

Queensland needs a right to data law

Author: Stephen Gates

Date: January 2015

Disclosure: I currently work for the Queensland Department of Transport and Main Roads. This work reflects my personal views and not those of my employer.

Licence: This work is licensed under the Creative Commons Attribution 4.0 International License. View this license at <http://creativecommons.org/licenses/by/4.0/>

Available at: https://dl.dropboxusercontent.com/u/67526403/OpenData/RightToDataLaw_StephenGates.pdf

Thank you for the opportunity to comment on the proposal to introduce open data legislation as outlined in the discussion paper, *A new law to cement open data?*¹.

I support the introduction of legislation to significantly increase confidence in the availability and quality of Queensland's open government data to generate economic and social benefit.

I understand that the benefits of the new law would be achieved by:

- Formalising a data-publishing framework.
- Requiring government agencies to make open data available and easily accessible.
- Promoting the publishing and sharing of open data consistent with open data principles.
- Requiring a dialogue with registered users of open data prior to a dataset being changed or removed (except for data updates, corrections and unintended disclosures).

A set of open data principles (i.e. government agency responsibilities) are proposed covering:

- Publishing data (open by default; maintaining data currency).
- Protecting privacy and security.
- Providing easy access (one central location; ideally free; minimal restrictions and technical barriers).
- Maintaining Open Data strategies (development, publishing and annual reviews).
- Making no representations or warranties about published data.
- Promoting the use of open data and the open data principles.

I would like to suggest a number of improvements for inclusion in the legislation, the principles and the operational systems that implement the new law. These suggestions include:

1. A definition for open government data.
2. Open licences to maximise reuse.
3. Better metadata to improve confidence in using open data.
4. Open standards to unlock innovation.
5. Using competition to drive continuous improvement.
6. Open discussions to support the use and improvement of open data.
7. Open to everyone – an inclusive open data platform for non-technical users.

1. An Open Government Data definition

The discussion paper¹ states, *“a myriad of definitions for open data have been published by organisations and agencies around the world”*. I believe legislation about open data needs to include a definition of open data. A vague definition of open data was one of the flaws in America’s open data implementation, according to Alon Peled in, *Re-Designing Open Data 2.0*².

I propose Queensland Government adopt the *Open Definition*³, to define open data.

The Open Definition sets out principles that define “openness” in relation to data and content. It makes precise the meaning of “open” in the terms “open data” and “open content” and thereby ensures quality and encourages compatibility between different pools of open material.

It can be summed up in the statement that: “Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness).”

Queensland Government open data should be defined as, *Open Government Data*⁴.

Open Government Data means:

- *Data produced or commissioned by government or government controlled entities.*
- *Data which is open as defined in the Open Definition – that is, it can be freely used, reused and redistributed by anyone.*

2. Open Data needs an Open Licence

One principle outlined in the discussion paper¹ aim is to, *“ensure there are minimal restrictions or technical barriers for published open data”*.

I propose that this principle explicitly state that open data requires an open licence.

Open Knowledge has documented a set of *Open Licences*⁵ that conform to the *Open Definition*³. These licences are aligned with the *AusGOAL Suite of Licences*⁶, recommended by the Australian Government Open Access and Licensing Framework (AusGOAL).

The *Premier’s message*⁷ about open data requires agencies to, *“apply clear licences (preferably open licences such as Creative Commons)”*.

3. Better metadata to improve confidence in using open data

Another of the discussion paper’s¹ open data principles encourages government agencies to, *“maintain the currency of open data”*.

The frequency of update is just one metadata attribute that describes the quality of open data. This and other metadata elements are the data publishers promise to data re-users, helping them:

- Discover the data and understand its scope.
- Understand the data quality to determine if it is fit for reuse.
- Understand the structure of the data so it is easy to use (e.g. column names and meanings, data types, valid values, code look up tables, and relationships to other data).
- Appreciate how the data may have been changed as part of the publishing process (e.g. anonymisation).

I propose a new open data principle requiring all open data to be published with standards-based metadata that assists in the data being discovered, and its structure, meaning, quality and provenance being understood.

The use of standards-based metadata will enable the Queensland Government to automatically earn Open Data Certificates. The criteria to earn a certificate are based on a number of global best practices⁸ that have been consolidated by the Open Data Institute.

