

# Australian Council for Educational Research (ACER)

---

From the Selected Works of Professor Kathryn Moyle

---

November 30, 2012

## Open data and open standards

Kathryn Moyle, *Charles Darwin University, Australia*

Leigh Blackall

Felix Hudson

Jon Mason, *Charles Darwin University, Australia*



This work is licensed under a [Creative Commons CC BY](https://creativecommons.org/licenses/by/4.0/) International License.



Available at: [https://works.bepress.com/kathryn\\_moyle/16/](https://works.bepress.com/kathryn_moyle/16/)

## Open Data and Open Standards

Report to Web Science Australia

of research conducted by the Centre for School Leadership, Learning and Development

Researchers:

Leigh Blackall, Felix Hudson and Jon Mason

Report prepared by

Kathryn Moyle, with assistance from Leigh Blackall, Felix Hudson and Jon Mason

30 November 2012

## Acknowledgements

This report has been prepared on behalf of the Institute of Metadata Management (IMM).

The researchers and authors would like to thank the Reference group for their support, suggestions and critical feedback. In particular:

- Anni Rowland-Campbell: Director, Interstica
- Armin Haller, Research Scientist Information Engineering Lab, CSIRO ICT Centre
- Oliver Bell, CTO, International Organizations, Microsoft Public Sector.

The researchers and authors would also like to express their thanks to all those who freely contributed to the survey, so that it was possible to gain insights into how different open data projects work and are progressing.

We hope you will find this report stimulating enough to want to go out and find out more about open data.

Professor Kathryn Moyle  
Executive Director,  
Centre for School Leadership, Learning and Development  
Charles Darwin University

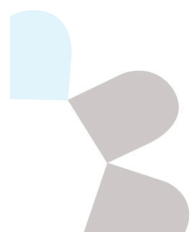


Centre for School Leadership, Learning and Development, 2012

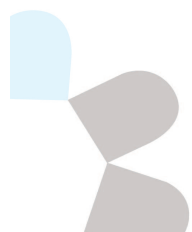


## Abbreviations

<b>ANDS</b>	Australian National Data Service
<b>API</b>	Application Programming Interface. Middleware language and format used by an application program to access another program.
<b>ArcInfo</b>	An ESRI product. Geographic Information System (GIS) software.
<b>AtomPub</b>	Atom Publishing Protocol.
<b>Catalogue</b>	A listing of agency data holdings
<b>CGM</b>	Computer Graphics Metafile. Open standard file format for two dimensional vector and raster graphics.
<b>CKAN</b>	Open-source data portal platform
<b>CSV</b>	Comma Separated Values. Plain text file that stores tabular data delineated by commas.
<b>D2RQ</b>	A platform for accessing relational databases as virtual, read-only RDF graphs
<b>DDI</b>	Data Documentation Initiative
<b>EMF</b>	Enhanced MetaFile
<b>EPS</b>	Encapsulated Post Script. Image description format for vector based graphics.
<b>FTP</b>	File Transfer Protocol
<b>GData</b>	Google Data Protocol. Relies on XML or JSON.
<b>Geodatabase</b>	Database containing geographic information and spatial data.
<b>GeoRSS</b>	Coding to enable the request, aggregation, sharing and mapping of geographically tagged feed
<b>GIS</b>	Graphic Information System
<b>HTML</b>	HyperText Markup Language. Cross platform language code used to create electronic documents for the Web.
<b>JPEG</b>	Joint Photographic Experts Group. Graphics file format that compresses digital images for transfer over the internet.
<b>JSON</b>	JavaScript Object Notation.
<b>KML/KMZ</b>	Keyhole markup language, suitable for viewing in Google Earth, Google Maps or other supported applications
<b>LOD</b>	Linked open data. See <a href="http://linkeddata.org/">http://linkeddata.org/</a>
<b>MapInfo</b>	MapInfo product. Geographic Information System (GIS) software package.
<b>MIF</b>	Maker Interchange Format
<b>OAI-PMH</b>	Open Archives Initiative Protocol for Metadata Harvesting



<b>OData</b>	A web protocol for querying and updating data, to provide access to information from a variety of applications, services, and stores. It applies to protocols such as HTTP, Atom Publishing Protocol (AtomPub) and JSON. Microsoft has released OData under its Open Specification Promise (OSP).
<b>ODS</b>	Open Document Spreadsheet
<b>OWL</b>	Web Ontology Language
<b>PDF</b>	Portable Document Format. Adobe Acrobat product. Universal file format that maintains document appearance regardless of the operating system used.
<b>PNG</b>	Portable Network Graphics
<b>RDF</b>	Resource Description Framework. Uses include modelling information and to represent metadata about digital artifacts
<b>RIF-CS</b>	Registry Interchange Format - Collections and Services
<b>RSS</b>	Rich Site Summary/Really Simple Syndication
<b>SDMX</b>	Statistical Data and Metadata Exchange (SDMX) model
<b>SEGY</b>	SEGY. Standard file format for storing geophysical data.
<b>Shape</b>	Shapefile. ESRI product. Geographic vector data format.
<b>SHP</b>	Shape file
<b>SPARQL</b>	SPARQL Protocol and RDF Query Language
<b>TIFF</b>	Tagged Image File Format
<b>TXT</b>	Raw Text files
<b>URI</b>	Uniform Resource Identifier
<b>VOMAP</b>	VO Metadata Application Profile
<b>W3C</b>	World Wide Web Consortium, who develop interoperable technologies, guidelines, software and tools.
<b>WFS</b>	Web Feature Service
<b>WMS</b>	Web Map Service
<b>XLS and XLSX</b>	Excel. Microsoft Office product. Spreadsheet file.
<b>XML</b>	Extensible Markup Language. Open standard text format derived from SGML used to define data elements on a Web page.
<b>Zip</b>	File compression format



## Executive Summary

This small study set out to gain an overview of which organisations within Australia and around the world are actively undertaking 'open data' initiatives. Data was collected through a desktop review of existing open data portals; and through an online survey. Many of the agencies reviewed for this project are government departments that seem to be in the first stages of their open data projects; however, there is also a broader base of relevant initiatives not associated with open government. Data from the not-for-profit organisation WikiLeaks has also been reviewed also submitted data.

There seem to be differing purposes for releasing open data. Most government-released datasets appear to consist of data that already existed prior to the start of their open data projects e.g., historical data, data created undertaking their normal business activities, and data that had already been published elsewhere (e.g., budgets and road toll figures). Government open data projects therefore seem to be making public, existing data as a public good. These initial releases of data are sometimes being used to promote the sharing of other datasets, including from within the same organisation. WikiLeaks on the other hand, is not generating its own data and has different motivations to those of governments, but is nonetheless accessing and harvesting large datasets and distributing these openly.

There are various levels of complexity in the delivery of open datasets. The most basic is to simply provide a download link to the various datasets. More sophisticated methods involve features that can query the data in real time, or allow the user to retrieve only the data in which they are interested, compared to having to download all of the available data, and have to sort through it manually.

Open data is comparable to many other kinds of published content. All content shares common attributes like authorship, ownership, publication origins, classifications and licences. Successful publishing in areas such as books, research reports, websites, and social media is due in part, to addressing and resolving these issues of provenance. Open data projects are likely to be successful when they identify and solve issues surrounding these 'ownership' and 'use' attributes.

