of the Treasury and the US Department of Agriculture. These cases are among the most advanced examples of OGD initiatives at the Federal level in the United States.

Data for the paper included notes and reports from the roundtables organized by the Center for Open Data Enterprise as well as document research gathered through the Internet in news, websites and government documents.

4 FINDINGS: ADVANCES ON OPEN GOVERNMENT DATA IN THE US

The United States has been a global leader in the OGD movement in the last years. President Obama executive orders on open government and open data, as well as his role in coordinating and launching the open government partnership are well-known factors in building the US current position. However, the open government data movement has a longer history in the US information policy both through laws and regulations as well as through executive orders.

Table 1. Selected US Statutes related to Open Government Data

Year	Title
1789	U.S. Constitution
1791	First Amendment to the U.S. Constitution
1873	Congressional Record established
1950	Federal Records Act, amended in 2014
1966	Freedom of Information Act, amended by 1986 Freedom of Information Reform Act, and by 1996 Electronic Freedom of Information Act
1976	Government in the Sunshine Act
1993	Government Performance and Results Act (GPRA)
2001	Data Quality Act
2002	Electronic Government Act of 2002
2014	Digital Accountability and Transparency Act

In terms of Federal Statutes (see **Table** 1), Open Government Data can be traced back to the Constitution and the Bill of Rights, which established a principle of openness by both Congress and the Executive. Moreover, the First Amendment constitute the basis for having access to information, including government information. The Congressional Record, established in 1873, constitutes a daily registry of Congress activity. Nowadays, the Congressional Record is kept online at https://www.congress.gov/congressional-record. The Federal Records Act establishes principles for data management at the Federal level. The last amendment of 2014 explicitly recognizes electronic records as government records. The 60's and the 70's witnessed the publication of several Freedom of Information Acts with the purpose of increase transparency in government. These acts where amended in 1996 to include the use of the Internet in the process. The Government Performance and Results Act was the first effort to introduce data-driven performance management and decision making. Another important statute was the Data Quality Act of 2001, which gives the Office of Management and Budget (OMB) the capacity of developing guidelines to improve data quality in Federal Agencies. The Electronic Government Act of 2002 creates the figure of the Federal Chief Information Officer in the OMB, and establishes a framework to use the Internet to increase citizen access to information and services. The latest approved statute relevant for OGD was the Digital Accountability and Transparency Act (known as DATA), which has the goal of making government expenditures more accessible and transparent. The Open Government Data Act (OGDA) is currently pending approval in the Senate. OGDA gives the force of law to the directive of open government data. The current draft emphasizes on the concept of data as an asset, and the need to keep an inventory of data assets in each agency, providing some basic criteria to decide which data is to be open to the public. The act makes the Director of the OMB the main leader of OGD in the US, working together with heads and CIOs of other agencies in the implementation of the Act. Data.gov remains as the main repository for open data, as well as shared best practices. The Act also emphasizes on data management practices and data quality, as well as analytic capabilities of each agency.

4.1 The Office of Management and Budget

In terms of directives and OMB guidance related to OGD, the history appears to be not as long (see **Table** 2). The selected directives included in the table are exemplars of directives that show the tension between openness and privacy or national security, which are all important social values. I would like also to note at this point that President Obama memoranda and Executive Orders on Open Government and Open Government Data include explicitly many of the principles of OGD in terms of accessibility, timeliness and free-licensing. However, it is also important to note that, although the Open Government Directive (M-10-06) had a wide scope on Transparency, Participation and Collaboration, the Open Data Policy (M-13-13), and Open Data Action Plan of 2014 emphasized on entrepreneurship and economic development, which appears to continue being the main emphasis of current OGD policy [7]. There have been not any additional directives coming from the White House in the last year in this policy issue. OMB has issued a couple of additional documents to offer additional guidance for the implementation of the DATA Act of 2014.

