State Open Data Policies and Portals

By Laura Drees & Daniel Castro | August 18, 2014

This report provides a snapshot of states' efforts to create open data policies and portals and ranks states on their progress. States creating new open data policies or portals, or refreshing old ones, have many opportunities to learn from the experiences of early adopters in order to fully realize the benefits of data-driven innovation.

INTRODUCTION

Given the significant opportunities that open data presents, all states should be developing open data policies and portals.

Many state governments in the United States have begun to embrace open data as a way to encourage transparency and accountability, increase public participation, and promote economic growth. "Open data" refers to data that is made freely available without restrictions.[1] States publish data sets on a wide variety of topics, such as education, health care, and public safety. By releasing open data, government agencies can foster data-driven innovation not only within government, but also among private-sector organizations and individuals who can make use of these data sets. The impact of releasing open data can be substantial. A 2013 McKinsey Global Institute report estimated that open data could add over $3 trillion annually in total value to the global economy.[2] Yet for all of the evidence of economic and societal benefits from open data, states vary widely in the degree to which they have embraced the idea. This report provides a snapshot of state efforts to create open data policies and portals and ranks states on their progress.

STATE OPEN DATA EFFORTS

States show their commitment to open data in two principal ways: by establishing open data policies and by creating open data portals. Open data policies specify what data the government will publish and how it will do so. Open data portals bring data from multiple government agencies onto a single website.

Open data policies differ by state, but most have some common elements. Key among these are requirements that data be open by default–that is, that state governments and agencies publish all information, such as public records, expenditure information, and legislative records, as a matter of course, unless there is an overruling justification against it, such as confidentiality and privacy reasons. Many open data policies also require that data be released in a non-proprietary, machine-readable format. Machine readability is crucial for ensuring that businesses, non-profits, and others can easily process and repurpose public data sets. Typically, policies also specify that data be made available to the public for any purpose, and often at no cost.

Open data portals are more consistent across states and generally have the following features: options to export data in machine-readable formats; application programming interfaces (APIs) for downloading particular parts of datasets; data from multiple sources within state agencies, including financial, expenditure, employment, population, mapping, education, and healthcare information; and options to sort and filter data according to different categories, such as type, source, and common elements.

To date, ten states have open data policies, established either via an executive order or by legislation. Many of these policies are relatively new. Within the past 2 years, five states have established open data policies, and four more have amended existing policies. In addition, twenty-four states offer open data portals, whether or not they have also established open data policies.

State	Open Data Policy	Year
Connecticut	Executive Order: Conn. Exec. Order No. 2014-39 (February 20, 2014)	2014
Hawaii	Legislation: HI Rev. Stat. §§ 27-41 - 45	2013
Illinois	Executive Order: Ill. Exec. Order No. 3 (Sept. 18, 2012); Legislation: §§ 20 ILCS 45/1 – 45/99	2012; amended 2014
Maryland	Executive Order: Md. Exec. Order No. 2012-18 (Jan 1., 2012); Legislation: Md. State Government Code Ann. §§ 10-1401 – 1404	2012; amended 2014
New Hampshire	Legislation: R.S.A. Tit. I, Ch. 21-R	2012
New York	Executive Order: NY Exec. Order No. 2013-95 (March 11, 2013)	2013

Oklahoma	Legislation: 62 Okl. St. §§ 34.11 – 34.11.5	2011; amended 2013
Rhode Island	Executive Order: R.I. Exec. Order No. 2013-01 (Jan. 10, 2013)	2013
Texas	Legislation: Tex. Gov't Code § 2054.1265	2011; amended 2013

