

HOUSE BILL REPORT

HB 2202

As Reported by House Committee On:
Government Operations & Elections
Appropriations

Title: An act relating to the establishment of an open data policy to facilitate sharing and publication of government data.

Brief Description: Concerning the establishment of an open data policy to facilitate sharing and publication of government data.

Sponsors: Representatives Carlyle, Pollet, Bergquist, Hudgins and Riccelli.

Brief History:

Committee Activity:

Government Operations & Elections: 1/24/14, 1/31/14 [DPS];
Appropriations: 2/10/14 [DP2S(w/o sub GOE)].

Brief Summary of Second Substitute Bill

- Requires the Office of the Chief Information Officer (OCIO) to coordinate implementation and expansion of an Open Data Portal for publication of government data maintained by, or on behalf of, executive branch state agencies.
- Requires state executive branch agencies to designate a data coordinator and prepare a data catalog and compliance plans for making all public data sets available online.
- Requires the OCIO to develop technical standards to facilitate publication of digital data in the public domain or otherwise free of restrictions on reuse, so that it can be used and implemented by anyone.

HOUSE COMMITTEE ON GOVERNMENT OPERATIONS & ELECTIONS

Majority Report: The substitute bill be substituted therefor and the substitute bill do pass. Signed by 9 members: Representatives S. Hunt, Chair; Bergquist, Vice Chair; Young, Assistant Ranking Minority Member; Carlyle, Christian, Manweller, Orwall, Robinson and Van De Wege.

This analysis was prepared by non-partisan legislative staff for the use of legislative members in their deliberations. This analysis is not a part of the legislation nor does it constitute a statement of legislative intent.

Minority Report: Do not pass. Signed by 1 member: Representative Taylor, Ranking Minority Member.

Staff: Jasmine Vasavada (786-7301).

Background:

In 1972 the voters in the state of Washington adopted Initiative 276, requiring that most records maintained by state, county, city governments, and all special purpose districts be made available to members of the public. These provisions, now codified in chapter 42.56 RCW, are known as the Public Records Act (PRA). The PRA applies to all state and local agencies and defines a "public record" broadly to include, among other things, "any writing containing information relating to the conduct of government or the performance of any governmental or proprietary function prepared, owned, used, or retained by any state or local agency regardless of physical form or characteristics." Many PRA exemptions exist, targeted to protect from disclosure certain information, the nondisclosure of which the Legislature has deemed important for the protection of interests such as personal privacy, confidentiality, security, law enforcement and investigative functions, and commercial competition. State law also establishes procedures by which public agencies preserve, store, transfer, manage, retain, destroy, or dispose of public records. The Division of Archives and Records Management in the Office of the Secretary of State is tasked with managing, centralizing, and making available to the public various archival public records.

In 2011 the Office of the Chief Information Officer (OCIO) was created within the Office of Financial Management. The OCIO is responsible for the preparation and implementation of a strategic information technology plan and enterprise architecture for the state, as well as the development of statewide data security standards. The OCIO must establish standards and policies for enterprise architecture, educate and inform the state on information technology matters, and work to achieve standardization and consolidation of information technology infrastructure.

In May 2013, President Obama issued an executive order "making open and machine-readable the new default for government information." The executive order required the Office of Management and Budget, in consultation with the Chief Information Officer and others, to issue an Open Data Policy to advance the management of government information as an asset.

State agencies in Washington post public data sets online via a number of different web portals. For example, in addition to data.wa.gov, separate government websites host data and reports from the Education Research and Data Center, state geospatial information, the Caseload Forecast Council, the Legislative Evaluation and Accountability Program Committee, and the Office of the Superintendent of Public Instruction. However, there is currently no requirement ensuring that the data is presented in open, machine-readable formats, or ensuring that data be standardized in a manner that would facilitate the combination of information from different data sources.

Summary of Substitute Bill:

Open Data Portal.

The Chief Information Officer shall coordinate implementation and expansion of an Open Data Portal to facilitate the sharing and publication of government data in an open format. Public data sets made available on the open data portal should be offered as open data in an open format and free of cost, wherever feasible, consistent with standards and policies developed by the OCIO. "Open" means freely available, machine readable, and formatted according to uniform technical standards to facilitate visibility and reuse of publishable data. "Open format" means "a published, free file format for storing digital data, where such format is in the public domain or otherwise free of restrictions on reuse, and can be used and implemented by anyone."