In terms of technology, the OMB has developed a Metadata standard and a web platform to provide links to all open government data that uses the Metadata standard. That is to say, data.gov is continuously harvesting the www, looking for datasets that are stored in agencies websites and then publishing links to all open data looking to improve accessibility and findability of data (see **Figure** 3). At the time of writing this paper, data.gov had about 197,000 datasets. Data.gov also includes a set of APIs that developers can use to harvest metadata and/or to develop applications. Another important resource developed by the Federal Government is the repository of tools and best practices

on open data called Project Open Data.[4] Project Open Data is a community platform that includes OGD principles, tools and best practices in the Federal Government. The portal also hosts a dashboard to follow on current progress and achievements of federal agencies in terms of their data inventories, publication of data, public engagement, privacy and security, human capital and use and impact.[5]

Table 2. Selected Directives and OMB Guidance related to Open Government Data

Year	Title
Ongoing	OMB Circular A-130 Management of Federal Information Resources
1995	EO 12958 Classified National Security Information
1998	Presidential Memorandum on Privacy and Personal Information in Federal Records
1999	Presidential Memorandum on Electronic Government
2001	EO Order 13228 (Establishing the Office of Homeland Security and Homeland Security Council)
2001	EO Order 13231 (Critical Infrastructure Protection in the Information Age)
2002	E-Government Strategy
2005	EO 13392 Improving Agency Disclosure of Information
2009	Presidential Memorandum, Freedom of Information Act
2009	M-10-06 Open Government Directive
2011	EO 13576 Delivering an Efficient, Effective, and Accountable Government
2013	EO 13642 Making Open and Machine Readable the New Default for Government Information
2013	M-13-13 Open Data Policy–Managing Information as an Asset
2014	Open Data Action Plan
2015	M-15-12 Increasing Transparency of Federal Spending by Making Federal Spending Data Accessible, Searchable, and Reliable
2016	MPM-2016-03 Additional Guidance for DATA Act Implementation: Implementing a Data-Centric Approach for Reporting Federal Spending Information

In terms of community, the US Federal government is a member of the Open Government Partnership (OGP), a multilateral initiative that group 75 national governments and 15 sub-national governments.[6] Although members of the OGP are national or local governments that have made a commitment to transparency and openness, the initiative has a Steering committee that includes also representatives of the civil society. Technical developments related to the open data project in the United States follow open standards and work in interaction with global communities promoting such developments. Data.gov, for example, is powered by CKAN and WordPress, two open source applications. The open data project itself used open platforms to document progress across agencies through a dashboard.[7] In an attempt of promoting better governance and exchange of best practices, OMB also created in 2016 the "Data Cabinet," which involved the Chief Data Officers of all Federal Agencies.

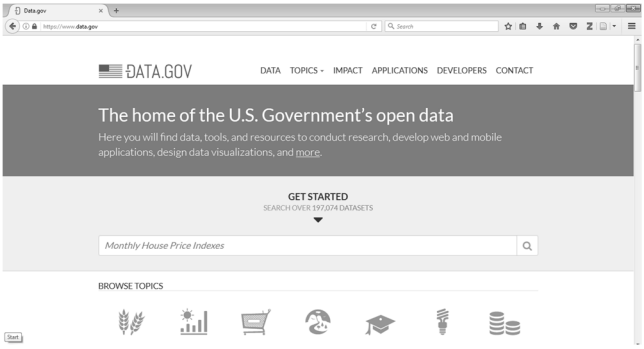

Figure 3. Data.gov, the Federal data repository

4.2 The US Department of the Treasury

The US Department of the Treasury is one of the federal agencies in the US with the longest tradition in opening data. In 1789, the Department of the Treasury published for the first time the Monthly Treasury Statement (MTS). The Monthly Treasury Statement is a summary of the receipts and outlays of the Federal Government. In addition to the MTS, the Department of the Treasury also publishes the Daily Treasury Statement, the Monthly Statement of the Public Debt, the Treasury Bulletin, the Budget of the United States Government and the Combined Statement of Receipts, Outlays, and Balances of the United States Government. A main source of data for all these periodic reports of the Department of the Treasury is the Central Accounting System, which in turn receives information from all Federal Agencies. Some of these reports are mandated by Congress, who uses them to monitor actual spending, find data discrepancies and make better analysis and decisions on budget proposals every year.

