

European Public Sector Information Platform

Topic Report No. 25

State of Play:

Public Sector Information in the United States

Author: Timothy Vollmer

Published: 25th February 2011

State of Play: Public Sector Information in the United States

Keywords

United States, public sector information, federal government, government information, copyright, public domain, data, freedom of information, regulatory policy

About the Author

Timothy Vollmer is Policy Coordinator for Creative Commons in the United States, and has worked as a policy fellow, business development assistant, and intern for Creative Commons. Prior to rejoining CC, Timothy was Assistant Director to the Program on Public Access to Information for the American Library Association Office for Information Technology Policy in Washington, D.C. Timothy is a graduate of the University of Michigan School of Information. While at Michigan, he was a research investigator for the Open.Michigan open educational resources initiative, and helped develop a student-centric open courseware publishing pilot there.

Abstract

This topic report examines the background of public sector information (PSI) policy and re-use in the United States, describing the federal, state and local government PSI environments. It explores the impact of these differences against the European framework, especially in relation to economic effects of open access to particular types of PSI, such as weather data. The report also discuss recent developments in the United States relating to PSI re-use, such as Data.gov, the NIH Public Access Policy, and new open licensing requirements for government funded educational resources.

Copyright information

© 2010 European PSI Platform - This document and all material therein has been compiled with great care; however, the author, editor and/or publisher and/or any party within the European PSI Platform or its predecessor projects the ePSIplus Network project or ePSINet consortium cannot be held liable in any way for the consequences of using the content of this document and/or any material referenced therein. This report has been published under the auspices of the European Public Sector Information Platform.



The report is released under the Creative Commons Attribution 3.0 License (<http://creativecommons.org/licenses/by/3.0/>) and may be used providing acknowledgement is made to the European Public Sector Information (PSI) Platform and Creative Commons. The European Public Sector Information (PSI) Platform is funded under the European Commission eContentplus programme.

Contents

1. Executive Summary 4

2. Introduction 4

3. PSI frameworks in the United States..... 5

 a. Federal framework 5

 b. State and local frameworks 9

4. Economic impacts of open access under U.S. federal framework 10

 National Oceanic and Atmospheric Administration 11

5. Recent PSI State of Play in U.S. 13

 a. Data.gov 13

 b. National Institutes of Health Public Access Policy 15

 c. Trade Adjustment Assistance Community College and Career Training Grant Program. 15

 d. Other projects..... 16

6. Conclusion 17

1. Executive Summary

The U.S. federal information framework is constructed primarily through legislation and regulatory policies, including the U.S. Copyright Act, Freedom of Information Act, Office of Management and Budget Circular A-130, and others. The U.S. federal framework is unique in that under Section 105 of the Copyright Act there exists no copyright protection for information created by the federal government. This policy typically does not extend to information created and disseminated at the state and local levels. At the state and local levels, there is a wide range of policy frameworks, and some rely on restrictive information management schemes in order to maintain control and in theory recuperate costs. However, some states and local entities have been proactively making PSI available.

U.S. federal PSI promotes economic activity because of the lack of intellectual property restrictions and non-adherence to strict cost recovery policies. Information created and shared by the National Oceanic and Atmospheric Administration (NOAA), particularly weather information, stimulates economic activity - approximately \$700 million annually - and leads to the creation of other value-added industries.

The U.S. has been active in disseminating and promoting the re-use of its public sector information. Three recent initiatives in the U.S. include Data.gov, the National Institutes of Health Public Access Policy, and the Trade Adjustment Assistance Community College and Career Training Grant Program. These and other projects show a groundswell in government and citizen interest in sharing and re-using PSI.

2. Introduction

The ePSIplatform has published several “state of play” reports discussing PSI (public sector information) activities in European jurisdictions and other countries too.¹ PSI is characterized in different ways and as representing a variety of perspectives. For instance, some view the dissemination and re-use of PSI as a means to increase the transparency and accountability of government. Others view PSI as primarily a means for improving internal government communication and efficiency. Some view PSI as a vehicle for promoting economic activity and innovation. Others are exploring ways for PSI to be used as a means for international diplomacy

¹ These include The Netherlands, Spain, Slovenia, Germany, and Australia. See http://www.epsiplus.net/topic_reports

State of Play: Public Sector Information in the United States

and global information sharing. Some see PSI as civic capital, working to increase citizen participation in government activities.²

In this topic report, we'll discuss the PSI framework within the United States. Next, we'll look at the National Oceanic and Atmospheric Administration as a demonstration of positive economic impact of open access under the U.S. federal framework. Finally, we'll describe some contemporary PSI work in the United States, including Data.gov and other initiatives working to disseminate information and support the re-use of PSI.