The requirement to post government information on the Portal does not apply to all government records, but instead is limited as follows:

- "data" means "final versions of statistical or factual information that are in alphanumeric form reflected in a list, table, graph, chart, or other non-narrative form, that can be digitally transmitted or processed; are regularly created or maintained by or on behalf of an agency and controlled by such agency; and record a measurement, transaction, or determination related to the mission of the agency";
- "statistical or factual information," which does not include image files, such as designs, drawings, photos, or scanned copies of original documents, but it does include geographic information system data and statistical or factual information about such image files; and
- "public data set" means data maintained on a computer system by, or on behalf of, an agency that the agency is permitted, required, or able to make available to the public. Numerous categories of information are exempted from this definition, including data that is not required to be disclosed under the Public Records Act, data the disclosure of which would impose undue financial, operative, or administrative burden on the agency, and data related to externally-funded research and development activities; and
- "agency" means "every state office, department, division, bureau, board, and commission of the state, and each state elected official who is a member of the executive department."

Duties of the OCIO.

The OCIO shall:

- publish draft uniform standards for the open format of data and guidelines for state agencies to use in cataloging their public data sets and making the data sets publicly available in an open format, within 90 days of the effective date of the act;
- implement an online forum to solicit public feedback on open data policies and availability of public data sets, within 90 days of the effective date of the act. "Public data set" means data maintained on a computer system by, or on behalf of, an agency that the agency is permitted, required, or able to make available to the public. "Public data set" does not include: (1) any portion of a data set to which an agency may deny access pursuant to the PRA or any other provision of a federal or state law, rule, or regulation or local law; (2) any data set that contains a significant amount of data to which an agency may deny access pursuant to other law or rule, where redacting the

data set would impose undue financial, operative, or administrative burden on the agency; (3) data that reflects the internal deliberative process of an agency or agencies, including but not limited to negotiating positions, future procurements, or pending or reasonably anticipated legal or administrative proceedings; (4) data subject to copyright, patent, trademark, confidentiality agreements, or trade secret protection; (5) proprietary applications, computer code, software, operating systems, or similar materials; or (6) data related to internal agency administration, including employment records, internal employee-related directories or lists, and facilities data. Also, "public data set" is defined as data maintained by, or on behalf of, an executive branch agency that the agency is permitted, required or able to make available to the public, so this definition by its terms excludes data maintained by and on behalf of another branch of government;

- post on the open data portal and submit to the Governor and Legislature an updated compliance plan, within 18 months and annually thereafter;
- oversee and grant waivers for agency compliance with established timelines, if the agency demonstrates certain conditions are met;
- provide technical guidance, assistance, and expertise to assist agencies and local governments in participating in the Open Data Portal. This includes authority to adopt rules, models, and guidelines;
- include a mechanism for ensuring that, if more than one agency includes a public data set or portion of the data set in its data catalog, such multiple versions of the public data set are published, similar and equal access is provided to all versions of the public data set, and the data of one agency is not systematically privileged over the data of another;
- develop a data licensing policy which shall include guidance on use of a public domain dedication tool; and
- post legal policies conspicuously on the Portal.

Duties of Executive Branch Agencies.

Each agency shall:

- designate a data officer with authority equivalent to that of a deputy director or division or department head to be responsible for the agency's compliance (within 30 days of the effective date of the act);
- prepare a proposed compliance plan that includes a catalog of the agency's public datasets and a timeline for making each data set publicly available in an open format (within 15 months of the effective date of the act);
- prioritize publication of data based on the extent to which information serves various objectives, such as the creation of economic opportunity or the increase of agency accountability and responsiveness, its current availability in an open format, and other objectives deemed important by the OCIO. If there are public data sets not proposed to make public within five years, the agency must state the actions that would be necessary to make the public data sets available;
- plan for and implement, within existing resources, a process for making public data sets available in an open format. Funding to meet the costs of preparation of data in open formats should come primarily from a range of public and private funding sources listed in existing statute, and agencies are encouraged to seek federal and private grants; and

- retain, pursuant to normal operating procedures and policies governing data retention by the agency, underlying source data that has been posted on the Portal, if the agency changes this data.