A common issue for open data projects seems to revolve around the licensing of existing data. As with all publishable data, issues such as who owns it, what are the privacy restraints, and what may be the down-the-line publishing contexts, all have to be addressed if the open data initiative is to be sustainable and dynamic.

The findings from this study suggest that the hosting location does not really matter, as the open datasets are intended to be shared. As such, secure hosting is less important than reliable hosting, although URL integrity is an issue for services requiring remote access. There also seems to be a higher amount of 'international cloud' hosting than other online environments might show. This maybe due to the open data movement starting around the same time as cloud hosting became a popular storage alternative to local terrestrial solutions.

Most of the open data projects reviewed were relatively young and immature, with several Australian state government departments only commencing their open data projects in 2011 or 2012. Large open data projects such as those managed by the US Government and associated agencies such as NASA are moving to enabling users to personalise the datasets they download. These initiatives go beyond discovering data to linking data to address real world issues. Although there does not seem to be a body of existing evidence at present about how open data may contribute to public or private sector efficiencies, the ability of users to freely and easily access, retrieve and analyse complex datasets does suggest that new opportunities for innovation are opening up.



## Introduction

Governments in Australia and around the world are publishing data they routinely collect, in open formats. Some government agencies are publishing datasets to provide more services to the public, as part of their open government strategies. Open data is also gaining credibility among academic researchers. Indeed on May 22, 2012 at the University of North Texas, academics, researchers, technology officers and librarians released the *Denton Declaration: An Open Data Manifesto* (University of North Texas, 2012), to promote open access to research data.

This research project set out to develop an overview of which organisations globally and within Australia are actively investigating, developing, undertaking and promoting 'open data' initiatives. The study was specifically interested in identifying the key organisations promoting open data in order to gain insights into how actively they are doing this, and the maturity of their open data activities.

Key issues about which the project sought information included:

- the range of storage and publishing formats being made available to enable open data publishing, access and management;
- the protocols supporting these initiatives (e.g., AtomPub, Gdata, Odata); and
- the metadata strategies deployed around the open data frameworks such as how these metadata strategies are being developed, managed and updated .

The research project was also interested in investigating the articulation of how the organisations expect to leverage their investment in open data to improve efficiencies.

This study was commissioned by the Institute of Metadata Management, which is a community of practice around the "... creation, management, use, promotion, capability development, innovation and articulation of information..." for better governance (Institute of Metadata Management, 2011, p1).

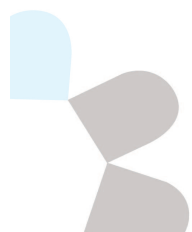
This study found that open data projects are growing in presence all around the world. Many of these projects are concerned with linking open data. As such, there is not a binary distinction between open and linked datasets. Both concepts are emerging leading to sophisticated uses of open data via the semantic web. At the same time, while some open data projects are continuing to allocate metadata manually, there is also an emerging tendency to automating the processes of metadata allocations. In Australia, the Australian National Data Service (ANDS) is at the forefront of some of this work.

## Defining the field

Like a lot of other terminology in contemporary settings, the phrase 'open data' gets used in different contexts and is used for a diversity of purposes. Meanings of open data are also linked to other key ideas and technical terms such as 'open government' and 'open metadata'.

### Open data

In this study, 'open data' was defined widely as that data that is freely available for people to use and republish without restrictions from copyright, patents or other mechanisms of control. Throughout the project however, this problem of definition was left open for review while looking at



a range of projects, to examine the ways in which they defined ‘open data’. Desktop research also reveals that there is a broader ‘open agenda’ that has been evolving since the 1980s when open systems and open architecture informed the development of open (interoperable) standards development and the emergence of open source software developments. The emergence of ‘open access’ and ‘open educational resources’ as important drivers of change within the education sector is also closely aligned with aspects of open data.

## Open Government

Open data initiatives are commonly linked to concepts of ‘open government’ where the provision of government services is seen to be underpinned by collaboration, participation and transparency. One of the main aims of government open data initiatives is to make public data universally accessible in easily accessible and engaging ways. This report focuses upon the government initiatives and software used to support open data initiatives.

## Open metadata

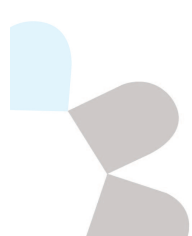
Open metadata is a key to being able to access and harvest open data. The foundation for this was first established by the Open Archives Initiative (OAI) in the development of the *OAI Protocol for Metadata Harvesting (OAI-PMH)*, first released in 1999, and widely adopted by government, cultural, and educational organisations since (such as the National Library of Australia). The critical feature of OAI-PMH is the exposure of metadata records enabling robots to automatically harvest it. The Wikibooks (2012, p1) *Open Metadata Handbook* defines ‘open metadata’ as

*“a piece of content or data is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and share-alike.”*

Such a definition gives emphasis to ownership and masks the obvious fact that the data actually has to be accessible in the first place. However, this *Handbook* (2012) also suggests conditions have to be fulfilled for metadata to be considered open. These include the following conditions:

- It is freely or cheaply accessible: i.e., it is publicly available preferably via the Internet, and is available free of charge (or available at a reasonable reproduction cost).
- It can be freely redistributed: i.e., it is possible for anyone to freely redistribute the metadata, either as it is, or as part of a broader dataset derived from many different sources.
- It can be reused: i.e., it is possible to modify or incorporate the metadata into derivative datasets, which can be distributed under the same terms as the original metadata.
- There are no technological barriers to its use: i.e., it is provided in such a form that it is interoperable having no technological obstacles to the performance of the metadata.
- The license must not restrict the distribution of other works: i.e., there are no restrictions on other works that are distributed along with the licensed dataset.

Open metadata can be considered the metaphorical cotton that holds together open government and open data, and enables data to be easily accessed and used. Importantly, such metadata is both human readable and machine executable. Metadata is also important for being able to link datasets, and as such has an important place in this study.





## Research approach

This study used two methods to collect the information included this report:

- a desktop, online review of what cities, states and national governments, as well as other quasi-government organisations and institutions are undertaking in relation to open data agendas; and
- a survey posted online and promoted through open source networks.

The research adopted an open approach to this open data research project by seeking to collect data using Wikibooks. The aim was to draw volunteers into the project, and to populate the survey using 'crowd sourcing' methods. Progress was documented in a Wikiversity space:

[http://en.wikiversity.org/wiki/Open\\_Data](http://en.wikiversity.org/wiki/Open_Data). Suggestions for sourcing active open data projects were also received from the Reference Group for this project. The fields for seeking data that were used in the online survey are available in Appendix One.

## Findings of desktop review

Organisations around the world are investigating, developing, undertaking and/or promoting open data initiatives. Responses to the online survey provided detailed information about the following 9 open data initiatives. Summaries of these projects can be found later in this report. Further information about other open data projects can be found at the Wikiversity site mentioned above.

Auspicing agency	Name of project
Australian Capital Territory Government	DataACT
Australian Government	Linked data for open government
Australian National Data Service (ANDS)	Research data
Italian National Research Council	Data.cnr.it
JISC	PROD
Centre for School Leadership, Learning and Development, Charles Darwin University	Open data project
South Australian Government	AusGOAL-DATA.SA
UK National Environment Research Council	Open Geoscience
WikiLeaks	WikiLeaks

The following agencies' open data initiatives were reviewed online.