3. PSI frameworks in the United States

Within the U.S., public sector information is disseminated from federal, state, county and municipal levels.³ The federal PSI policy environment originates from a variety of locales, including legislative, executive, judicial actions. The most prominent of these pieces include legislation and regulatory policies. We'll explore in more detail legislative components of Section 105 of the U.S. Copyright Act and the Freedom of Information Act, since they provide the backbone to PSI legislative policy at the federal level.⁴ The most important regulatory policy dealing with PSI in the US is the Office of Management and Budget Circular A-130.

a. Federal framework

Copyright Act Section 105

The U.S. Copyright Act is of particular interest due to the unique handling of information produced by the federal government. As described in Section 105 of Title 17, “copyright protection under this title is not available for any work of the United States Government...”⁵

² *The Knight Commission on the Information Needs of Communities in a Democracy* underscored the ways that public information is vital and important to a lively democracy. See <http://www.knightcomm.org/>

³ A recent study by Socrata indicates that U.S. government entities are increasingly turning toward mandated sharing of PSI, noting that 59.5% of federal, 51.9% of state, 43.5% of county, and 56.0% of municipal governments mandate sharing information with the public. *2010 Open Government Data Benchmark Study*. January 2011. Available at <http://www.socrata.com/benchmark-study/>

⁴ Other legislation includes The Sunshine in Government Act, Paperwork Reduction Act, Privacy Act, E-government Act, and others.

⁵ 17 U.S.C § 105, Available at <http://www.copyright.gov/title17/92chap1.html#105>

State of Play: Public Sector Information in the United States

Information created by the U.S. federal government is free to use, for any purpose.⁶ Paul Uhlir, Director of the Board on Research Data and Information at the U.S. National Academies, notes that there are several benefits of releasing information created by the U.S. federal government directly into the public domain, including the following: supporting the transparency of government activities; promoting democratic ideals; encouraging the dissemination of information to enhance public health, safety, and general welfare; and supporting scientific and technical research functions.⁷

However, there are several caveats to Section 105, spelled out in the latter half of the legislative text, "...but the United States Government is not precluded from receiving and holding copyrights transferred to it by assignment, bequest, or otherwise."⁸ For example, Section 105 does not bar the U.S. from exercising copyright over works transferred to it. In addition, Section 105 only applies to U.S. government works, thus does not include works created by contractors. Also, Section 105 prohibits the attachment of copyright under U.S. law, but doesn't preclude a U.S. government work from being protected by copyright under the laws of other jurisdictions. So, the work may only be "free" as a practical matter within the United States. Outside its borders, the U.S. government could assert that, for example, one of its works is protected under French copyright law, and then enforce its copyright in France.⁹ The U.S. government might

⁶ Robert Gellman notes several examples of information created by the U.S. government not eligible for copyright protection, including "the President's annual budget; unemployment and economic statistics; crop reports and agricultural information; the census; financial filings with the Securities and Exchange Commission; the Federal Register, Commerce Business Daily, the Congressional Record; and proposed legislation and agency regulations." *Twin Evils: Government Copyright and Copyright-Like Controls Over Government Information*, 45 Syracuse Law Review 999 (1995), at 2. Available at <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA394923>

⁷ Uhlir describes "public domain" more broadly than its legal definition—the absence of intellectual property rights. He says public domain is "what is left outside the scope of copyright or other forms of statutory protection...[and] also refers to information of an intrinsically public nature; that is, certain types of information that are produced by public authorities ("government" in the broad sense) in the course of their duties, and that are seen as a public good." (v) Within the document referred to here, UNESCO defines public domain information as "publicly accessible information, the use of which does not infringe any legal right, or any obligation of confidentiality. It thus refers on the one hand to the realm of all works or objects of related rights, which can be exploited by everybody without any authorization, for instance because protection is not granted under national or international law, or because of the expiration of the term of protection. It refers on the other hand to public data and official information produced and voluntarily made available by governments or international organizations." (3) Uhlir, Paul F. *Policy Guidelines for the Development and Promotion of Governmental Public Domain Information*, UNESCO, 2004, at v-vi. Available at <http://unesdoc.unesco.org/images/0013/001373/137363eo.pdf>

⁸ 17 U.S.C § 105, Available at <http://www.copyright.gov/title17/92chap1.html#105>

⁹ The legislative report accompanying the Copyright Act specifically contemplates this scenario: "The prohibition on copyright protection for United States Government works is not intended to have any effect on protection of these works abroad. Works of the governments of most other countries are copyrighted. There are no valid policy

State of Play: Public Sector Information in the United States

enforce its copyright in other jurisdictions, for example, as a mechanism to negotiate with other countries for access to their works.