Data Problems, Data Changes, and Legal Policies.

The OCIO must ensure the Portal includes the following features:

- a mechanism by which a member of the public who has identified deficiencies in a public data set can provide feedback to the OCIO about the deficiency;
- a log that records changes made by an agency after source data has been posted on the Portal, to include at a minimum: (1) the date the change was made; (2) a description by the agency of the change that was made, and (3) an explanation of why such change was necessary;
- a disclaimer of state liability for any deficiencies in a public data set and a disclaimer of warranties of the completeness or accuracy of information posted online; and
- notice that no private right of action is created.

Effect on Local Governments and Other Branches of State Government.

The new requirements do not apply to local governments or outside of the executive branch of state government. However, local governments and other branches of state government are encouraged to prepare a catalog and timeline for publication of local government data and to submit the data to the OCIO for inclusion in the Open Data Portal, and the OCIO is authorized to provide them technical guidance, assistance, and expertise.

Other.

A reference to a 1996 policy task force report is removed. A pilot project for electronic access to public records is decodified, due to the project being time-delimited.

Substitute Bill Compared to Original Bill:

Substitute House Bill 2202 makes changes in the following areas, as compared to the original bill.

Definitions.

The substitute bill:

- changes the definition of "open" to mean "freely available, machine-readable, and structured in a way that enables the data to be fully discoverable and usable by end users;"
- changes the definition of "public record" to remove reference to legislative and court records;
- changes the definition of "state agency" to clarify that it only includes executive branch entities;
- adds a definition for "open format" to mean "a published, free file format for storing digital data, where such format is in the public domain or otherwise free of restrictions on reuse, and can be used and implemented by anyone;"
- exempts from the definition of "public data set" certain data: (1) related to externally-funded research and development activities; (2) related to and specifically used for teaching; or (3) that cannot feasibly be converted to an open format without undue financial, operative, or administrative burden; and

- removes "maps" from the list of information that is excluded from the definition of "statistical or factual information."

Open Data Portal.

The substitute bill:

- provides that certain public data sets already available in open format on an existing state government data portal need not be separately uploaded onto the Open Data Portal;
- encourages the Office of the Chief Information Officer (OCIO), in developing uniform standards, to consider industry best practices and to develop standards addressing certain industry best practices;
- expands the bases for agencies' prioritization of data for publication;
- establishes that public data sets made available on the open data portal should be offered as open data in an open format and free of cost, wherever feasible, consistent with standards and policies developed by the OCIO; and
- requires the OCIO to include a mechanism for ensuring that, if more than one agency includes a public data set or portion of the data set in its data catalog, such multiple versions of the public data set are published, similar and equal access is provided to all versions of the public data set, and the data of one agency is not systematically privileged over the data of another.

Participation by Local Governments and Other Branches of State Government.

The substitute bill:

- encourages participation from other branches of state government, in addition to local governments; and
- authorizes but does not require the OCIO to provide assistance to local governments and other branches of state government.

Timelines and Compliance.

The substitute bill:

- requires agencies to provide proposed compliance plans to OCIO within 15 months instead of 180 days, but requires interim progress reports;
- requires the OCIO to post and report to the Legislature the compliance plans within 18 months, instead of one year, of the act's effective date;
- authorizes the OCIO to grant an agency a waiver from compliance with established timelines, if certain conditions are met; and
- clarifies that an agency may comply by appointing a data officer from within current staff, and provides that an agency may appoint a separate data coordinator who is directly accountable to the data officer.

Policies and Technical Guidance.

The substitute bill:

- requires the OCIO to develop a data licensing policy to facilitate the removal of barriers to access and re-use of government data, while providing adequate notice in the limited cases where the use of public data sets must be subject to license terms and conditions; and

- requires the OCIO to provide agencies guidance on the use of a standard public domain dedication tool to place the data into the public domain to the greatest extent allowed by law.

Changes to Data.

The substitute bill:

- prohibits the OCIO from making changes to the source data of a public data set posted by an agency, and establishes that only the agency who posts the public data set has authority to alter the source data;
- requires changes made by an agency to source data after it has been posted on the Open Data Portal to be recorded in a log that is readily accessible and includes, at a minimum, the date the change was made, a description by the agency of the change that was made, and an explanation of why such change was necessary; and
- requires the agency to retain, pursuant to normal operating procedures and policies governing data retention by the agency, the underlying source data that has been changed.