Table 1: States with Open Data Policies

Of the ten states with open data policies, nine offer open data portals. New Hampshire has an open data page, but it is not yet complete and offers no datasets. Fourteen states, while not offering open data policies themselves, have individual cities with open data policies and portals. Cities such as Burlington, Vermont; Philadelphia, Pennsylvania; and South Bend, Indiana could serve as models and catalysts for statewide open data policies within their respective states. States vary as to whether their open data policies were established by their governors or by their state legislatures. Of the ten states with open data policies, three—Connecticut, New York, and Rhode Island—established them via executive order. Executive orders, being unilaterally issued by state governors, provide a faster means of making policy than legislation, but they are limited by the existing constitutional powers of the governor and may be rescinded by another executive order or by legislation. Executive orders often require no legislative review—less than one quarter of states explicitly require review, and even in these states it is often limited.[3] Although it is sometimes difficult and time-consuming to pass legislation initially, this approach may help establish a long-term state policy since it is equally difficult and time-consuming to later overturn. Legislation also may allow for more public participation since it involves an extensive approval process, including public hearings and other opportunities for constituents to influence their representatives.

Although both executive orders and legislation can be effective means of setting policy, one compromise between the speed of an executive order and the permanence and detail of legislation is for the governor to issue an executive order establishing a policy and the legislature to pass related legislation afterward. This has occurred in a few states. For example, both Maryland and Illinois established open data policies via executive order in 2012 and amended those policies via legislation in 2014. The Illinois legislation specifically provided that it superseded and repealed any contradictory or inconsistent executive orders. Similarly, in 2013, the New York state legislature introduced an open data bill seven months after the

governor issued an executive order establishing an open data policy, although that bill stalled in committee.[4]

STATE TRANSPARENCY EFFORTS

Some states have policies or portals narrowly focused on government transparency, which means that they publish information on only a few topics, such as government expenditures, contracts, or legislative records. Transparency policies and portals can be useful for opening up some types of government data, and they can also be precursors to more comprehensive open data efforts.

The following table presents trends in open data and transparency policies. Policies differ significantly from state to state, but in general, transparency policies tend to differ from open data policies in similar ways across states. As a rule, open data policies tend to pick up where earlier transparency policies have left off. Some states even seem to have treated transparency policies as foundations on which to build their open data policies, which can be effective because an efficacious open data policy can accomplish both transparency and other open data goals.

	Transparency Policy	Open Data Policy
Title of legislation or executive order	Often contains some combination of the words "expenditure," "transparency," and "finances."	Often contains the phrase "open data," "open government," or "electronic data." Does not specify financial data in title.
Purpose of policy	To increase government transparency and accountability and public access to government data, especially revenue and expenditures.	To increase government transparency and accountability, encourage other beneficial public and private data use and innovation.
Types of data required to be released under policy	Requires government revenue, expenditure, employment, and contract data; sometimes links to other state agency portals, such as health and education departments.	May require government revenue, expenditure, employment, and contract data; other general agency data; mapping and population data collected and compiled by state government; or simply contain "open by default" language.
Format of data on portal	Often requires that data be searchable and presented in a way that readers can understand.	Often requires that data be searchable and machine readable.

Method of making information available	May provide for development of a portal.	May provide for development of a portal. May establish an advisory board to facilitate data access.

Table 2: Differences Between Transparency Policies and Open Data Policies

While open data portals offer wide ranges of expenditure and general state and population data, as well as information from other state agencies and sources on a single website, transparency portals focus almost exclusively on financial and expenditure data, sometimes including public records and other information through links to other individual state agencies' sites. In addition, unlike most open data portals, transparency portals do not always provide access to data in machine-readable formats, making the data much more difficult to extract for additional uses.

METHODOLOGY

In this report, states are evaluated based on the contents of their open data policies and open data portals and awarded points according to the structure of the portal, source of the policy, availability of financial and government data, and machine readability. The states are ranked from most to fewest points.

DATA SOURCES

Multiple sources were used to compile the list of open data policies and open data portals. The initial list of open data policies was compiled from the Sunlight Foundation.[5] In addition, basic web searches and LexisNexis research were conducted to locate open data policies that were not listed or not current on the Sunlight Foundation website. The Data.gov website was used to compile the initial list of open data portals. [6] The open data sites listed on Data.gov were evaluated to confirm that they were current and actually open data websites and not expired pages, general state government websites, or transparency sites. In addition, basic web searches and LexisNexis database research were conducted to confirm that the open data sites listed were accurate and to locate open data sites that were not listed by either source.