Freedom of Information Act

Like many nations around the world, the United States has a freedom of information law.¹⁰ The U.S. Freedom of Information Act (FOIA) is a law that allows the public to request access to previously unreleased executive branch government information.¹¹ FOIA was signed into law in July 1967. Executive branch departments must comply with requests for information. However, there are nine exemptions to FOIA.¹² The FOIA statute expressly provides that agencies may only recover the direct costs of search, duplication, or review, and various fee structures apply depending on the status of the requester and/or purpose of request.¹³ FOIA is not meant to be a cost recovery vehicle for the U.S. government in making available PSI. Gellman notes, “there is no doubt that the FOIA recovers only a fraction of its costs, and this is clearly the intent of the law. Recoverable FOIA costs bear no relationship to the commercial value of the information.”¹⁴ Even though FOIA operates at a financial loss, it is generally understood that the public benefit in being able to access information outweighs the cost associated with making the information available.

reasons for denying such protection to United States Government works in foreign countries, or for precluding the Government from making licenses for the use of its works abroad.” H.R. Rep. No. 94-1476 (1976), at 59. See [http://en.wikisource.org/wiki/Page:H.R. Rep. No. 94-1476 \(1976\) Page 059.djvu](http://en.wikisource.org/wiki/Page:H.R. Rep. No. 94-1476 (1976) Page 059.djvu)

¹⁰ For a listing of other jurisdictions with Freedom of Information legislation, see http://en.wikipedia.org/wiki/Freedom_of_information_legislation

¹¹ See http://www.justice.gov/oip/foia_updates/Vol_XVII_4/page2.htm

¹² These exemptions include 1) national security, 2) internal personnel rules and practices of an agency, 3) information exempted by other statute, 4) trade secrets and commercial or financial information, 5) inter-agency or intra-agency memoranda not available outside litigation, 6) personnel and medical files related to personal privacy, 7) records or information compiled for law enforcement purposes, 8) information involved with the regulation or supervision of financial institutions, and 9) geological and geophysical information and data, including maps, concerning wells. See <http://www.justice.gov/oip/foi-act.htm>

¹³ Gellman notes three fee categories: 1) Commercial users may be charged for the cost of search, duplication, and review; 2) educational or noncommercial scientific users or members of the media may only be charged for duplication costs; 3) all other users may be charged for search and duplication. There are other complex metrics for whether fees will be charged to specific users. For example, no fees may be charged for any noncommercial request for the first two hours of search time or for the first 100 pages of duplication. *Twin Evils: Government Copyright and Copyright-Like Controls Over Government Information*, 45 Syracuse Law Review 999 (1995), at 12-13. Available at <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA394923>

¹⁴ Gellman, at 13.

State of Play: Public Sector Information in the United States

Office of Management and Budget Circular A-130

The Office of Management and Budget Circular A-130 is a crucial piece of U.S. federal regulatory policy with respect to the dissemination and re-use of PSI. Circular A-130 was adopted in 1985 and last updated in 2000. It establishes policies for the management of federal information resources. Circular A-130 observes that the federal government is the largest single producer, collector, consumer, and disseminator of information in the United States. In addition to communicating basic assumptions with regard to the handling of U.S. PSI, Circular A-130 discusses specifically how government agencies can “avoid establishing or permitting others to have arrangements that are exclusive, restricted, or otherwise interfere with making information available on a timely and equitable basis.”¹⁵

Nancy Weiss, General Counsel of the U.S. Institute of Museum and Library Services, describes the main tenets of Circular A-130 as follows: “Agencies will: 1) Disseminate information in a manner that achieves the best balance between the goals of maximizing the usefulness of the information and minimizing the cost to the government and the public; 2) Disseminate information dissemination products on equitable and timely terms; 3) Take advantage of all dissemination channels, federal and nonfederal, including State and local governments, libraries and private sector entities, in discharging agency information dissemination responsibilities; and 4) Help the public locate government information maintained by or for the agency.”¹⁶