Other.

The substitute bill:

- includes additional statement of legislative intent, referencing Washington's history of ensuring open government and potential benefits of the Open Data policy; and
- codifies the new provisions in a new chapter, but does not recodify existing OCIO provisions into the new chapter.

Appropriation: None.

Fiscal Note: Requested on January 24, 2014.

Effective Date of Substitute Bill: The bill takes effect 90 days after adjournment of the session in which the bill is passed.

Staff Summary of Public Testimony:

(In support) This is not just a technology bill. It is about making data usable. When you look at transit information, applications on mobile devices to see when a bus is coming, so many uses are really about managing data. President Obama issued an executive order to create an Open Data policy, and the website is called data.gov, and its about making information usable. Washington State has been a leader, and we have a state portal called data.wa.gov. What is exciting is that this bill creates a default that data should be open and transparent, and with common standards, internal agency uses of data will become more effective. There is a macro-trend across the country, for which Washington is and should be a leader. In jurisdictions that have embraced this structure, there is a dramatic decline in Freedom of Information Act requests because members of the public are better able to find the information they need. This directs the Chief Information Officer to create a framework of a policy that will allow us to organize the public's data in a much more responsible way. It is permissive, in the sense that agencies have authority to manage their compliance in a way

that makes sense for their operation. The data covered by this bill is a great deal of the core data you see in charts and graphs. This may be the foundation of a new approach to e-government and efficient service delivery in a way that will not only save the public money but will make the information come alive.

Washington is one of the most transparent states in the nation; we value openness and we believe its important for good government. This bill is in line with those values. It is also in line with a trend where governments all over the world are changing the default to ensure that data is available in an open, accessible, and findable format. This bill holds agencies accountable to publishing data in a clear way, and is very clear about what data should be published. This bill is also an opportunity for the government to make our data open and easy to access, in a way that will have a positive economic impact. It has sparked opportunity for entrepreneurs to use the data in ways we cannot even imagine. Open data and the applications and services it fuels are already making a difference. In Seattle you can look up business licenses, parks, new development in your neighborhood—it helps people make important decisions in their lives. At the state level, open data will be critical to Results Washington. This policy will shape the openness, transparency, and accessibility of the people's data. The OCIO has worked closely with the Office of Financial Management to survey agencies to craft a meaningful fiscal note, but this bill is primarily a swap between ensuring that data is operational only to one agency and ensuring that the data is the people's data, available to all. This will make state government more effective.

You can see on the Seattle Police Department website information about 911 calls, police reports duly redacted under the Public Disclosure Act. In Washington, D.C., a private individual created an app called "Stumble Safely" that, based on your geographic information system data and the time and day of week, advises you on the incidence of crime on a given route. Here in Washington, there is an app that pulls up thousands of lakes, ponds, and reservoirs stocked with fish, species and images, and recommends the best time to fish. Much of the underlying data came from the state, but it would be a lot easier to create these applications if the data were online in an open format. New York City has an open data law or ordinance that is similar to this one and they have published their data sets online. San Francisco and Oakland also publish their data sets. This meshes well with the Governor's policy goals for the Results Washington website, to make more data sets be open and transparent and available on data.wa.gov.

This bill will be critical to reducing our long-term costs in the public records arena. Disclosure will not be as much of a problem as it has been in the past 10 years, during which time the rates of public requests have accelerated. Quality of information will improve, and questions of what is redactable will be addressed at the front-end and not each time a request comes in, so the costs of compliance with the PRA will go down. The reason courts are not included is that they have decided that they are not a state agency, and they created a separate commission to govern the release of court records. It is important that you can see conflicting data sets and the public can reconcile them on their own, instead of asking government to reconcile and seek truth. The legal policy in this bill is that "we don't know, and you can't hold us accountable." But the public is able to see the truths and the mistakes made by agencies, and hold the government accountable. It does not change any exemptions under the PRA.

(Neutral) When two agencies have conflicting data, it is often a question of whether data is being described accurately. Time and time again, the description of the data is too simplified and the "devil is in the details". Sometimes the data may be correct on the chart that the agency has provided, but subsequent characterizations are inaccurate. Agencies make a constant effort.