SCORES

States were awarded scores based on their open data policies and portals. As described below, states could earn a maximum of eight points based on the presence and quality of their policies and portals.

PRESENCE OF AN OPEN DATA POLICY (2 POINTS MAX)

A state was awarded two points if it had a statewide open data policy or a single point if it had at least one city with an open data policy. No extra points were awarded for the existence of city policies within states that had state open data policies. Furthermore, no additional points were awarded for general transparency policies, public records acts, or open data policies where the law provided for a group that was charged with updating an open data portal but did not require open data more generally.

QUALITY OF OPEN DATA POLICY (3 POINTS MAX)

Each state was awarded two points if its open data policy applied broadly to all government data and one point if its policy specified that only certain types of government data must be provided, e.g., only spending information. An additional point was awarded if a state's open data policy specified machine-readability. No points were awarded for the quality of city open data policies in states without state policies.

PRESENCE OF AN OPEN DATA PORTAL (2 POINTS MAX)

Each state was awarded two points for the existence of an open data portal and one point for the existence of a transparency portal if there was no open data portal. Only official government-run open data portals were considered. Open data portals are relatively easy to distinguish from transparency websites because transparency sites tend to offer only financial and expenditure information, provide fewer if any machine readability options, and offer other state agency information only through links if at all. No points were awarded for open records access sites or portals that focus exclusively on geospatial or mapping information.

QUALITY OF OPEN DATA PORTAL (1 POINT MAX)

States were awarded one point if the data sets on their open data or transparency portals were machine readable. Machine readability was assessed by identifying whether more than 50% of the data files on a portal could be downloaded in CSV, JSON, KML, or other such file formats. Many states' open data portals also provided Application Programming Interfaces (APIs) for effective data downloading and use. APIs allow users to extract subsets of full datasets, which can be convenient for users who do not require every data point in what would otherwise be a particularly large bulk download. Files provided in the PDF or DOC format and files that were provided without download links (e.g., charts or lists embedded into a web browser) were not considered to be machine readable.

FINDINGS

The six top-scoring states are Hawaii, Illinois, Maryland, New York, Oklahoma, and Utah. Each of these states has established an open data policy; only New York did so exclusively by means of executive order; the rest either passed legislation or issued both an executive order and legislation. These states' policies require their open data portals to publish basic government data, such as expenditure information, as well as other agency data. They also require that the data be published in a machine-readable format. The corresponding portals contain extensive catalogs of open data, are relatively simple to navigate, and provide data in machine-readable formats. The portals also provide links to APIs to download particular data and have other information designed specifically for developers looking to build applications using the data. The next highest-ranked states, Connecticut and Texas, offer similarly serviceable, machine-readable open data portals that provide wide varieties of information, but Connecticut's policy does not require machine readability, and Texas's does not require data beyond expenditures. Of the next ranked two states, Rhode Island's policy requires neither machine readability nor government data beyond expenditures; New Hampshire's policy requires machine readability and many types of data, but its open data portal is not yet fully functional.