While Section 105, FOIA, and Circular A-130 provide the backbone of U.S. federal PSI framework, sometimes those policies are implemented or interpreted in ways that limit the dissemination of that information. Gellman argues that some federal agencies enact “copyright-like” controls on government information that limits the utility of innovative protocols like Section 105 and FOIA.¹⁷ These efforts, coupled with the increasingly strong push against

¹⁵ For instance, agencies are directed to 1) Avoid establishing, or permitting others to establish on their behalf, exclusive, restricted, or other distribution arrangements that interfere with the availability of information dissemination products on a timely and equitable basis; 2) Avoid establishing restrictions or regulations, including the charging of fees or royalties, on the reuse, resale, or re-dissemination of Federal information dissemination products by the public; and, 3) Set user charges for information dissemination products at a level sufficient to recover the cost of dissemination but no higher. They must exclude from calculation of the charges costs associated with original collection and processing of the information. For more information, see Management of Federal Information Resources Notice (Transmittal Memorandum No. 4, 2000), OMB Circular No. A-130. Available at http://www.whitehouse.gov/omb/Circulars_a130_a130trans4/

¹⁶ Weiss, Nancy. *Overview of US Federal Government Information Policy*. US CODATA—OECD Workshop, January 2008. Slides available at <http://www.oecd.org/dataoecd/28/0/40047022.pdf>

¹⁷ Gellman suggests that the U.S. FOIA framework allows agencies to deny public access to information not covered under the statutory exemptions. These include “loopholes created by unfortunate or erroneous interpretations of the law, by lack of resources, or by poorly drafted legislation; the ease of exercising dominion over information in electronic formats; the absence of organized opposition to restrictive agency activities; the lack of effective

State of Play: Public Sector Information in the United States

openness coming from various sectors through national security restrictions, protections for proprietary information, and the commercialization and control of public information by the private sector, offer continuing challenges.

b. State and local frameworks

All 50 states, the District of Columbia, and some territories have some form of freedom of information legislation that governs PSI at the state and local level.¹⁸ But, state and local governments, for various reasons, do not approach the standardization in openness that has characteristic of federal information covered under the mechanisms described above. Unlike at the federal level, there are few laws forbidding state and local governments in attaching copyright for the information they produce and disseminate. Sometimes state and local governments apply additional licensing or other restrictions in order to recuperate costs. However, some states like Oregon have adopted policies that expressly remove copyright protection from PSI, like state statutes.¹⁹ The state of Virginia passed legislation that indicates a preference for state-funded materials to be released with a Creative Commons (or equivalent open content) license.²⁰ The New York State Senate releases text, pictures and graphics, third-party submissions, and Senate video under a Creative Commons license.²¹

Some local entities have been able to experiment with projects supporting the dissemination and re-use of local PSI, but these activities vary depending on municipality size, leadership, and oftentimes, interest from knowledgeable volunteers and other community members. Cities such as Washington, D.C. and San Francisco have demonstrated a commitment to disseminating PSI.²² In November 2010, San Francisco passed an Open Data Policy, which requires that city departments and agencies “shall make reasonable efforts” to publish data under their control.^{23 24}

oversight and enforcement by the Congress and the executive branch; and misplaced agency zeal, entrepreneurial or otherwise.” Gellman, at 30.

¹⁸ See http://en.wikipedia.org/wiki/Freedom_of_information_in_the_United_States

¹⁹ See <http://resource.org/oregon.gov/>

²⁰ See <http://leg1.state.va.us/cgi-bin/legp504.exe?091+ful+CHAP0791>

²¹ See <http://www.nysenate.gov/copyright-policy>, which recommends the Creative Commons Attribution-NonCommercial-NoDerivatives U.S. license (<http://creativecommons.org/licenses/by-nc-nd/3.0/us/>).

²² For example, see <http://data.octo.dc.gov/> and <http://datasf.org/>.