Government financial data has the potential of creating value by promoting a more transparent and accountable government, making possible for users and innovators to monitor the different ways in which tax money is being invested and spent. Spending data in particular can be useful for all types of recipients of federal funds, including state and local governments, NGOs and private contractors, who can better understand trends on federal spending, sources of grants and other federal funding.

[4] See https://project-open-data.cio.gov/
[5] See https://labs.data.gov/dashboard/offices
[6] See https://www.opengovpartnership.org/about/about-ogp
[7] See https://project-open-data.cio.gov/ for a general description of the project and https://labs.data.gov/dashboard/offices for the agency dashboard.

The DATA Act is the first open data law in the United States to make federal spending data open. DATA expands the Federal Funding Accountability and Transparency Act of 2006, with the main purposes of (1) disclosing federal agency expenditures linking them to specific programs, (2) establish government-wide standards to report financial information, (3) simplify reporting requirements and (4) improve the quality of the data.[8] The Act makes OMB and the Department of the Treasury the leaders of the effort and the leaders in developing data standards and a platform to make data available to the public. Another important characteristic of the act is that all data published in the platform is required to be partially audited –through a statistical valid sampling method—by the Inspector General of each agency and the Comptroller General of the United States. The Office of Management and Budget has also provided additional guidance on the implementation of the act through M-15-12 "Increasing Transparency of Federal Spending by Making Federal Spending Data Accessible, Searchable, and Reliable" and MPM-2016-03 "Additional Guidance for DATA Act Implementation: Implementing a Data-Centric Approach for Reporting Federal Spending Information." In these memoranda, OMB describes the general and new reporting requirements associated to the act.

At the technical level, and following the mandate of the DATA Act, OMB has successfully developed the DATA Act Information Model Schema (DAIMS), which is a set of standard definitions that are already being used by federal agencies to report their spending information at the USASpending.gov platform (see Figure 4). The development of both platforms has been a participatory process involving feedback from federal agencies and open to feedback from any user through the platforms themselves. The community around this project includes governments, NGOs and private companies that are recipients of federal funds. This community of users has been the one that has been reached out for development purposes.

Figure 4. Beta USASpending.gov site. Current working site can be visited at www.usaspending.gov

Main challenges identified by participants in these developments include:

1. Making data more discoverable and accessible.- A challenge in making financial data understandable for a variety of users, and the need to develop basic applications and filters to facilitate navigation through the data.
2. Improving data quality and relevance.- A key challenge of any complex system that gathers data from a distributed group of actors consists of having complete and timely data input, key characteristics of data quality.
3. Making datasets interoperable.- Financial data in the federal government is processed through many different systems, and value of data is increased by linking this disparate sources.
4. Engaging with data users.- Communicating with key users about portal requirements, and finding effective ways of getting their feedback has been perceived as a challenging task.

4.3 The US Department of Agriculture

The Department of Agriculture (USDA) was founded in 1862 by President Lincoln with a very wide mission in the areas of food, agriculture, economic development, science, and natural resource conservation. USDA research and reporting has played an important role in supporting universities, research centers and farmers in developing better agricultural practices and innovation. Data and information resources associated with the mission areas of the US Department create value for the development of rural areas, increasing food safety as well as resiliency in case of natural disasters. Agricultural research data has a long tradition of being transformed into practice in the US through the extension function of the land-grant universities in the US, which were created in the late 19th Century. The USDA has played a key role in promoting research and extension since the beginning through the National Institute of Food and Agriculture (NIFA) and all its preceding agencies.