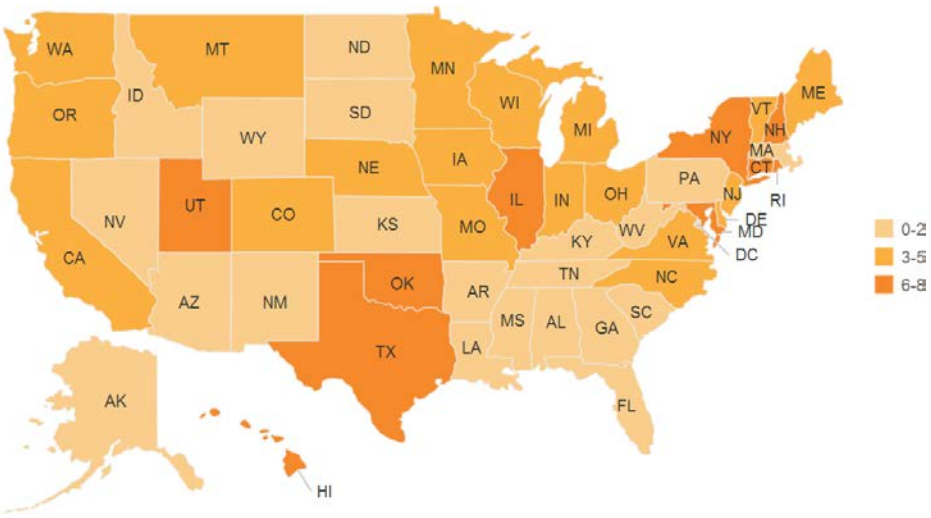

Figure 1: Map of State Open Data Scores

State	Policy	Policy Quality	Portal	Portal Quality	Total
Alabama	0	0	1	0	1
Alaska	0	0	1	0	1
Arizona	0	0	1	1	2
Arkansas	0	0	1	1	2
California	1	0	2	1	4
Colorado	0	0	2	1	3
Connecticut	2	2	2	1	7
Delaware	0	0	2	1	3
Florida	0	0	1	1	2
Georgia	0	0	1	1	2
Hawaii	2	3	2	1	8
Idaho	0	0	1	1	2
Illinois	2	3	2	1	8
Indiana	1	0	1	1	3
Iowa	0	0	2	1	3
Kansas	0	0	1	0	1
Kentucky	1	0	1	0	2
Louisiana	0	0	1	0	1
Maine	0	0	2	1	3
Maryland	2	3	2	1	8
Massachusetts	1	0	0	0	1
Michigan	1	0	2	1	4
Minnesota	0	0	2	1	3
Mississippi	0	0	1	1	2
Missouri	1	0	2	1	4

State	Policy	Policy Quality	Portal	Portal Quality	Total
Montana	0	0	2	1	3
Nebraska	0	0	2	1	3
Nevada	0	0	1	0	1
New Hampshire	2	3	1	0	6
New Jersey	1	0	2	1	4
New Mexico	0	0	1	1	2
New York	2	3	2	1	8
North Carolina	1	0	1	1	3
North Dakota	0	0	1	1	2
Ohio	1	0	1	1	3
Oklahoma	2	3	2	1	8
Oregon	1	0	2	1	4
Pennsylvania	1	0	1	0	2
Rhode Island	2	1	2	1	6
South Carolina	0	0	1	1	2
South Dakota	0	0	1	0	1
Tennessee	1	0	1	0	2
Texas	2	2	2	1	7
Utah	2	3	2	1	8
Vermont	1	0	2	1	4
Virginia	0	0	2	1	3
Washington	0	0	2	1	3
West Virginia	0	0	1	1	2
Wisconsin	1	0	1	1	3
Wyoming	0	0	1	0	1

Many states have not yet upgraded their transparency portals to open data portals. While a general transparency portal is a good start, open data portals can help increase transparency and accountability by opening up all government data, not just certain types of records.[7] In addition, government transparency websites often link to other state agency sites that may or may not provide additional data, and each agency website may be organized differently. Open data portals organize data sets and make them searchable, thereby saving both public and private users a great deal of time and resources when looking for data. Most open data portals are hosted at a URL such as "data.statename.gov," "statename.gov/data," or "open.statename.gov." Some transparency websites have more unorthodox web addresses. For example, Florida's transparency portal, www.floridahasarighttoknow.com, is sponsored by the Florida governor and is accessed via a link from the official state portal.

A frequently updated transparency portal that offers at least some machine-readable data may be more useful than an open data portal that is rarely updated or that contains very few datasets. For example, Massachusetts established a seemingly well-organized open data catalog in 2009 that offered links and access to many datasets, some machine readable; however, the state website has since been reorganized and much of its Open Data Initiative wiki has not been updated since 2012, so the data are less useful than they could be. In contrast, Arizona still offers only a transparency portal, but it is searchable, it is current, and the data is machine readable.