Auspicing agency	Name of project
British Government	<a href="http://data.gov.uk">data.gov.uk</a>
European governments	<a href="http://www.publicdata.eu">www.publicdata.eu</a>
Government of Victoria	<a href="http://data.vic.gov.au">data.vic.gov.au</a>
New South Wales Government	<a href="http://data.nsw">data.nsw</a>
New Zealand Government	<a href="http://Data.gov.nz">Data.gov.nz</a>
NASA – United States of America	<a href="http://data.nasa.gov">data.nasa.gov</a>
South Africa not-for profit	Open Data & Democracy Initiative
United States of America Federal Government	<a href="http://Data.gov">Data.gov</a>
World Bank	<a href="http://data.worldbank.org/">http://data.worldbank.org/</a>

The following approaches within the respective open data initiatives were identified for further review:

- storage of data;
- publishing formats of data;
- access to and management of data;
- protocols;
- metadata management strategies; and
- maturity of the open data initiatives.

## Storage

Judging by the projects reviewed in this study, it appears that storage of the data is being undertaken in several different ways:

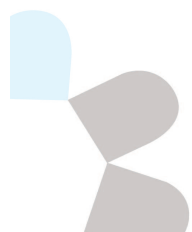
- international cloud hosting;
- national data collection/aggregation services; and
- local hosting.

Local data hosting and internal cloud hosting are being used by individual agencies undertaking specialised work, e.g., in science, or are leveraging national datasets. The state governments of Australia appear to be linking into a national approach to data storage led by the Australian Government.

## Publishing formats

A common suite of publishing formats are used by all the projects reviewed in this study. A list of common publishing formats is included in the 'Abbreviations' earlier in this report. The most common formats seen in the projects reviewed, however, are as follows:

- |           |   |
|-----------|---|
| • CSV/XLS | Comma Separated Values. Plain text file that stores tabular data delineated by commas.                              |
| • KML/KMZ | Keyhole markup language, suitable for viewing in Google Earth, Google Maps or other supported applications.         |
| • RDF     | Resource Description Framework, used for modelling information  |
| • TXT     | Raw Text files  |
| • XML     | Extensible Markup Language. Open standard text format derived from SGML used to define data elements on a Web page. |



## Access and management

Different projects reported different licensing arrangements for the use of the available datasets. While the Creative Commons Attribution is common for several projects, the Open Geoscience project in the UK has its own terms and conditions. Of those projects reported through the online survey:

- none identified that they have 'all rights reserved';
- three reported they are in the public domain;
- four reported the use of 'Creative Commons Attribution';
- one reported the use of 'Creative Commons Attribution Non Commercial'; and
- two reported the use of 'Creative Commons Attribution No Derivatives'.

## Protocols

Most if not all of the projects use common web standards to access and publish the datasets and metadata. These standards consist of HTTP utilising HTML, RSS and XML. While other access methods (such as ftp) may be provided they are complementing the existing mechanisms.

## Metadata strategies

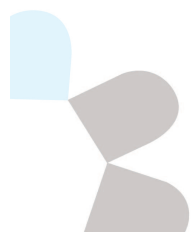
The study investigated metadata strategies deployed around the 'open data' frameworks to see how these metadata strategies are being developed, managed and updated. Metadata has been used by all of the identified open data projects. Certain common metadata attributes such as ownership, publication date, title and data format are implemented on almost all websites. Other more specific types of metadata such as geo-location metadata are used most often when the dataset has some kind of geographic data component. For example boundary charts, railways and road data.

The metadata strategies identified in this study vary between the projects reviewed: some are using fully automated metadata strategies; while other projects are using fully manual metadata strategies; and another group are using both automatic and manual metadata strategies.

Mr Arofan Gregory (2011), Executive Manager, of the *Open Data Foundation*, has been providing expert advice to the Australian Bureau of Statistics in Australia and to the Open Data projects in the UK and Europe during 2011-12. He outlines in a recent paper how collaborations between open government initiatives, the Linked Open Data and RDF communities, and experts in the field of statistics and research data management could improve the usability of government data on the Web. He argues that two good standard metadata models exist that may help open data initiatives achieve their aspirations:

- the Statistical Data and Metadata Exchange (SDMX) model, which tends to be used by official statistical organisations to mark up their aggregated data in XML; and
- the Data Documentation Initiative (DDI), which tends to be used as a metadata model for describing survey data and other types of microdata.

Both of these standards are being adopted across the world and are enabling the creation of generic tools which can work with any standardised data. Perennial issues around metadata schemes however remain, and in particular the ability to agree on a common naming and keyword system. If



a common agreement could be achieved, comparisons between datasets could become even more useful. Examples of different types of metadata approaches prepared in this UK can be found at: <http://obd.jisc.ac.uk/examples>.

## Project maturity

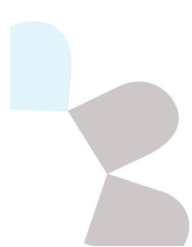
Project maturity can be judged along a spectrum, with new projects harvesting and releasing data that already exists. More mature projects such as those reviewed overseas, are working on enabling the personalisation of open data to meet the demands of users.

Most of the Australian projects reviewed are in their early stages of generating open data – apart from those who have been using open metadata formats for many years. Governments in Australia are largely in their initial phases and are publishing existing datasets. But the more mature open data projects, such as those being undertaken by NASA and the US Government, are aiming to create open datasets. Some large scale projects such as WikiData are in very early stages of development but, with backing from Google, are poised to have significant impact in promoting and enabling the sharing of open data (<http://meta.wikimedia.org/wiki/Wikidata>).

## Issues identified during research

The following comments are observations based upon communications with stakeholders:

1. Academics, while overwhelming in their support for the benefits of open access and free software, sometimes do not want to share their data because of 'intellectual property' claims.
2. Services such as ANDS (Australian National Data Service) have currently put in place infrastructure that may serve as a useful foundation for open data and greater public access to publicly-funded research. However, most entries in the ANDS registry to date require further investigation to track down how to access the resources listed or described.
3. While the Wikiversity pages have been useful for collaborative documentation they do not lend themselves easily toward stimulating discussion.
4. There appears to be a conflation by some that the 'open data' agenda is being led by the Open Government movement (see for example [www.shareable.net/blog/open-data-emerges-as-a-global-movement](http://www.shareable.net/blog/open-data-emerges-as-a-global-movement)). There are however, other facets to openness associated with the Internet (eg architecture, source code, standards, educational resources, access). As such, there are many synergies available for 'connecting the dots' between various open data, open government, and openly interoperable initiatives.
5. In September 2012, Wikileaks was entered into the survey collecting information about open data projects. Receiving this response poses a number of interesting questions around the similarities and differences between the motivations, methods and principles of a project like Wikileaks, and the seemingly wide growth of open government data projects worldwide, and the Open Government or 'gov2.0' movement.
6. There was a poor response rate to the open survey (9 responses), despite the survey going out to major forums, twitter channels, to Offices of Information Ministers known to be



engaging in open data initiatives, and direct telephone calls with offers to enter information to the survey on behalf of projects. This poor response rate raises a number of old and new issues, such as

- whether open surveys are an effective instrument and method;
- whether the design and implementation of this particular survey was suitable (albeit, it was reviewed by several critical friends);
- whether the reach went far enough into the open data networks (suggested by the majority of the entries being Australian); and/or
- whether open government providing access to open data fosters democracy.