Open Data Certificates provide a number of benefits for open data re-users, publishers and policy makers as outlined in the, *Introduction to Open Data Certificates*⁹, and include:

1. Re-users will be informed of:

- What they can legally do.
- Whether they can rely on the data.
- What the data means.
- Where to get some support.

2. Publishers will be provided with:

- A simple, semi-automated process to earn certificates.
- A badge to display with their data to indicate its quality.
- Automated suggestions for the next steps to improve data quality.

3. Policy makers will have a mechanism to:

- Use Open Data Certificates as shorthand for set of standard requirements.
- Set a standard for publishing open data, potentially based on the data's value.
- Set a standard for the procurement of new systems and data.
- Use as the basis to measure the quantity and quality of open data.

Deloitte, in *What do businesses really look for in open data?*¹⁰, supports the use of Open Data Certificates if businesses are expected to take open data seriously.

I propose that the Queensland Government create a policy to ensure that all of its open data is capable of attaining an ODI Open Data Certificate. This will demonstrate that the data is current, aligns with the open data principles and provide confidence to businesses to utilise Queensland's open government data, stimulating the economy and providing value to the community.

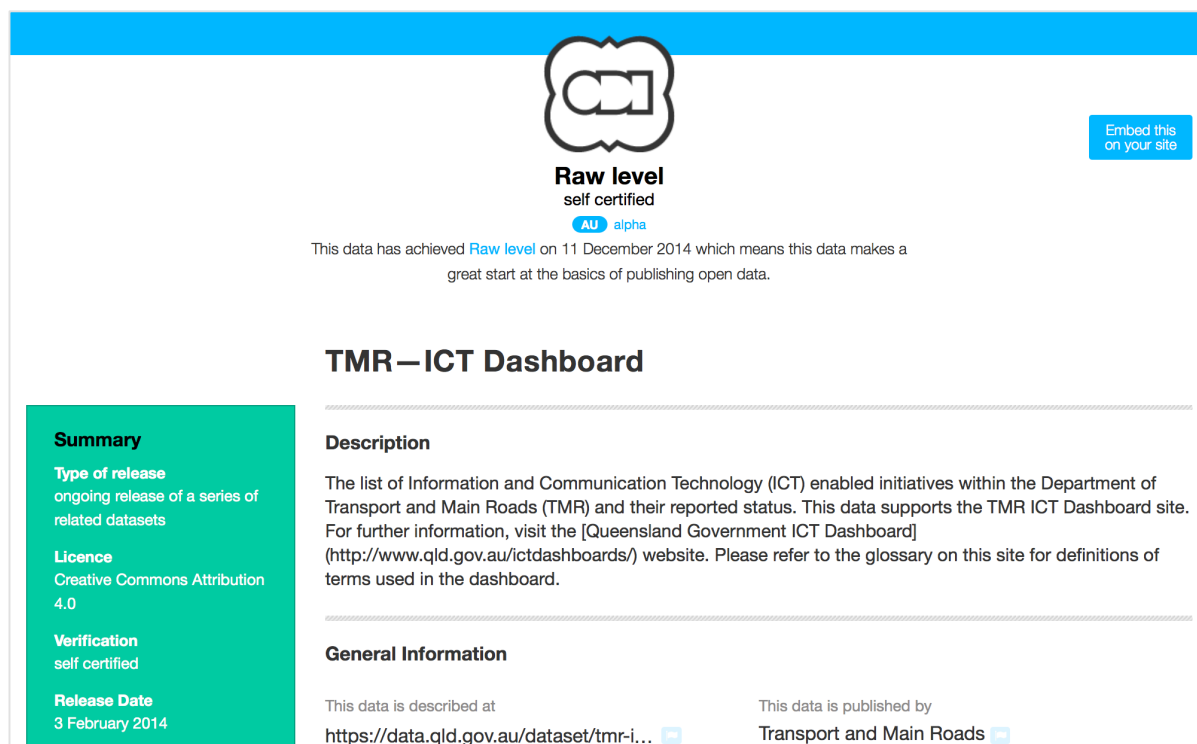


Figure 1. The first Australian Open Data Certificate was awarded to the Department of Transport and Main Roads for the ICT Dashboard data in December 2014 and can be viewed at <https://certificates.theodi.org/datasets/23694/certificate>

4. Open Standards to unlock innovation

As highlighted in McKinsey & Company's report, *Open Data: Unlocking innovation and performance with liquid information*¹¹, making data open is insufficient for capturing its full value. Other actions are required, such as:

- Creating a suitable platform for sharing data and to perform analysis.
- Mandating standards to make data from multiple sources easier to combine and compare.
- Providing metadata to make open data more easily used.