Some states established transparency policies in the 2000s or early 2010s. Examples are Alabama's 2009 executive order and the 2008 Kansas Taxpayer Transparency Act, policies that helped start the process of opening up government spending information.[8] However, while transparency portals that provide PDFs or even spreadsheets of financial data are a vast improvement over the process of ordering paper copies of records, they are not useful for obtaining raw data sets.

It is important for states to have both an open data policy that outlines where, how, what formats, and what types of data should be made available; and an open data portal that provides the data in a single accessible and machine-readable location. New Hampshire and Oregon are states that have developed only either a policy or a portal, not both: New Hampshire's open data policy is thorough and clearly states who should provide data and how, but its transparency portal only offers some data in machine-readable bulk downloads. By contrast, Oregon's open data portal is well organized, machine readable, and provides an API for exporting data; however, Oregon still does not have an open data policy. Most middle-ranked states, such as California, Michigan, and Missouri, similarly

offer open data portals but no clear open data policies. Open data policies provide opportunities to require machine readability, ensure accountability by requiring participation from all state agencies, keep data current by mandating regular updates of specific types of data, and ensure that unnecessary restrictions are not placed on public data. The mere existence of an open data portal does not necessarily guarantee continued effectiveness and contribution by various entities any more than the mere existence of an open data policy will immediately open up all government data. The purpose of open data portals is to provide government accountability and data that can be used for socially and economically beneficial purposes, and they are more likely to continue to be updated and maintained if they are backed up by state policies, just as policies are more likely to be effective if there is a place to publish the data they require.

Of the bottom-ranked twenty states, three have cities that have established their own open data policies and portals. These states, along with the five middle-ranked states that also have cities with open data policies and portals, are uniquely suited to developing the most effective policies and portals, because they can learn from their cities' open data efforts when they form statewide policies. For example, New York City is one of the most active cities in terms of beneficial use of open data, both public and private.[9] In 2013, two years after New York City passed its open data policy, New York Governor Cuomo signed an executive order ordering a statewide open data policy and portal.[10] Since then, New York State has sponsored its own data innovation challenges for effective use of statewide data.[11] With a population of just over 100,000, South Bend, Indiana is one of the smallest cities to offer an open data policy and portal, but its policy is particularly clear and organized, including both detailed definitions of what constitutes open data and requiring machine readability.[12] South Bend has also committed to various programs to encourage use of its open data, as when it partnered with Code for America to encourage local innovators to build city data-based applications.[13] Although Indiana still offers only a state transparency portal and a policy that requires disclosure of financial information, South Bend could set an excellent example for the rest of the state, just as New York City did for the state of New York.

RECOMMENDATIONS AND CONCLUSION

Many states are actively improving their open data efforts by passing legislation, issuing executive orders, and developing new open data portals. For example, the New Jersey legislature has already pre-filed an open data initiative for consideration in the 2014 legislative session, and several other states, such as Ohio, Alaska, and Minnesota, have considered open data policies in recent sessions. This issue is constantly

evolving even when state legislatures are not in session: both New Jersey and Montana launched open data portals in July 2014.

Given the significant opportunities that open data presents, all states should be developing open data policies and portals as an initial step towards greater use of open data. States creating new open data policies or portals, or refreshing old ones, have many opportunities to learn from the experiences of early adopters. One option for states planning to legislate in this area is to amend their existing transparency policies to form open data policies, thereby modernizing already effective statutes. For example, Utah's 2014 legislation established an open data policy and portal by amending the duties required of the Utah Transparency Advisory Board and Utah Public Finance Website.[14]

Policymakers should pay attention to the implementation of open data efforts to ensure they maximize their return on investment. Government data that is buried somewhere on a state website, but not readily accessible and understandable, adds little value. Data sets should be well-documented and easy to interpret so that others, including average citizens, journalists, and developers, can effectively make use of them.