Data gathering and dissemination initiatives at the Department of Agriculture have been motivated by a series of laws and regulations related to the land-grant institutions such as the Smith-Lever Act of 1914, which establishes the extension services in land-grant universities, and the USDA Reorganization Act that creates NIFA in 1994. More recently, just like any other Federal Agency, open data initiatives respond to executive orders of President Obama. The Department of Agriculture OGD activities have been also oriented by President Obama participation and commitments in the development of the Global Open Data for Agriculture and Nutrition (GODAN), an international initiative to promote open data to ensure world food security.[9]

The USDA has developed a series of technical platforms to share OGD relevant for its mission. Three platforms are at the core of their OGD program. The first platform, named Discovery Tool for New Farmers,[10] is a personalized service to get access to information, financial and data resources for new farmers across

[8] The full text of the act can be found at https://www.govtrack.us/congress/bills/113/s994/text

[9] See http://www.godan.info/about

[10] See https://newfarmers.usda.gov/discovery

the United States. The main rationale of the platform is that most farms in the US will go through an ownership transition in the coming years, and the tool is a form of proactively engaging with the future generation of farmers in the USA. A second technical platform for open data is the National Agricultural Statistics Service (NASS).[11] NASS gathers, organizes and publishes data in service to the US agriculture through surveys and census. Finally, the National Agricultural Library (NAL) has released a beta version of the AG Data Commons (see Figure 5). Ag Data Commons is a collaborative repository of datasets from funded research that uses open standards and platforms.

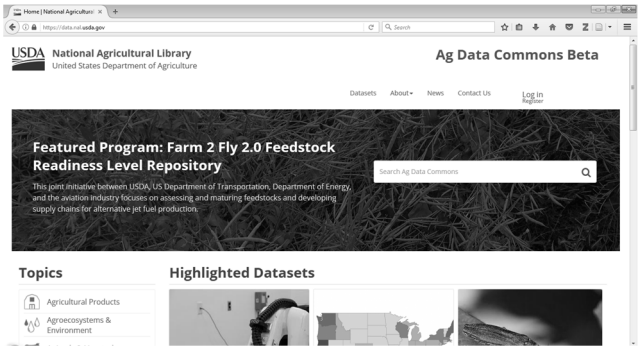

Figure 5. Ag Data Commons at the National Agricultural Library.

These three platforms also suggest the type of community engagement involved in the USDA OGD initiatives. The discovery tool represents engagement with farmers and final users of USDA data, and the Ag Data Commons reflects engagement with the research community and other funding agencies. Each platform is at the center of two different types of OGD ecosystems. It is also important to note that USDA is involved in an international community of open data through the Global Open Data for Agriculture and Nutrition. Some key challenges for USDA are the following:

1. Data volume and velocity.- The amount of data and research being produced in the area pose a challenge to keep it timely and updated.
2. Data digitalization.- There is a number of data resources in older media that require to be digitized and made available in machine-readable formats.
3. Data interoperability.- Preparing data to be linked with data in different datasets in a meaningful way.

5 DISCUSSION

The cases introduced in the previous sections represent two examples of the type of policies, projects and communities that have been formed or strengthened through the OGD information policy in the United States. The first important element to note is that OGD is not necessarily a new movement, but a new instance of a movement that charges government to gather, manage and disseminate data to facilitate an information environment that improves policy and decision making. Some of these practices can be traced back to the second part of the 19th century. In addition, OGD initiatives can be also traced back to movements that promote a more transparent, accountable and democratic government. The last components of the US Information Policy, however, emphasize on applications to promote economic development and government efficiencies, giving much less importance to other social and democratic values that –in the history of the US—have been associated with this movement.

There are common features to both projects introduced in the previous section (see Table 3), the most obvious are related to the importance of EO 13642, M-13-13 and other commitments from the President for all projects. In addition, leadership from OMB, and the National CIO have been also key in the implementation of the directives. The Project Open Data platform is definitively an excellent example of the ways in which OMB and the CIO keep leading and following-up on the national initiative. Another commonality is related to the interest in using open software, open standards and open APIs. To different extent in both cases, but there is also an interest in engaging with relevant communities for each project.