²³ See http://www.sfbos.org/ftp/uploadedfiles/bdsupvrs/bosagendas/materials/bag110910_101155.pdf

4. Economic impacts of open access under U.S. federal framework

The EU Directive on the re-use of public sector information prohibits exclusive arrangements, encourages transparency, and says “[w]here charges are made, the total income from supplying and allowing re-use of documents shall not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment.”²⁵ When it comes to charging and usage policies, countries implement a variety of different processes. Just as with some U.S. state and local governments, some jurisdictions treat PSI as controlled and controllable information, with access based upon the ability for the entity to recover the costs associated with making the information available.²⁶ However, some countries have adopted open licensing frameworks in order to provide access to their PSI.²⁷

Pollock describes three charging policies for the recuperation of costs associated to making available PSI: profit maximizing, average-cost (cost recovery), and marginal cost.²⁸ He points out that profit-maximizing and average-cost pricing are associated with retaining strong control over re-use and redistribution through the use of copyright or cost-recovery agreements, while a marginal cost pricing structure would be more associated with a government entity releasing the

²⁴ For more examples of U.S. city-level open government and open data initiatives see <http://wiki.civiccommons.org/>

²⁵ “Where charges are made, the total income from supplying and allowing re-use of documents shall not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment.” See Article 6, *Directive 2003/98/EC of the European Parliament and of the Council*. Available at http://ec.europa.eu/information_society/policy/psi/docs/pdfs/directive/psi_directive_en.pdf

²⁶ Peter Weiss notes, “In other countries, particularly in Europe, publicly funded government agencies treat their information holdings as a commodity used to generate short-term revenue. They assert monopoly control on certain categories of information to recover the costs of its collection or creation. Such arrangements tend to preclude other entities from developing markets for the information or otherwise disseminating the information in the public interest.” Weiss, Peter, *Borders in Cyberspace: Conflicting Public Sector Information Policies and their Economic Impacts*. February 2002, at 2. Available at http://www.weather.gov/sp/Borders_report.pdf

²⁷ Some jurisdictions with open licensing frameworks include the United Kingdom Government Licensing Framework, New Zealand Government Open Access and Licensing Framework, Spain’s Aporta Project, and the Australian Government Information Licensing Framework.

²⁸ “Given the demand faced by the government entity; average-cost or cost-recovery describes setting prices equal to average long-run costs; marginal-cost describes setting prices equal to the short-run marginal cost, that is the cost of supplying data to an extra user.” Pollock, Rufus, *The Economics of Public Sector Information*. University of Cambridge, November 2008, at 8-9. Available at http://www.rufuspollock.org/economics/papers/economics_of_psi.pdf

State of Play: Public Sector Information in the United States

PSI in such a way so that it would be free to re-use or redistribute.²⁹ From the perspective of the U.S. federal government, the OMB A-130 Circular describes an optimal situation: “In order to minimize the cost and maximize the usefulness of government information, the expected public and private benefits derived from government information should exceed the public and private costs of the information...”³⁰ Nancy Weiss explains that A-130 “encourages agencies to set user fees for government information products at the marginal cost of dissemination...As a result, agencies often post information on the Internet, where the marginal cost of dissemination is zero.”³¹ Peter Weiss concludes, “Charging marginal cost of dissemination for public sector information will lead to optimal economic growth in society and will far outweigh the immediate perceived benefits of aggressive cost recovery. Open government information policies foster significant, but not easily quantifiable, economic benefits to society.”³²

National Oceanic and Atmospheric Administration

Peter Weiss notes that in the U.S., open and unrestricted access to public sector information has resulted in the rapid growth of information-intensive industries.³³ Uhlir and others argue that the U.S. federal government should treat public information as a major social and economic infrastructural investment. He notes several advantages to unrestricted re-use of publicly funded scientific data, of which the benefits can be extrapolated to PSI generally: it promotes new types of research and avoids duplication of research, enables the development of tools that can aid in search and discovery of information, promotes transparency and validation of government funded information, maximizes the return on investment for government funded PSI, promotes

²⁹ Pollock, at 9.

³⁰ This section also notes, “the benefits to be derived from government information may not always be quantifiable.” Management of Federal Information Resources Notice (Transmittal Memorandum No. 4, 2000), OMB Circular No. A-130. Available at http://www.whitehouse.gov/omb/Circulars_a130_a130trans4/

³¹ Weiss, Nancy. “Overview of U.S. Federal Information Policy,” *The Socioeconomic Effects of Public Sector Information on Digital Networks*. The National Academies Press, 2009, at 4. Available at http://www.nap.edu/catalog.php?record_id=12687

³² Weiss, Peter. *Borders in Cyberspace: Conflicting Public Sector Information Policies and their Economic Impacts*. February 2002, at 17. Available at http://www.weather.gov/sp/Borders_report.pdf

³³ Peter Weiss, at 2.