## Summary

There are a wide range of storage and publishing formats being made available to enable open data publishing, access and management. There is an increasing demand for access to inter-related services and linked datasets that answer specific individual queries. Furthermore, there is a 'thirst' for the datasets to be presented in attractive graphical interfaces. The links between open government and open data are being seen by policy makers and researchers as a cost effective way to foster research across disciplines and fields in ways previously impossible. The concepts and inter-relationships between open data and open government are still in their infancy, yet it can be seen that there are many open data projects being undertaken worldwide, and these projects are growing daily.

## Open data initiatives reviewed in Australia and overseas

The *World Map of Open Government Data Initiatives*, a semantic web which is part of the *Austrian Linked Data* portal (last updated December 2011), provides an overview of several open data projects or pilots being conducted in the US, UK, Europe, Australia, New Zealand, Canada, Kenya and a small number of Latin American countries. Most open data projects are either being undertaken by governments or are government funded.

There are 40 open data projects being conducted by local or regional governmental authorities in 14 countries. In addition, there are 15 national government open data projects; 9 nationwide private initiatives in 7 countries; and 4 international or transnational open data projects. In the UK, Germany, USA there are also local and regional private initiatives being undertaken. This study has viewed a selection of these projects and some projects initiated in 2012. Those projects reviewed as part of this study are indicated with an \*. A summary of the countries and projects is provided at Appendix Two.



## Australian open data initiatives

In Australia, the Australian Government and several state and territory governments are implementing open data initiatives. The following information summarises these initiatives.

### Data.gov.au

The Australian Government is implementing open data initiatives through the auspices of the Department of Finance and Deregulation. The portal Data.gov.au has been established and provides downloadable datasets cases and links to other data catalogues or sources. This initiative grew out of work by the *Government 2.0 Taskforce* and implements the *Declaration of Open Government* announced by the Australian Government in July 2010:

*“The Australian Government now declares that, in order to promote greater participation in Australia’s democracy, it is committed to open government based on a culture of engagement, built on better access to and use of government held information, and sustained by the innovative use of technology”* (Australian Government Department of Finance and Deregulation, 2010).

The datasets used on Data.gov.au are drawn from several different government agencies. The dataset provided on Data.gov.au are presented using one or more of the following file formats:

- CSV/XLS
- TXT
- XML
- RDF
- KML/KMZ
- Shapefiles
- Catalogue

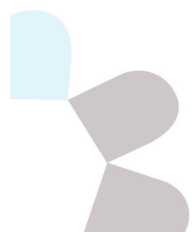
The respective datasets have been tagged and sorted into categories, with many of the datasets provided in a compressed (zipped) format and have to be uncompressed (unzipped), for use (see <http://data.gov.au/about/>). The Department of Finance and Deregulation is responsible for the development and ongoing operation of the Data.gov.au portal.

### \*Linked Data for Open Government

<http://lab.environment.data.gov.au>

This project aims to develop an Australian demonstrator of the technological capability of the data integration approach known as linked open data (LOD), developed and promoted by the W3C standards body as part of the semantic web platform. It commenced in 2012. The project publishes a flagship data product of the Bureau of Meteorology, ACORN-SAT (Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT), in the LOD style to promote uptake amongst the LOD user community. Strategically, the project aims to attract interest in LOD for open government across the Australian Government sectors. It is copyright is the Public Domain.

ACORN-SAT is an open data set, recording historical temperature observations in Australia. The dataset employs analysis techniques that provide a daily temperature record over the last 100 years. The project delivered a Linked Data API SPARQL for the ACORN-SAT climate data set available openly



on the Web. The project also developed a number of ontologies that are openly accessible on the Web as well. Linked Data is at: <http://lab.environment.data.gov.au>. The UK initiative <http://environment.data.gov.uk/id/bathing-water> inspired this project.

The project is using RDF(s) and OWL file formats, and the metadata is generated automatically. The data is generated using automatic mapping via D2RQ from a relational data model to RDF.

#### **\*Research Data Australia - an ANDS Service**

<http://researchdata.ands.org.au/>

*Research Data Australia* is a discovery service for Australian research and researchers, enabling broader access to publicly funded research. It is a mesh of searchable web pages describing, and where possible linking to Australian research data collections. As such, it is not yet a service that publishes open datasets – just information linking to them. The data connections strategy of *Research Data Australia* incorporates common referencing methods for researchers, research groups, research activities, places, research datasets, research fields, and scholarly or scientific terminology. *Research Data Australia* is a project by the Australian National Data Service (ANDS) in partnership with research institutions and data producing agencies. ANDS commenced in 2008.

ANDS supports public access to as much publicly funded research data as can be provided within the constraints of privacy, copyright and technology limitations. ANDS leads the creation of a cohesive national collection of research resources aimed at developing a rich data environment that is intended to:

- Support the better use of Australia's research outputs;
- Enable Australian researchers to publish, discover, access and use data; and
- Enable new and more efficient research.

ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy Program and the Education Investment Fund (EIF) Super Science Initiative. The underlying model for the ANDS Registry is the international standard ISO 2146:2010 *Information and documentation – Registry services for libraries and related organizations*. ANDS is a federated registry service that contains descriptive and administrative metadata: for collections; for related services, parties and activities; and for identifying the relationships between them. ANDS provides access to data for use by the public and researchers, with copyright being Public Domain and Creative Commons Attribution.

ANDS uses a wide range of file formats and protocols. Technical documents are available at <http://ands.org.au/resource/techguides.html>. These files include documentation for the metadata protocol RIF-CS OAI-PMH Data Provider at <http://ands.org.au/resource/rifcs-provider.html>.

#### **Australian state government initiatives**

##### **\* Australian Capital Territory government**

<https://data.act.gov.au/>

dataACT was launched on 30th August 2012. It is designed as an open data service initiative to enable individuals, businesses and people in the ACT public sector to access, explore and build upon existing government datasets. It currently has about 45 datasets.



The beta dataACT is built on Socrata platform and makes data available in a machine readable and in a consistent range of data formats. This functionality enables static datasets to be downloaded in readily usable formats (e.g., JSON). The platform includes some basic data visualisation tools to make graphs and maps available. These can be embedded into external websites, enabling customised views of datasets to be developed. The platform has been designed to enable data to be uploaded from across the ACT Public Service through data integration systems, so that the data can be kept up to date.

Metadata tagging is undertaken both manually and automatically. Common open data files used in the ACT include: CSV, JSON, PDF, RDF, RSS, XLS, XLSX and XML.

#### **\* New South Wales**

<http://data.nsw.gov.au/>

data.nsw was launched in 2009. It is a searchable portable that includes the following categories: publications, spatial data, data feeds, apps, and apps4nsw (archived).

The file formats used by data.nsw are as follows:

- CSV
- PDF
- XLS
- HTML
- XML
- EPS
- EMF
- Shape
- ArcInfo
- SEGY
- CGM
- Geodatabase
- MapInfo
- API
- JSON
- Zip
- JPEG

#### **\* Victoria**

[data.vic.gov.au](http://data.vic.gov.au)

data.vic.gov.au is a new portal administered by the Department of Business and Innovation that assists users to find datasets that are housed on other government websites. This site was launched in 2012, following on from other public online sites hosted by the State of Victoria. data.vic.gov.au employs a Creative Commons Attribution copyright license.

#### **South Australia**

##### **\*AusGOAL - DATA.SA**

<http://www.ausgoal.gov.au/>

The South Australian Government is a partner in the AUSGOAL initiative. AusGOAL is the Australian Governments Open Access and Licensing Framework. It provides advice and support to state governments and related sectors to facilitate open access to publicly funded information. The AusGOAL Framework uses an open licensing framework.