There is still much work to be done, but all states should be working to ensure that they are making progress towards fully realizing the economic and social benefits of open data.

REFERENCES

1. Open Data Handbook, “What is Open Data?”, Open Knowledge Foundation, 2012, http://opendatahandbook.org/en/what-is-open-data/.

2. James Manyika et al, “Open Data: Unlocking Innovation and Performance with Liquid Information,” McKinsey Global Institute, 2013, http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information.

3. “Governors’ Powers and Authority,” National Governors Association, accessed August 4, 2014, http://www.nga.org/cms/home/management-resources/governors-powers-and-authority.html#executive.

4. “New York Assembly Bill 8197,” Open States, accessed July 31, 2014, http://openstates.org/ny/bills/2013-2014/A8197/.

5. “Open Data Policies At Work,” Sunlight Foundation, accessed July 28, 2014, http://sunlightfoundation.com/policy/opendatamap/.

6. “Open Data in the United States,” Data.gov, accessed July 20, 2014, https://www.data.gov/open-gov/.

7. Emily Shaw, “Why Open Data and Accountability Are Not the Same Thing,” (Sunlight Foundation, June 4, 2014), http://sunlightfoundation.com/blog/2014/06/04/why-open-data-and-accountability-are-not-the-same-thing/

8. Ala. Exec. Order No. 46, (2009), http://www.archives.alabama.gov/executiveorders/executiveorders/2009_46.pdf; K.S.A. §§ 74-72, 122 – 74-72, 126, http://www.kansas.gov/KanView/pdf/2008/KSA2008Supp74-72122thru72126.pdf.

9. “Featured Datasets for NYC BigApps 2014,” NYC OpenData, accessed July 31, 2014, https://nycopendata.socrata.com/.

10. NY Exec. Order No. 95, (2013), http://www.governor.ny.gov/executiveorder/95.

11. “NYS Health Innovation Challenge,” Health 2.0, accessed July 31, 2014, http://www.health2con.com/devchallenge/nys-health-innovation-challenge/.

12. “State and County QuickFacts: South Bend, Indiana,” U.S. Census Bureau, accessed August 8, 2014, http://quickfacts.census.gov/qfd/states/18/1871000.html; City of South Bend Exec. Order No. 2, (2013), http://www.ci.south-bend.in.us/sites/default/files/files/Mayor_2013-Open%20Portal%20Exec%20Order.pdf.

13. “Mayor’s Office Programs and Initiatives,” City of South Bend, Indiana, accessed July 31, 2014, http://www.ci.south-bend.in.us/government/content/programs-and-initiatives.

14. Utah Code Ann. §§ 63A-3-401 – 406, http://le.utah.gov/~2014/bills/sbillenr/sb0070.pdf.

ABOUT THE AUTHORS

Laura Drees is a policy intern at the Center for Data Innovation. She is entering her third year at the University of Colorado Law School, where she has volunteered with CU Restorative Justice, worked as a Research Assistant, and served as a Legislative Research Fellow for State Senator Aguilar. She graduated magna cum laude and Phi Beta Kappa from the University of Kansas with a B.A. in English and Psychology and a minor in German.

Daniel Castro is the director of the Center for Data Innovation where he leads the Center's research efforts. Mr. Castro is also a senior analyst at the Information Technology and Innovation Foundation. Previously, he worked as an IT analyst at the Government Accountability Office. He has a B.S. in Foreign Service from Georgetown University and an M.S. in Information Security Technology and Management from Carnegie Mellon University.

ABOUT THE CENTER FOR DATA INNOVATION

The Center for Data Innovation is the leading think tank studying the intersection of data, technology, and public policy. Based in Washington, DC, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as predictive analytics, open data, cloud computing, and the Internet of Things. The Center is a non-profit, non-partisan research institute proudly affiliated with the Information Technology and Innovation Foundation.

contact: info@datainnovation.org

datainnovation.org