State of Play: Public Sector Information in the United States

interoperability between different sets of government information, and supports socioeconomic and good governance.³⁴

Uhlir acknowledges the inherent difficulty in measuring the value of public sector information, but notes, “information with the lowest barriers to access and use will potentially have the widest audience, and the positive effects of public domain information can be increased by enormous proportions when such information is placed on global digital networks...”³⁵ One prime example that demonstrates how U.S. federal information is promoting economic activity is information and data made available by the National Oceanic and Atmospheric Administration.³⁶ The broad availability of data disseminated by NOAA, particularly weather information, stimulates economic activity and leads to the creation of value-added industries. Rodney Weiher, former Chief Economist at NOAA, noted that the agency adheres to the Circular A-130 guidelines, “set[ting] user fees at a level sufficient to recover the cost of dissemination but no higher, and, in particular, it does not charge prices to recover the capital costs.”³⁷

In 2008, Weiher wrote that NOAA real-time weather data supplied the private weather service industry with sales of over \$700 million annually.³⁸ There are several related economic benefits that arise from the use of NOAA data, as outlined below:

- Benefits of real-time oceanographic data (PORTS) in the Houston/Galveston Bay port is \$15 million annually
- Benefits to world ship routing from NOAA polar satellite data is estimated at \$95 million annually

³⁴ Uhlir, Paul. *Policy for Publicly Funded Scientific Data in the U.S.* Symposium on the Value of Shared Access and Reuse of Publicly Funded Scientific Data, December 2010, at 14. Available at http://sites.nationalacademies.org/PGA/brdi/PGA_059258

³⁵ Uhlir, Paul F. *Policy Guidelines for the Development and Promotion of Governmental Public Domain Information*, UNESCO, 2004, at 8. Available at <http://unesdoc.unesco.org/images/0013/001373/137363eo.pdf>

³⁶ See <http://www.noaa.gov/>

³⁷ The decision to not attempt to recover costs to cover capital expenses is explained by Weiher: “First, it is difficult to exclude users, which makes it difficult to charge for the data in order to recoup the cost of the capital. Also, the marginal cost of producing additional information is essentially zero, so to charge for it would be non-optimal because it would exclude many users who value the data.” (48) Weiher, Rodney. “Assessing the Economic and Social Benefits of NOAA Data Online,” *The Socioeconomic Effects of Public Sector Information on Digital Networks*. The National Academies Press, 2009, at 48. Available at http://www.nap.edu/catalog.php?record_id=12687

³⁸ Weiher, at 50.

State of Play: Public Sector Information in the United States

- Installation of National Weather Service Doppler radars reduced tornado fatalities and injuries by 40% from the levels in late 80's and early 90's
- Average value of all US daily weather forecast information is around \$109 per household
- U.S.'s \$8-10 billion and growing annual Weather Derivatives financial industry relies on NOAA's seasonal weather data and records
- NOAA's forecasts and warnings and associated emergency responses result in \$3 billion [savings] in a typical hurricane season
- U.S. electricity generators save \$166 million annually using 24-hr temperature forecasts³⁹

NOAA has been held up as a success story with regard to innovation and a demonstration that economic activity can be generated in the absence of copyright and other restrictions. NOAA maintains a website dedicated to showcasing the economic and social benefits of their PSI and products.⁴⁰

5. Recent PSI State of Play in U.S.

Three relatively recent activities demonstrate the increasing U.S. interest in the dissemination and re-use of public sector information: Data.gov, the National Institutes of Health Public Access Policy, and the Trade Adjustment Assistance Community College and Career Training Grant Program.

a. Data.gov

Data.gov is a website maintained by the U.S. federal government whose purpose is “to increase public access to high value, machine-readable datasets generated by the Executive Branch of the Federal Government.”⁴¹ In 2009 President Obama released a Memorandum on Transparency and

³⁹ Weiher, Rodney. *Assessing the Economic & Social Benefits of NOAA Data*, NAS/OECD Conference, Paris. February 2008, at 17-18. Available at <http://www.oecd.org/dataoecd/12/31/40066192.pdf>

⁴⁰ See <http://www.economics.noaa.gov/>

⁴¹ See <http://www.data.gov/about>. In addition, Lakhani et al. note that one of the priorities in developing Data.gov was to push government information through a single web portal. They write that at the beginning of 2009, the U.S. government held 24,000 .gov domains containing 30 million web pages. Karim R, Lakhani, Robert D. Austin, and