The South Australian project has recently commenced in 2012, and is an initial step toward creating a 'DATA.SA' site. The site is currently a 'beta' version. Experiences from delivering this site will inform future directions in South Australia. The South Australian site presents a catalogue of raw





data released by SA Government and available to all (i.e., researchers, the business community, public) for re-use. Data is not hosted on this site. The site contains metadata that describes the data and links to where the data is currently hosted on agency websites. It is the intent of the South Australian Government to use formats that are as open as possible, and to use licence datasets under AusGOAL ([www.ausgoal.gov.au](http://www.ausgoal.gov.au)), which are primarily Creative Commons licences. Indeed, the copyright for Data SA is a Creative Commons Attribution licence. Metadata is handled manually.

The purpose of the DATA SA site is to provide as much information as possible under open licences that is freely available to use, re-use and distribute. The South Australian Government is keen to promote information in more open formats. The project forms a focal point for drawing attention to data held by South Australian Government that has been released publicly. It is an initiative to encourage open data champions in South Australian Government agencies to continue and to accelerate their efforts to advocate for and to release data held by SA Government.

The common files used include links to raw datasets, APIs, and RSS feeds to other jurisdictions' sites. In addition, functionality is being added to the code of the originally built site.

### **Northern Territory**

***\*Centre for School Leadership, Learning and Development Open Data Project***

[http://en.wikibooks.org/wiki/Open\\_Data](http://en.wikibooks.org/wiki/Open_Data)

This small 2012 study has been conducted as a means to collect data about open data projects. That is, the project design has modelled the topic of research in order to gather information on open data projects, determine what they are, who is running them, how they are publishing open data, why they are doing it, from where they emerged, and for how long have they been operating. Information about the project is presented on Wikiversity using a Creative Commons Attribution licence. The file formats being used are CSV, ODS, HTML, TXT and XLS.

### **Open data initiatives overseas**

The US and UK Governments and several European Governments are undertaking open data initiatives. The main European site is <http://www.publicdata.eu/>. Probably the most comprehensive public data initiative next to the UK and the US is the French public data portal: <http://www.data.gouv.fr/>. A brief summary of the UK, US and European government-funded open data initiatives is outlined below.

To support governments to create open government platforms, in 2012, the General Services Administration (GSA) of the US Government released the open source code of their Open Government Platform (OGPL). This software was developed by the US Data.gov in collaboration with the National Informatics Centre in India. This software is now available to any government to evaluate and adopt. The OGPL is based on Drupal, with the core software including a data management system, web site, and social networking support functionality. It is available from: <http://www.opengovplatform.org/>



**United States of America*****Data.gov***

Data.gov is an official website of the US Government which operates an open data portal to provide a one-stop-shop for the access and re-use of data. This site is one of the most mature open data government sites, globally. It is the place from which the US Government releases open source software. In September 2012 there were

- 378,529 raw and geospatial datasets
- 1,264 government apps
- 236 citizen-developed apps
- 103 mobile apps
- 172 agencies and subagencies

***NASA – United States of America***

<http://data.nasa.gov/>

data.nasa.gov portal is an open data project that forms part of the NASA Open Government Initiative. It provides direct links to more the 500 datasets. The portal uses an API which was created from the JSON-API Wordpress Plugin.

**United Kingdom**

***data.gov.uk*** - <http://www.data.gov.uk> [beta site]

data.gov.uk brings together public data into one searchable portal. Currently there are 8655 datasets pulled together from central UK government departments, other public sector bodies and local authorities. Each UK government department has published its own Open Data strategy which sets out their release plans of new data that will be published over the next two years from 2012-14.

***\*JISC project – PROD***

<http://prod.cetis.ac.uk/>

PROD is a database of IT development projects sponsored by JISC. The project commenced in 2009. The basic project data have been augmented by as much information as possible about the technologies and standards the individual projects use. The website <http://kasabi.com/dataset/jisc-cetis-project-directory> is generated by open data. There is an institutional commitment to the use of open technologies, driven by a desire to help build an open world wide web of data. There is also a demand for augmenting JISC datasets with those of others. The site is covered by a Creative Commons Attribution licence. The most common file formats used are HTML, RSS and RDF, and five star linked data (SPARQL).

***\*OpenGeoscience***

<http://www.bgs.ac.uk/opengeoscience>

OpenGeoscience is a web service that commenced in 2009, to publicly provide a range of geological information including maps, photos and digital data that can be combined with other environmental



information to help enable people to understand the world around them. One of the purposes of OpenGeoscience is to respond to increasing UK and European Union demands for open access to public sector information to complement existing business delivery channels. OpenGeoscience provides materials free for non-commercial private study, research and educational activities. The business user community can also use the materials for innovation towards new commercial products and services.

The project is led by the Natural Environment Research Council in the UK. Data is generated from records held by the British Geological Survey (BGS), which is the UK's National Centre for earth science information and the foremost supplier of geoscience solutions and impartial advice to the governments of the UK and developing nations. The BGS maintains a vast store of data and information that has been gathered since the inception of its ancestor, the Geological Survey of Great Britain, in 1835. Various datasets are obtained under statutory requirements such as the Mining Industry Act 1926 and the Water Resources Act 1991. BGS acquires and maintains up-to-date geoscientific knowledge of the UK and its continental shelf, by means of systematic geological, geophysical, geochemical, hydrogeological and geotechnical surveys, collecting and using high quality data.

OpenGeoscience includes a WMS service and web-based interactive viewer for its 1:50,000 scale digital geology data for the whole of Great Britain – the first time in the world that national coverage, attributed, street-level scale geology data has been made available via the internet. OpenGeoscience received 20 million hits on its day of release. OpenGeoscience web services have since been used to deliver the data via the iGeology smartphone app for both iPhone and Android phones. iGeology has had over 100,000 downloads. Other resources made available include scans of over 1 million borehole records and 50,000 detailed geological photographs.

Websites that are served by Open Geoscience include the following:

- <http://www.bgs.ac.uk/opengeoscience/>
- <http://www.bgs.ac.uk/igeology>
- <http://www.bgs.ac.uk/data/services/mash-ups/home.html>

Open Geoscience has its own Terms and Conditions for use. The file formats used include WMS, KML, GeoRSS, Linked Open Data, WFS, jpg, GIS (shapefiles and MIF), png, TIFF, excel, pdf. Metadata is allocated both automatically and manually.

## European Union

### ***\*Public Data EU - Europe's public data***

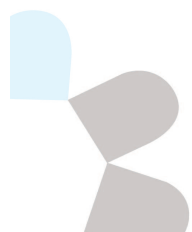
<http://www.publicdata.eu>

This project aims link and combine large datasets from a wide range of different sources to increase the transparency of governments and to improve public service delivery. Public Data EU is developed by the Open Knowledge Foundation. PublicData.eu is powered by CKAN, which is a large, open source data catalogue system. The file formats used include CV, TXT and LOD.

## Italy

### ***\*Italian National Research Council: data.cnr.it***

[data.cnr.it](http://data.cnr.it) is an initiative of the Italian National Research Council (CRN) aimed at providing public access to the information collected by this CNR organisation. It commenced in 2010.



The technologies of the data.cnr.it are based on RDF and follow the paradigm of the Linked Open Data initiative. The data is publicly available and query-able. The data is generated through the receipt of ad-hoc queries which are then run on multiple databases; the data is then transformed into RDF and hosted on a Quad-store. Queries can be extracted as spreadsheets, or as XML. This project is working towards a SPARQL endpoint: see <http://data.cnr.it/sparql/>. It is covered by a 'Creative Commons Attribution Non Commercial, Creative Commons Attribution No Derivatives' licence. The metadata is generated automatically.