Complying with data standards enables value to be generated from open data through the new insights gained from combining multiple datasets. Data standards enable:

- Similar data across government agencies to be combined (e.g. grants, contracts, budgets).
- Similar data to be combined across different levels of government.
- Related data from business, community groups and governments to be combined (e.g. to support forming a combined response to natural disasters).
- Historic analysis to be performed but only when the data is retained in a permanent location and a stable format, as proposed by the Sunlight Foundation's, *Open Data Policy Guidelines*¹².

I propose that the open data principles be strengthened to require the use of International, Australian, or other commonly used data standards, where available, to unlock innovation.

5. Competition to drive continuous improvement

Countries, regions and cities currently compare progress through the *Open Data Census*¹³. They are scored on how well they publish and engage around key open datasets. I propose that competition be used to drive improvement within Queensland Government agencies and across state governments.

Competition within Queensland Government agencies

A league table could be established to measure the performance of each department on publishing open data and adhering to the open data principles. Measures could include:

- The number of datasets and resources published.
- The quality of the data published (e.g. the level of Open Data Certificate achieved).
- The utilization of datasets (e.g. number of downloads or access requests).
- The number of case studies published demonstrating the value provided by Queensland’s open government data.

Competition between States

Free assessment tools, such as the Open Data Census, could be used to compare progress between State Government open data initiatives. Personally, I would love to see Queensland win the first Open Data State of Origin.

The World Wide Web Foundation and the Governance Lab brought together open data assessment experts to develop common methods for assessing open data¹⁴. I recommend reviewing this and subsequent works in developing a competitive framework to drive continuous improvement.



Some open data assessment results that could be applied at a State Government level are shown below.



Figure 2. Global Open Data Index: Survey, by Open Knowledge, is Public Domain, viewed at <http://global.census.okfn.org>