State of Play: Public Sector Information in the United States

Open Government, which described the overarching principles for government operation: transparency, participation, and collaboration.⁴² Obama's memo led to the development of the Open Government Directive, requiring government agencies to 1) publish government information online; 2) improve the quality of government information; 3) create and institutionalize a culture of open government; and 4) create an enabling policy framework for open government.⁴³ The United States government efforts are collectively called the Open Government Initiative.⁴⁴ Data.gov was launched on May 21, 2009, on the 120-day anniversary of the President's Memo on Transparency and Open Government. Most of the PSI made available through Data.gov is "wholesale" (raw data) instead of "retail" (displayed and interpreted).⁴⁵ Data.gov contains 3,191 records in its raw data catalog, 789 records in its tool catalog, and 305,055 records in its geodata catalog.⁴⁶ Crucial leadership within the Obama administration has provided a driving force to promoting efforts to make available PSI through Data.gov. These include Chief Information Officer Vivek Kundra, Chief Technology Officer Aneesh Chopra, and

Yumi Yi. *Data.gov Case Study*. Harvard Business School, May 2010, at 7. Available at http://www.data.gov/documents/hbs_datagov_case_study.pdf

⁴² See http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment/. See also Orszag, Peter R. "Memorandum for the Heads of Executive Departments and Agencies Re: Open Government Directive." December 8, 2009 M-10-06. Available at http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf

⁴³ See <http://www.whitehouse.gov/open/documents/open-government-directive>

⁴⁴ See <http://www.whitehouse.gov/open>

⁴⁵ Some argue that Data.gov should allow the typical citizen to visit and interact directly with government data. Lakhani et al note that within the U.S. government, there was some debate about the best way to present the data, but eventually settled on the "raw data" approach, even though there was concern that the raw data might be misinterpreted or misconstrued by downstream users. Lakhani, Karim R, Robert D. Austin, and Yumi Yi. *Data.gov Case Study*. Harvard Business School, May 2010, at 7. Available at http://www.data.gov/documents/hbs_datagov_case_study.pdf. This is aligned with the position recommended by Robinson et al: "Rather than struggling, as [the government] currently does, to design sites that meet each end-user need, it should focus on creating a simple, reliable and publicly accessible infrastructure that "exposes" the underlying data." Robinson, David, Harlan Yu, William P. Zeller, and Edward W. Felten, *Government Data and the Invisible Hand*, 11 Yale Journal of Law & Technology, 2009, at 161. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1138083

⁴⁶ See <http://www.data.gov/catalog/>. At the Gov2.0 Summit in September 2010 in Washington, D.C., Sunlight Foundation Director Ellen Miller said, "the drive for transparency has stalled," criticizing the quantity and quality of datasets—at least those relating to government transparency and accountability—posted to Data.gov. She also underscored the need for robust and accessible underlying data, instead of an emphasis on web presentation and visualization. See <http://sunlightfoundation.com/blog/2010/09/07/gov2-0-presentation-an-open-government-scorecard/>

State of Play: Public Sector Information in the United States

former Deputy Chief Technology Officer and Director of the Open Government Initiative Beth Noveck.

b. National Institutes of Health Public Access Policy

PSI activity in the U.S. extends beyond simply access to data. The National Institutes of Health (NIH) spends \$28 billion on research each year, resulting in about 65,000 peer-reviewed articles. The NIH Public Access Policy “requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central upon acceptance for publication.”⁴⁷ Through the PubMed repository, these manuscripts and materials are made available to the public. The original policy was voluntary and went into effect in May 2005. Later, the policy became mandatory, and was made permanent in March 2009. The Obama administration, particularly the Office of Science and Technology Policy, supports expanding public access policies to the outputs of other federally funded programs. The Federal Research Public Access Act (FRPAA) would require that eleven U.S. government agencies with annual extramural research expenditures over \$100 million make manuscripts of journal articles stemming from research funded by that agency publicly available via the Internet.⁴⁸

c. Trade Adjustment Assistance Community College and Career Training Grant Program

There has been recent support in the U.S. in promoting public access to publicly funded educational materials. The Trade Adjustment Assistance Community College and Career Training Grant Program (TAACCCT) is a grant fund cooperatively administered through the U.S. Department of Labor and U.S. Department of Education.⁴⁹ The program will make \$2 billion available over the next four years for grants that will provide community colleges and other eligible institutions of higher education with funds to expand and improve their ability to deliver education and career training programs. The solicitation for grant applications requires that all resources created using these funds must be released under the Creative Commons Attribution 3.0 license.^{50 51} Unlike the NIH Public Access Policy, which only requires that