## New Zealand

### \*data.govt.nz

data.govt.nz is an initiative of the New Zealand Government Information Services, Department of Internal Affairs, which commenced in 2011. It coincides with the release of the *Declaration on Open and Transparent Government*, which was approved by Cabinet on 8 August 2011. This site uses a range of copyright licenses including the following:

- All Rights Reserved;
- Creative Commons Attribution;
- Creative Commons Attribution Share Alike; and
- Creative Commons Attribution Non Commercial.

This project uses a combination of manually and automated metadata strategies. The main file formats used include CSV, KML/SHP, and PDF.

## South Africa

<http://codingfordemocracy.org>

A very new initiative is the Open Data & Democracy Initiative (ODADI) which was launched with an event called *Coding for Democracy*, held on 3-5 August 2012 in Cape Town, South Africa. ODADI is a community of citizens, activists, technologists, journalists, and entrepreneurs who are dedicated to developing and applying practical open technologies and promoting open data as a means to improve the efficiency of government, increase transparency, improve service delivery and the empower South Africans.

## International projects

### \*WikiLeaks

<http://wikileaks.org/>

WikiLeaks commenced in 2006, and is an international, online, not-for-profit organisation publishing submissions of private, secret, and classified media from anonymous news sources, news leaks, and whistleblowers. WikiLeaks combines security technologies with journalism and ethical principles to make available data that is submitted by anonymous people. WikiLeaks investigates the materials and the source, writes a news piece around the leaked information, and provides access to the leaked information to support the news story. This practice of publishing the data that informs their news writing makes WikiLeaks an open data project. WikiLeaks uses a range of common open data files to support audio, video, graphics and text. Metadata is allocated both automatically and manually.



**\*World Bank**

<http://data.worldbank.org/>

The World Bank has created an extensive open data portal using the open source Drupal and currently uses the following three APIs to provide access to different datasets: s

- Indicators (or time series data) using API, XML and JSON;
- Projects (or data on the World Bank's operations) using API, Atom; and
- World Bank financial data (World Bank Finances API) using API, XML, JSON and RDF.

**\*WikiData**

<http://meta.wikimedia.org/wiki/Wikidata>

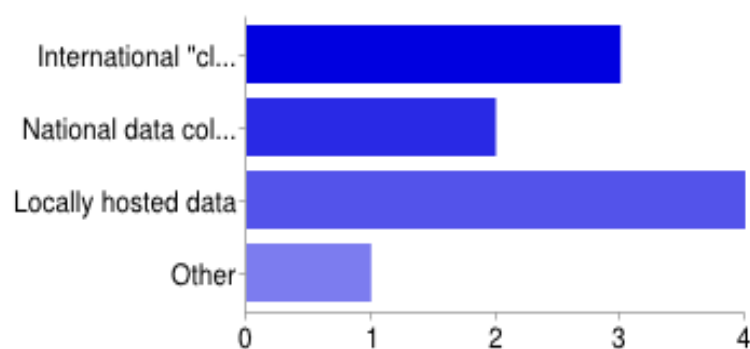
The goal of WikiData project is to create a common data repository, principally for WikiMedia projects. This will make it possible to collect and curate data to be used within all WikiMedia projects and by third parties.

## Graphical presentation of findings from online the survey

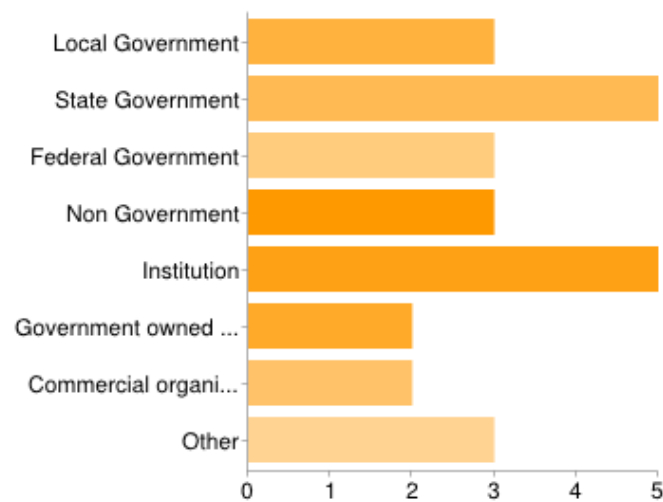
The following graphs provide the collated responses to each of the online survey questions. The full detailed responses informing these graphs can be found at:

[http://en.wikiversity.org/wiki/Open\\_Data](http://en.wikiversity.org/wiki/Open_Data).

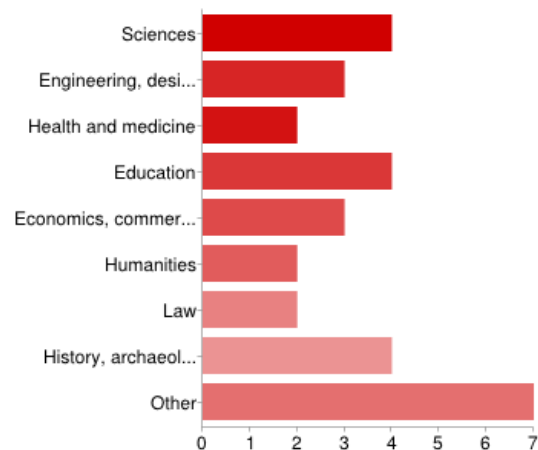
**Figure 1: Host server owner/manager**



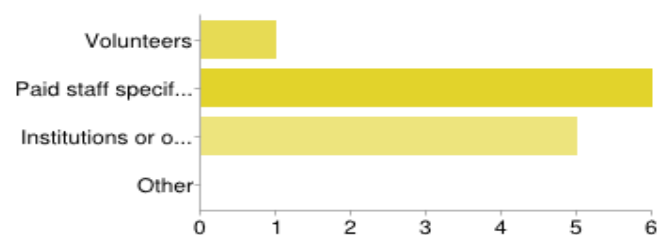
This graph shows that of the nine responses, the host servers of the projects reported are located locally, nationally and internationally, depending on the project reported.

**Figure 2: Generating source of the data**

This graph shows that the agencies responding to the survey show that their sources for generating open data are varied across government and non-government agencies and range from local to Federal levels in scale.

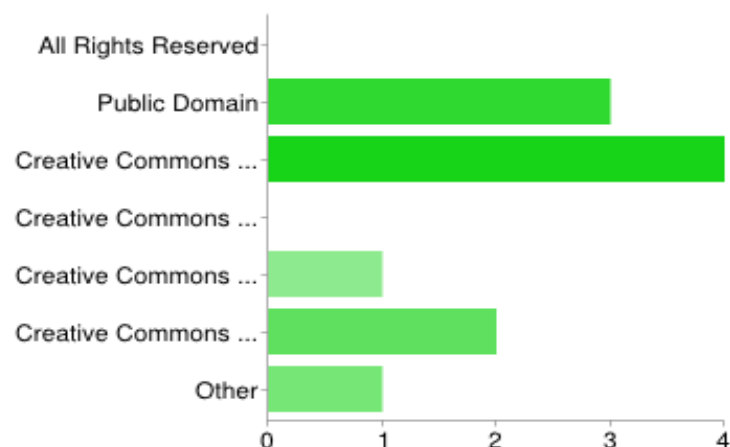
**Figure 3: Subject areas of the open data**

This graph shows that the agencies that responded to the online survey covered a broad range of discipline areas, depending on the agency.