Table 3. Key Dimensions of OGD

Agency / Dimension	OMB	Department of the Treasury	Department of Agriculture
Information Policy	EO 13642 and M-13–13	DATA Act and OMB Guidance Memoranda	EO 13642, Smith-Lever Act and GODAN commitments
Data Standards	Open Metadata Schema	DAIMS	Diverse data standards for different projects
Technology Platform	CKAN and Wordpress, Project Open Data	USASpending.gov	Diverse platforms including DKAN and Wikis, APIs
Community	US Federal Agencies	Federal agencies and recipients of federal funds	Farmers, researchers and the international community
Distinguished Feature	Pending approval of the Open Government Data Act	Supported by specific regulations and standards under a clear leadership and governance	User-centered platform development

[11] See https://www.nass.usda.gov/

A key finding from the description of the projects above is that the OGD project is not a monolithic program, but a manifestation of several communities and groups that are interested around topics and problem areas that can be better tackled using data. For example, USASpending.gov is a community of federal agencies, recipients of federal funds, policy makers, watchdog organizations and citizens, interested on different aspects of US Federal spending. The DATA Act, standards and data in USASpending.gov, absolutely key for this community, have little relevance to the wider community represented in data.gov, which is mostly a community of federal agencies around a platform to catalogue and publish OGD. In this sense, the OGD Ecosystem can be understood as the aggregate smaller communities. Given the diversity of datasets that can be made public in each federal agency and across all US Federal government, selecting key datasets constitute a key strategic exercise that requires leadership and citizen engagement. There are a couple of NGOs that have played a key role in promoting some of these conversation in close collaboration to OMB, the GovLab,[12] the Center for Open Data Enterprise,[13] and the Sunlight Foundation.[14]

As it is shown in **Table** 3, projects and communities also are different in some key features or characteristics. Some of them, for example, are related to problems that transcend in one way or another national boundaries. For example, the Global Open Data for Agriculture and Nutrition community includes governments and private organizations interested in improving food safety in the world. Projects also differ in how agencies distribute the effort, some of them concentrate on a single project, and some others have a number of ongoing projects, some allocate more effort to community building, some others on platform development, some others in developing guidelines, plans and legislations. All of these tasks are relevant for the development of a healthy OGD ecosystem.

Finally, Federal Agencies in the US share a number of challenges that need to be overcome:

1. Data gathering and management.- Quality of published data is a reflection of data gathering and management practices. Continuous improvement and updating of these techniques is a challenge for every agency.
2. Human capital.- Data gathering, management and analysis requires a work force with a combination of technology and data analysis skills scarce in the market.
3. Data quality.- Establishing and implementing processes to ensure data timeliness, accuracy and completeness.
4. Data integration and interoperability.- Linking data increases value, but requires effort on developing standards and curating data.
5. Lack of resources.- No new resources to manage OGD projects has been added to any Federal Agency. At least one of the projects reported here did not advance because of lack of resources.
6. Developing and updating standards.- Data and metadata standards are not easy to develop and enforce in environments with multiple stakeholders. Moreover, in changing environments, standards need to be continuously updated.
7. Making data more discoverable and accessible.- Making data available in a way that a variety of users can find and use, developing the right filters and visualizations to facilitate navigation.
8. Engaging with data users.- Communicating with key users about data requirements and applications, as well as finding effective ways of getting their feedback has been perceived as a challenging task.
9. Data volume and velocity.- The amount of data and research being produced in all different projects reported pose a challenge to keep it timely and updated.
10. Data digitalization.- There is a number of data resources in older media that require to be digitized and made available in machine-readable formats.

6 CONCLUSIONS AND RECOMMENDATIONS

In this final section of the paper, I would like to briefly answer the questions guiding the document: what are the current results of the open data policy in the US? and what are the main recommendations to continue and strengthen this policy?

The findings section of this document provides detailed information on the first and the second questions. OGD information policy in the US consists of a set of acts and formal regulations, as well as a set of administrative and executive orders (see Tables 1 and 2). Many Federal Agencies have taken the OGD policy and incorporated it in a long-standing mission of gathering and publishing information for both policy makers and the public. In this sense, each Federal Agency has its own tradition and policy with regards of opening information and data.