(With concerns) Higher education institutions support the intent and philosophy of this bill, because we have seen the benefits of providing open data and what it can do for the public. Washington has already created a nationally-recognized higher education dashboard. Applications like "One Bus Away" helps thousands of people get to work on time through the King County metro and transit systems. Some concerns that need to be addressed are potential costs and narrowing the "open data set" definition so that we will not have to convert data sets to open format and subsequently publish those. The definition in the bill goes beyond routine agency practice and open data sets that are already provided to the public by higher education agencies.

(Opposed) None.

Persons Testifying: (In support) Representative Carlyle, prime sponsor; Michael Cockrill and Bill Schrier, Office of the Chief Information Officer; and Rowland Thompson, Allied Daily Newspapers.

(Neutral) Katie Blinn, Office of the Secretary of State.

(With concerns) Margaret Shepherd, Council of Presidents.

Persons Signed In To Testify But Not Testifying: None.

HOUSE COMMITTEE ON APPROPRIATIONS

Majority Report: The second substitute bill be substituted therefor and the second substitute bill do pass and do not pass the substitute bill by Committee on Government Operations & Elections. Signed by 19 members: Representatives Hunter, Chair; Ormsby, Vice Chair; Chandler, Ranking Minority Member; Carlyle, Cody, Dunshee, Green, Haigh, Hudgins, S. Hunt, Jinkins, Kagi, Lytton, Morrell, Pettigrew, Seaquist, Springer, Sullivan and Tharinger.

Minority Report: Do not pass. Signed by 11 members: Representatives Ross, Assistant Ranking Minority Member; Wilcox, Assistant Ranking Minority Member; Buys, Christian, Fagan, Haler, Harris, G. Hunt, Parker, Schmick and Taylor.

Staff: Alex MacBain (786-7288).

Summary of Recommendation of Committee On Appropriations Compared to Recommendation of Committee On Government Operations & Elections:

The second substitute bill:

- Exempts data or data sets from publication on the Open Data Portal (Portal) if an agency head determines that publication would be detrimental to the public interest, and the agency publishes a detailed rationale in support of its determination;
- Requires each agency with custody of a public data set and authority to update the public data set to include it in the agency's data catalog, instead of only encouraging each agency to include such data when multiple versions exist;
- Encourages the Office of the Chief Information Officer (OCIO) to develop standards that address a method of time-stamping data set changes; and
- Adds a null and void clause that makes the bill null and void unless funded in the budget.

Appropriation: None.

Fiscal Note: Available.

Effective Date of Second Substitute Bill: The bill takes effect 90 days after adjournment of the session in which the bill is passed. However, the bill is null and void unless funded in the budget.

Staff Summary of Public Testimony:

(In support) There is a major national trend driven by a website called data.gov, which opens up public information. The default for statistical or factual information should be open and accessible. Washington has been a leader in this area, but the information that is available on the data.wa.gov website is not searchable and readable. This bill directs the OCIO to work with agencies to determine what information should be public and to prioritize that information. Every year there are four to five bills requiring an individual type of data set to be available to the public. Under this foundational bill, all of the data sets would be eligible to be public without individual legislation. The substitute bill that passed out of the Government Operations Committee had a nine to one vote. The fiscal note for the bill is indeterminate. There has been some information from individual agencies that costs could be high. Those estimates are being driven by fear, doubt, and uncertainty about the timelines in the bill. In the substitute bill, the timelines have been pushed out to provide plenty of time. The proposed second substitute includes a process for exempting data sets which should also help reduce costs.

(Opposed) As an agency trying to assess the fiscal impact, the issues driving costs include the amount of data sets that would ultimately be made public on the data portal and the complexity of doing so. The cost estimate for the Department of Natural Resources (DNR) was about \$1.7 million per year for at least the first three years. The DNR is a very data rich agency with about 200 terabytes of data, so the costs to identify, compile, and prepare the data to be easily searchable by the general public are estimated to be relatively high.

Persons Testifying: (In support) Representative Carlyle, prime sponsor.

(Opposed) Lisa Largent and Lenny Young, Department of Natural Resources.

Persons Signed In To Testify But Not Testifying: None.