**Figure 4: How the open data project is managed**

The graph at Figure 4 shows that the majority of the open data projects reported in the survey are maintained by paid professionals rather than relying on volunteers.

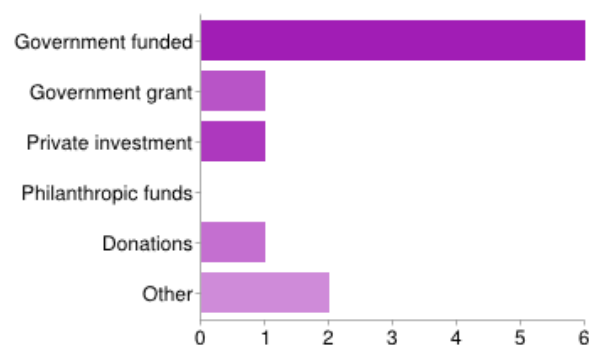
**Figure 5: What copyright is being used for the data?**



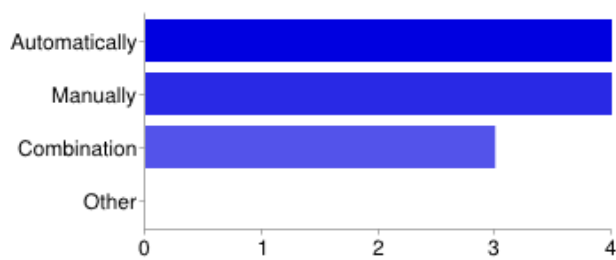
All Rights Reserved	0	0%
Public Domain	3	33%
Creative Commons Attribution	4	44%
Creative Commons Attribution Share Alike	0	0%
Creative Commons Attribution Non Commercial	1	11%
Creative Commons Attribution No Derivatives	2	22%
Other	1	11%

This data suggests that open data projects thoughtfully selected and use a range of license arrangements, depending on the project, albeit that most are based on the Creative Commons parent license.

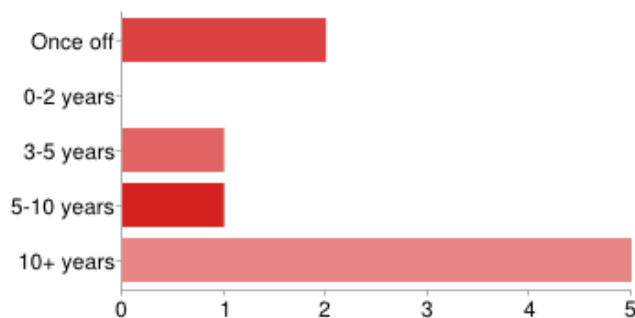
**Figure 6: How is this project funded?**



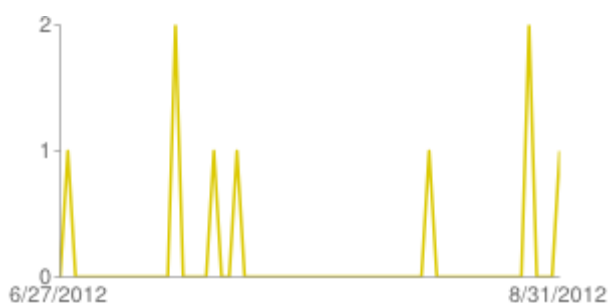
This figure indicates that most of the open data projects reported through the online survey were government funded projects.

**Figure 7: How is metadata generated?**

This graph suggests that the open data projects reported upon through the online survey used a mixture of automatic and manual approaches to allocating the metadata.

**Figure 8: What is the expected duration of the project?**

This graph suggests that while some of the open data projects are 'one off', there are plans for several open data projects to have extended project lives, of up to a decade and more.

**Figure 9: Survey response Timings**

It can be seen by this last figure, that there were spikes in the submission of data through the online survey. These spikes tended to be associated with then the researchers actively encouraged agencies to complete the survey.



## Appendix One: Online survey questions

The following fields of information were sought through the use of the online survey:

- Name of the project?
- Domain owner/manager?
- Project URL?
- Host server owner/manager?
- Contact details?
- Subject areas of the open data?
- Overview of the project
- How the project is managed?
- What services are being generated from the open data?
- What copyright is being used for the data?
- Why is this an open data project?
- What initiatives or activities have inspired this project?
- How is this project funded?
- How is the data generated?
- List the data formats that are used?
- How is metadata generated?
- When did the project start?
- Is the project ongoing?
- Under what legal jurisdiction is this project being conducted, if any?
- How is the data served?
- Date published?
- Generating source of the data?
- How can the success of the project be known?
- Subject specific keywords
- How the project is managed?
- What is the expected duration of the project
- Feedback on this survey



## Appendix Two: Key organisations undertaking open data projects

The following lists provide an overview of the key organisations globally that are undertaking open data projects. This list has been derived from *The World Map of Open Government Data Initiatives*, a semantic web, which is part of the *Austrian Linked Data* portal (last updated December 2011).

### Local or regional governmental authorities

Country	Initiative	Agency
Australia	data.nsw	NSW Government Chief Information Office
Australia	data.vic.gov.au	Department of Innovation, Industry and Regional Development
Austria	data.wien.gv.at	Vienna City Administration
Austria	data.linz.gv.at	Open Commons Linz Region
Brazil	Governo Aberto SP	Governo do Estado de São Paulo
Canada	Toronto Open	City of Toronto
Canada	Open Data Edmonton	City of Edmonton
Canada	Public Data Catalogue	City Of Calgary
Canada	Open Data Feeds	City of Nanaimo
Canada	dataott.org	City of Ottawa
Canada	Vancouver Open Data	City of Vancouver
Canada	Open Data Catalogue	City Government of Windsor, Ontario
Finland	Helsinki Region Infoshare	Consortium
Finland	Infoshare	City of Helsinki Urban Facts
France	ParisData	Ville de Paris
France	data.nantes.fr	City of Nantes
France	in constitution	City of Paris Department for Innovation and Science
France	data.rennes-metropole	Direction générale de l'information et de la Communication de Rennes Métropole
France	Open Geo Data	City of Brest
Germany	Offene Daten Berlin	Senatsverwaltung für Wirtschaft, Technologie und Frauen



Country	Initiative	Agency
Ireland	Fingal Open Data	Fingal County Council
Italy	dati.piemonte.it	Piemonte Regional Government
Northern Ireland	opendatani.info	UK Department of Finance & Personnel
Spain	Conjunto de Datos	City of Zaraoza
Spain	Datos de Asturias	Government of the Principality of Asturias
Spain	Open Data Euskadi	Administración General de la Comunidad Autónoma de Euskadi
Uruguay	Datos abiertos	Montevideo City Council
UK	London Datastore	Greater London Authority (GLA)
UK	Lichfield Open Data	Lichfield District Council
UK	Warwickshire Open Data	Warwickshire County Council
USA	CivicApps	Portland Development Commission
USA	data.seattle.gov	City of Seattle
USA	DataSF	City & County of San Francisco
USA	NYC BigApps	NYC Economic Development Corporation
USA	City of Chicago's Data Portal	City of Chicago
USA	Rhode Island Data	State of Rhode Island