As of today, the OGD information policy in the US has promoted the development of an ecosystem in which Federal Agencies have contributed in the development of a shared repository of data resources, data.gov. The Director of OMB, as well as the Federal Chief Information Officer have played a key leadership role in the development of the necessary technology standards for the development of data.gov, and the General Services Administration have successfully managed the data portal. In the development of this repository, the US Federal Government has consistently favored open standards and open source software. In terms of governance, a Data Cabinet, formed by Chief Data Officers of Federal Agencies was formed. Although the Open Government Policy has been guided by principles of Transparency, Collaboration and Participation, the OGD policy has had a clear bias towards innovation and economic development as main goals.

Besides their contribution to the data.gov repository and their inventories of data assets, each federal agency has engaged in specific OGD projects around specific user communities. The exemplars included in this report suggest that projects share common features, such as the preference for the use of open

[12] See http://www.thegovlab.org/
[13] See http://opendataenterprise.org/
[14] See https://sunlightfoundation.com/

standards and open platforms. Moreover, all the introduced exemplars have also as a characteristic the focus on a specific community. Communities in these projects are most of the time communities of data owners and data providers that are trying, with different levels of success, to engage with users and innovators. In some cases, users and innovators are at the center of the design of technology platforms and systems. In a sense, findings suggest that the OGD ecosystem is in fact an aggregate of smaller communities engaging in specific domains and topics. Findings from the cases suggest that current progress is related to the following main Enablers and Challenges.

6.1 Main Enablers:

1. Information policy.- Regulations, acts and executive and administrative orders have played an important role enabling change. Although acts such as FOIA or the Data Quality Act provide important basis for OGD, M-13-13 and EO 13642 have been at the center of all OGD policy in the Federal Government. The DATA Act has also played a clear enabling role in the case of opening spending data.
2. Leadership.- Direct support and commitment from the President and the Director and CIO at OMB constitute a second important enabler by setting a vision as well as specific basis for implementation and project management.
3. Technology development.- Current progress on the OGD policy are possible because technology developments that allow for the creation of standard vocabularies and schemas, as well as platforms that allow for making data widely available.
4. Civil society involvement.- Particularly the leadership of few NGOs who engage with government in promoting community engagement in OGD projects such as the GovLab, the Center for Open Data Enterprise, and the Sunlight Foundation.

6.2 Main Challenges:

1. Data management.- Improving and controlling processes of data gathering and curation as well as improving data quality practices.
2. Resources.- Besides the lack of new resources to respond the open data initiative, data management and analysis require a work force with skills scarce in the market.
3. Data integration and interoperability.- Developing and updating technical standards and definitions to allow linking OGD as well as digitizing data resources existing in older formats.
4. Making data more discoverable and accessible.- Making data available in a way that a variety of users can find and use, developing the right filters and visualizations to facilitate navigation, and keeping pace with the volume of data generated every day.
5. Engaging with data users.- Communicating with key users about data requirements and applications, as well as finding effective ways of getting their feedback has been perceived as a challenging task.

Finally, I would like to conclude with a set of recommendations for the current administration to continue and reinforce the Open Government Data Program in the United States:

1. Finalize approval of the Open Government Data Act.- Although this is really in hands of Congress, Executive leadership and support for the policy may ease the approval process. OGDA elevates to the status of law many of the contents of the current executive and administrative orders that are fueling the Open Government Data program in the Federal Government.
2. Confirm support and leadership.- Although players at OMB are committed to the OGD policy and program, confirming presidential support to OGD is key for the success of the program. This support can be shown by confirming/appointing key leadership for the program such as the CIO at OMB. Support from OMB to promote the formal appointment of Chief Data Officers or Chief Data Scientists in all Federal Agencies will also contribute to a stronger leadership.
3. Governance.- Having the formal position of Chief Data Officer/Scientist will also contribute to have a Data Cabinet that will lead information policy and develop a true strategy for open government data. Such a model of governance should also include interactions between the Data Cabinet and communities and ecosystem developed around specific datasets within and across agencies. Roundtables like the ones organized by OMB in collaboration with the Center for the Open Data Enterprise are a good model to continue and improve.
4. Strategy.- Current efforts on OGD at the Federal level in the US has been guided by the DATA Act and a set of Executive and Administrative orders that do not have any clear strategy. As it was suggested by the cases, the Ecosystem is really an aggregate of small ecosystems organized around domain topics. The cases also showed that each Federal Agency has hundreds of data resources that can be used to create an ecosystem around a community of government data providers, users and innovators. Prioritizing and selecting which data sets to open is no trivial task that requires a clear intentionality, as well as clear thinking in terms of sustainability and value creation. A Data Cabinet, under the leadership of OMB maybe the right place to develop such a strategy.
5. Scope of the Information Policy.- Although current emphasis on creating efficiencies and economic value through OGD is a valuable approach, current international trends on data collaborative suggest that a focus on problems on social, democracy or environmental issues are also valuable areas of focus that need to be included in the scope of the US information policy. Such scope needs to be clearly established in the information policy through acts and regulations, but also through executive and administrative

orders. International trends also suggest that value comes from combining data from public and private sources. A revitalization of the Smart Disclosure policy and a more formal incorporation of it in the OGD information policy is also desirable.

6. Citizen engagement and education.- One of the main challenges of current information policy implementation is related to community engagement. International trends also include good practices in citizen education and engagement. Successful examples of such initiatives such as Monithon are the result of combining NGO effort with government support.

ACKNOWLEDGMENTS

The research reported here was partially funded by the IBM Center for the Business of Government. Any opinions expressed in this material are those of the author and do not necessarily reflect the views of IBM.

REFERENCES

[1] Braman, S. 2009. *Change of State: Information, Policy, and Power*. The MIT Press.

[2] Dahlander, L. et al. 2008. Online Communities and Open Innovation. *Industry and Innovation*. 15, 2 (Apr. 2008), 115–123. DOI:https://doi.org/10.1080/13662710801970076.

[3] Dawes, S.S. 2017. Information Policies: Value-Oriented, Instrumental, and Managerial Choices for Governing an Information Society. *Routledge Handbook on Information Technology in Government*. Y.-C. Chen and M.J. Ahn, eds. Routledge Handbooks Online. 27-.

[4] Dawes, S.S. et al. 2016. Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly*. 33, 1 (Jan. 2016), 15–27. DOI:https://doi.org/10.1016/j.giq.2016.01.003.

[5] Galvin, T.J. 1997. Chapter 1: Public information policy: Concepts, definitions, analytical typology. *Rights in Conflict: An Introduction to Public Information Policy*.

[6] Harrison, T.M. et al. 2012. Creating Open Government Ecosystems: A Research and Development Agenda. *Future Internet*. 4, 4 (Oct. 2012), 900–928. DOI:https://doi.org/10.3390/fi4040900.

[7] Howard, A. 2017. Notes from a roundtable on open data at the White House. *Sunlight Foundation*.

[8] Janssen, M. et al. 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*. 29, 4 (Fall 2012), 258–268. DOI:https://doi.org/10.1080/10580530.2012.716740.

[9] Linked Data: 2009. *http://www.w3.org/DesignIssues/LinkedData.html*. Accessed: 2013-08-21.

[10] Pollock, R. 2011. Building the (Open) Data Ecosystem. *Open Knowledge International Blog*.

[11] Styrin, E. et al. 2017. Open Data Ecosystems: An International Comparison. *Transforming Government: People, Process and Policy*. (2017).

[12] Yin, R.K. 2003. *Case Study Research. Design and Methods*. Sage Publications.

[13] Zuiderwijk, A. et al. 2014. Innovation with open data: Essential elements of open data ecosystems. *Information Polity: The International Journal of Government & Democracy in the Information Age*. 19, 1 (Mar. 2014), 17–33. DOI:https://doi.org/10.3233/IP-140329.