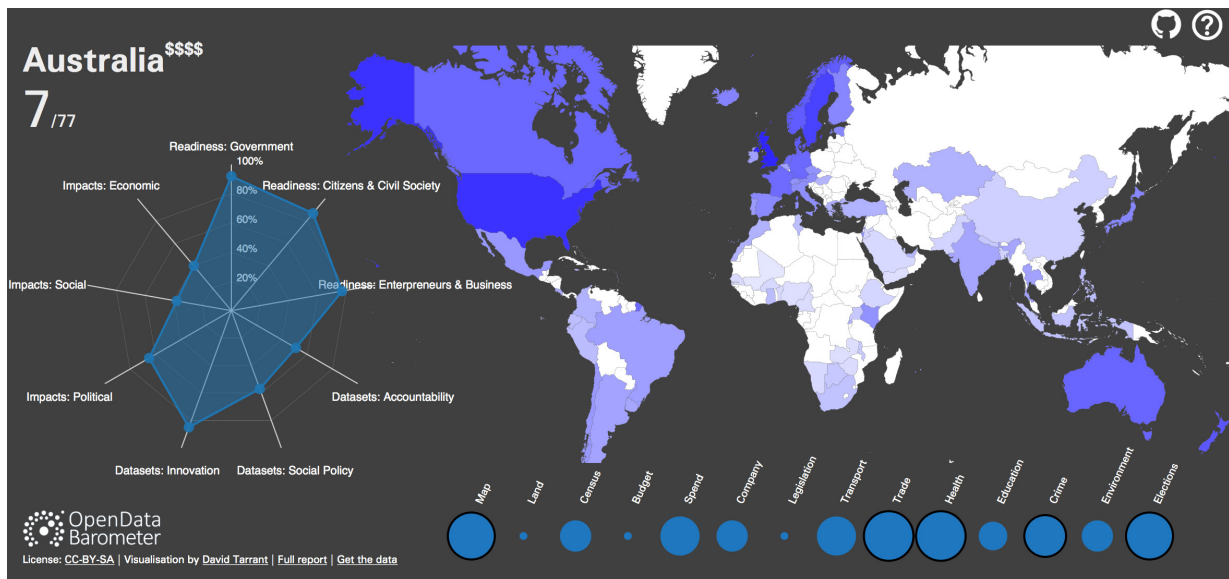


Figure 3. Open Data Barometer Data, by the Open Data Institute, used under Creative Commons Attribution-ShareAlike 2.0 UK: England & Wales viewed at <http://theodi.github.io/open-data-barometer-viz/>

6. Open Conversations

Conversations between data publishers and data re-users are currently hidden in emails or web forms. There is no open way for conversations to be held about a specific dataset. The benefits of having open conversations are:

- Data re-users can support each other, potentially reducing enquires to government.
- Data publisher responses are visible to all data re-users, avoiding the need to answer the same question multiple times.
- Data re-users can submit ideas for improvement and based on levels of support from others, improvements can be better prioritised.

There are some risks in having open conversations on the Internet but this can be mitigated by controls such as moderating comments before they are visible to the public or automatically checking for the use of bad language.

I propose the Queensland Government Open Data Portal be enhanced to support open conversations.

7. Open for everyone

In line with the principles of minimising technical barriers and promoting the use of open data, I propose the government consider providing improved visualisation and analysis tools for open data. These tools would support the broader use of open data by community groups and others who may not have access to expensive technology required to explore different combinations of open and closed data on charts, maps, and other analysis techniques.

8. Conclusion

I congratulate the Queensland Government on the progress made in the open data initiative. I encourage you to consider the proposals in this response. I look forward to a new right to data law that provides certainty around the publishing of open government data.

¹ *A new law to cement open data?* by The State of Queensland, is used under the CC Attribution-NoDerivs 3.0 Australia Licence viewed at <http://queenslandplan.qld.gov.au/gov-response/assets/Open-Data-Discussion-Paper.pdf>

² *Re-Designing Open Data 2.0*, by Alon Peled, viewed at <http://www.jedem.org/article/view/219/180>

³ *The Open Definition*, by Open Knowledge, is used under the CC Attribution 4.0 International License. The Open Definition is summarised at <http://opendefinition.org> and the formal definition can be viewed at <http://opendefinition.org/od/>

⁴ *Open Government Data*, by Open Knowledge, is used under the Creative Commons Attribution 3.0 License (all jurisdictions), viewed at <http://opengovernmentdata.org>.

⁵ *Conformant Licences*, by Open Knowledge, viewed at <http://opendefinition.org/licenses/>

⁶ *AusGOAL Suite of Licences*, by AusGOAL, viewed at <http://www.ausgoal.gov.au/the-ausgoal-licence-suite>

⁷ *Premier's message*, by The State of Queensland, viewed at <https://data.qld.gov.au/department-strategies/about>

⁸ *Open Data Certificate – Vision*, by the Open Data Institute, viewed at <https://github.com/theodi/open-data-certificate/wiki/Vision>

⁹ *Introduction to Open Data Certificates*, by the Open Data Institute, viewed at <https://www.scribd.com/doc/246127801/Introduction-to-Open-Data-Certificates>

¹⁰ *What do businesses really look for in open data?* by Harvey Lewis, Deloitte, viewed at <http://www.computerweekly.com/blogs/Data-Matters/2014/11/what-do-businesses-really-look-for-in-open-data.html>

¹¹ *Open data: Unlocking innovation and performance with liquid information*, by McKinsey & Company, viewed at http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information

¹² *Open Data Policy Guidelines*, by the Sunlight Foundation, viewed at <http://sunlightfoundation.com/opendataguidelines/>

¹³ *Open Data Census – Meta*, by Open Knowledge, viewed at <http://census.okfn.org>

¹⁴ *Towards common methods for assessing open data: workshop report and draft framework*, by The Governance Lab, viewed at <http://thegovlab.org/towards-common-methods-for-assessing-open-data/>