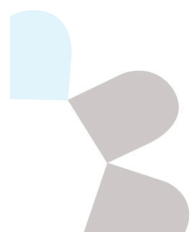
### Regional private initiatives

Country	Initiative	Agency
Germany	OpenBerlin	Open Knowledge Foundation Deutschland Open Data Network e.V. Liquide Democracy e.V. Open Everything
UK	Open Data Manchester	Open Data Manchester
USA	Open Colorado	Colorado Smart Communities



**Nationwide government authorities**

<b>Country</b>	<b>Initiative</b>	<b>Agency</b>
Australia	<a href="http://data.australia.gov.au">data.australia.gov.au</a>	Department of Finance and Deregulation
Canada	Open Data Pilot Project	Treasury Board of Canada
Chile	Datos públicos del Gobierno de Chile	Ministerio Secretaría General de la Presidencia
Denmark	Digitaliser.dk	The Ministry of Science - National IT and Telecom Agency
Finland	Apps4Finland	Suomen Verkkodemokratiaseura
Greece	<a href="http://geodata.gov.gr">geodata.gov.gr</a>	Office of the Prime Minister
Kenya	<a href="http://opendata.go.ke">opendata.go.ke</a>	Kenya ICT Board under the Ministry of Information and Communication
New Zealand	<a href="http://data.govt.nz">data.govt.nz</a>	Department of Internal Affairs
Moldova	<a href="http://data.gov.md">data.gov.md</a>	Republica Moldova
Norway	<a href="http://data.norge.no">data.norge.no</a>	Ministry of Government Administration, Reform and Church Affairs
Spain	Aporta Project	Spanish Ministry of Industry, Tourism and Commerce
Sweden	<a href="http://opengov.se">opengov.se</a>	Swedish Government
The Netherlands	<a href="https://data.overheid.nl/">https://data.overheid.nl/</a>	The Netherlands Government
United Kingdom	<a href="http://data.gov.uk">data.gov.uk</a>	HM Government
United States	<a href="http://data.gov">data.gov</a>	US Government

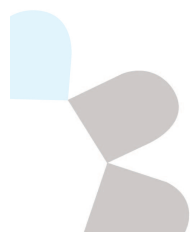


**Nationwide private initiatives**

<b>Country</b>	<b>Initiative</b>	<b>Agency</b>
Albania	open.data.al	Open Society Foundation
Belgium	OpenBelgium.be	Krimson - Drupal architects
Italy	DataGov.it	Associazione Italiana per l'Open Government
Germany	apps4democracy.de,	Opendata Network e.V. Government 2.0 Netzwerk Deutschland i.G. echo source e.V.
Slovakia	Datanest	Aliancie Fair-Play
Russia	opengovdata.ru	NPO
The Netherlands	Open Data Overheid	Lex Slaghuis

**Transnational catalogue**

<b>International catalogue</b>	<b>Initiative</b>	<b>Agency</b>
Energy Efficiency Catalogue	data.reegle.info	Renewable Energy and Energy Efficiency Partnership
OECD	OECD iLibrary	OECD
United Nations	UNdata API	UNdata
WikiData	WikiData	Wikimedia Foundation
World Bank	World Bank's Open Data initiative	World Bank



## Appendix Three: Models, Specifications & Protocols, Licenses

Name	Location
British Library Data Model	<a href="http://talis-systems.com/2011/07/british-library-data-model-overview/">http://talis-systems.com/2011/07/british-library-data-model-overview/</a>
City of Vancouver Open Data Formats	<a href="http://data.vancouver.ca/dataFormats.htm">http://data.vancouver.ca/dataFormats.htm</a>
Common Data Format	<a href="http://cdf.gsfc.nasa.gov/">http://cdf.gsfc.nasa.gov/</a> Led by NASA
GOM Inspection Exchange Format	<a href="http://www.gom.com/3d-software/open-data-formats.html">http://www.gom.com/3d-software/open-data-formats.html</a> 3D datasets
OData	<a href="http://www.odata.org/">http://www.odata.org/</a> An initiative of OASIS
Open Archives Initiative	<a href="http://www.openarchives.org/OAI/openarchivesprotocol.html">http://www.openarchives.org/OAI/openarchivesprotocol.html</a>
Open Conceptual Data Model	<a href="http://virtuoso.openlinksw.com/whitepapers/open%20conceptual%20data%20models.html">http://virtuoso.openlinksw.com/whitepapers/open%20conceptual%20data%20models.html</a>
Open Data Commons Licenses	<a href="http://opendatacommons.org/licenses/">http://opendatacommons.org/licenses/</a>
Open Data Format	<a href="http://ewh.ieee.org/cmte/psace/CAMS_taskforce/format.htm">http://ewh.ieee.org/cmte/psace/CAMS_taskforce/format.htm</a>
Open Data Model	<a href="http://www.opendatamodel.com/">http://www.opendatamodel.com/</a> Focused on database design
OpenMAMA Project	<a href="http://www.openmama.org/">http://www.openmama.org/</a> Middleware Agnostic Messaging API (MAMA)
Microformats	<a href="http://microformats.org/">http://microformats.org/</a>
XTF eXtensible Text Format	<a href="http://xtf.cdlib.org/">http://xtf.cdlib.org/</a>



## References

- Australian Government, Department of Finance and Deregulation, (2010). *Declaration of Open Government*, downloaded on 9 September 2012, from <http://www.finance.gov.au/e-government/strategy-and-governance/gov2/declaration-of-open-government.html>
- Austrian Government (2011). *World Map of Open Government Data Initiatives*, downloaded on 9 September 2012, from <http://gov.opendata.at/site/node/38>
- Cornish, L. (2012). *Data.gov.au and open data in Australia – where are we at?* News Queensland, 29 February 2012, downloaded on 9 September 2012, from <http://www.geonext.com.au/wp-content/uploads/2012/03/LISA-CORNISH.pdf>
- Gregory, A. (2011). *Open Data and Metadata Standards: Should We Be Satisfied with “Good Enough”?*, Open Data Foundation, downloaded on 30 September 2012, from <http://odaf.org/papers/Open%20Data%20and%20Metadata%20Standards.pdf>
- Institute of Metadata Management (2011). *Who we are*, Institute of Metadata Management, downloaded on 10 October 2012, from <http://www.metalounge.org/aboutIMM-who-we-are>
- Minister of State for the Cabinet Office and Paymaster General (2012). *Open data white paper. Unleashing the potential*. HM Government, downloaded on 9 September 2012, from [http://www.data.gov.uk/sites/default/files/Open\\_data\\_White\\_Paper.pdf](http://www.data.gov.uk/sites/default/files/Open_data_White_Paper.pdf)
- Office of the Prime Minister (UK) (2011). *Letter to cabinet ministers on transparency and open data*, downloaded on 9 September 2012, from <http://www.number10.gov.uk/news/letter-to-cabinet-ministers-on-transparency-and-open-data/>
- University of North Texas (2012). *The Denton Declaration, An Open Data Manifesto*, University of North Texas, downloaded on 30 November 2012 from [http://openaccess.unt.edu/denton\\_declaration](http://openaccess.unt.edu/denton_declaration)
- Wikibooks, (2012), *Open Metadata, Open data handbook*, Wikibooks, downloaded on 30 September 2012, from [http://en.wikibooks.org/wiki/Open\\_Metadata\\_Handbook/Open\\_Metadata](http://en.wikibooks.org/wiki/Open_Metadata_Handbook/Open_Metadata)