⁴⁷ See <http://publicaccess.nih.gov/>

⁴⁸ See <http://www.taxpayeraccess.org/issues/frpaa/index.shtml>

⁴⁹ See <http://www.federalgrantswire.com/trade-adjustment-assistance-community-college-and-career-training-taacctt-grants.html>

⁵⁰ The grant solicitation states: “In order to further the goal of career training and education and encourage innovation in the development of new learning materials, as a condition of the receipt of a Trade Adjustment Assistance Community College and Career Training Grant (“Grant”), the Grantee will be required to license to the

State of Play: Public Sector Information in the United States

manuscripts be made available online, the TAACCCT program is new in its requirement of open licensing.

d. Other projects

There are dozens of interesting new projects related to the dissemination and re-use of public sector information in the U.S., with involvement from interested citizens, community organizers, developers, non-profit organizations, businesses, city governments, schools, universities, and more. Some of these web sites include Open311,⁵² RECAP,⁵³ Civic Commons,⁵⁴ OpenCongress,⁵⁵ Smithsonian Commons,⁵⁶ Sunlight Labs,⁵⁷ Apps for Democracy,⁵⁸ GovLoop,⁵⁹ Public.Resource.Org,⁶⁰ SeeClickFix,⁶¹ Flickr Commons,⁶² FlyOnTime,⁶³ and countless others.

public (not including the Federal Government) all work created with the support of the grant (“Work”) under a Creative Commons Attribution 3.0 License (“License”). This License allows subsequent users to copy, distribute, transmit and adapt the copyrighted work and requires such users to attribute the work in the manner specified by the Grantee. Notice of the License shall be affixed to the Work.” See <http://www.doleta.gov/grants/pdf/SGA-DFA-PY-10-03.pdf>

⁵¹ The Washington State Board for Community and Technical Colleges (SBCTC) has adopted a similar open licensing policy such that “all digital software, educational resources and knowledge produced through competitive grants, offered through and/or managed by the SBCTC, will carry a Creative Commons attribution license; and the open licensing policy applies to all funding sources (state, federal, foundation and/or other fund sources) that flow through SBCTC as a competitive grant to any party, is not retroactive to past grants, and does not apply to state or federal formula driven funding.” See http://www.sbctc.edu/general/admin/Tab_9_Open_Licensing_Policy.pdf

⁵² See <http://open311.org/>

⁵³ See <https://www.recapthelaw.org/>

⁵⁴ See <http://civiccommons.org/>

⁵⁵ See <http://www.opencongress.org/>

⁵⁶ See <http://www.si.edu/commons/prototype/>

⁵⁷ See <http://sunlightlabs.com/>

⁵⁸ See <http://www.appsfordemocracy.org/>

⁵⁹ See <http://www.govloop.com/>

⁶⁰ See <http://public.resource.org/>

⁶¹ See <http://www.seeclickfix.com/citizens>

⁶² See <http://www.flickr.com/commons>

State of Play: Public Sector Information in the United States

Events, such as the Gov2.0 Summits⁶⁴ and various bar camps⁶⁵ help to professionalize and foster leadership in this field.

6. Conclusion

Carl Malamud, a longtime advocate for open access to PSI in the United States, says that government information is “the raw material of innovation, creating a wealth of business opportunities that drive our economy forward. Government information is a form of infrastructure no less important to our modern life than our roads, electrical grid, or water systems.”⁶⁶

Interest in and dissemination of PSI in the U.S. is increasing, and many innovative economic activities may be due in part to the U.S.’s unique handling of information created by the federal government. At the same time, the U.S., like governments around the world, continue to struggle with balancing access to information with other priorities, including national security, privacy, and financial sustainability. However, the U.S. has demonstrated that in at least some situations, unrestricted access can promote other desirable effects, like the creation of value-add businesses that pop up around the free flow of information such as weather data.

While many U.S. state and local information management policies have not yet caught up with the more liberal federal PSI framework, there are numerous experiments being led not only by forward-looking government entities, but also cultural heritage institutions, interested individuals, nonprofit organizations, and innovative startups.

⁶³ See <http://flyontime.us/>

⁶⁴ See <http://www.gov2summit.com/gov2010>

⁶⁵ Such as <http://transparencycamp.org/>

⁶⁶ Malamud, Carl. “*By the People*” speech given at Gov2.0 Summit, Washington, D.C., September 2009, at paragraph 27. Available at http://public.resource.org/people/3waves_pamphlet.pdf