Welsh Government

Open Data Plan

Front Cover Image: © Edhar Yralaits | Dreamstime.com

Any enquiries regarding this publication should be sent to us OfficeoftheChiefDigitalOfficer@wales.gsi.gov.uk

Ministerial Foreword

Data is a valuable resource that has the power to transform the services we provide and the decisions we make. None more so than open data which is free to access, free to use and can be shared by anyone.

Opening up access to our data, and encouraging other public bodies to do so, will not only improve transparency and accountability but also drive innovation. By changing the way we work with data and embracing these changes we have an opportunity to identify and predict trends which will enable us to improve the services we provide.

Open data also enables citizens to put data to use. It will help people hold us all to account more effectively and drive improvement. The perspectives of citizens, of communities and of businesses offer a completely different lens to look at data. By publishing more open data we can reap the benefits of this new insight and the innovative solutions they identify.

Within Welsh Government we are already striving to increase the accessibility to our data through websites such as Lle and StatsWales. I want us to build on this excellent work and fully realise the opportunities that open data can provide.

I believe this Open Data Plan will, through raising awareness and consolidating ongoing work, demonstrate Welsh Government's commitment to open data. I also think it provides a practical opportunity to work with and encourage public service organisations to increase their publication and use of open data.

Leighton Andrews AM
Minister for Public Services

Contents

1. Introduction

Within Wales the public service generates and publishes a vast amount of non-personal data. If this data was published openly and shared freely this could provide numerous opportunities and benefits for the people of Wales, businesses, the public service sector and the Welsh Government.

The Welsh Government is committed to publishing and sharing data that is meaningful, accessible and re-usable. This open data plan not only outlines Welsh Government's ongoing commitments in this area, but aims to increase awareness of open data across Welsh Government. It also hopes to prompt thinking about what we in Welsh Government can be doing to help realise the benefits and opportunities related to open data.

2. What is Open Data?

Open data is defined as data that is made freely available in clear understandable formats that can be used, re-used and distributed by anyone.

The main reason for opening up data, besides improving Government openness and transparency, is to help inform and improve services for the public. For example open data can be used to highlight what government money is being spent on, how schools or hospitals are performing, identify problem areas or it can be used to create online services for people.

Whilst open data is not a new concept, in recent years there has been a growing interest in open data both internationally and locally. However, not all data can be made open. There are legitimate reasons for not opening up data, such as the data being personal, the risk of information being misused or the data presenting a security risk.

But, it's not as straight forward as data being either open or closed. In some cases data may be shared with others but only in a certain format or by a specific agreement. This is demonstrated by the Open Data Institute (ODI) data spectrum which shows how data can range from being closed to shared through to being openly available.

© Open Data Institute (ODI) - (https://theodi.org/data-spectrum)
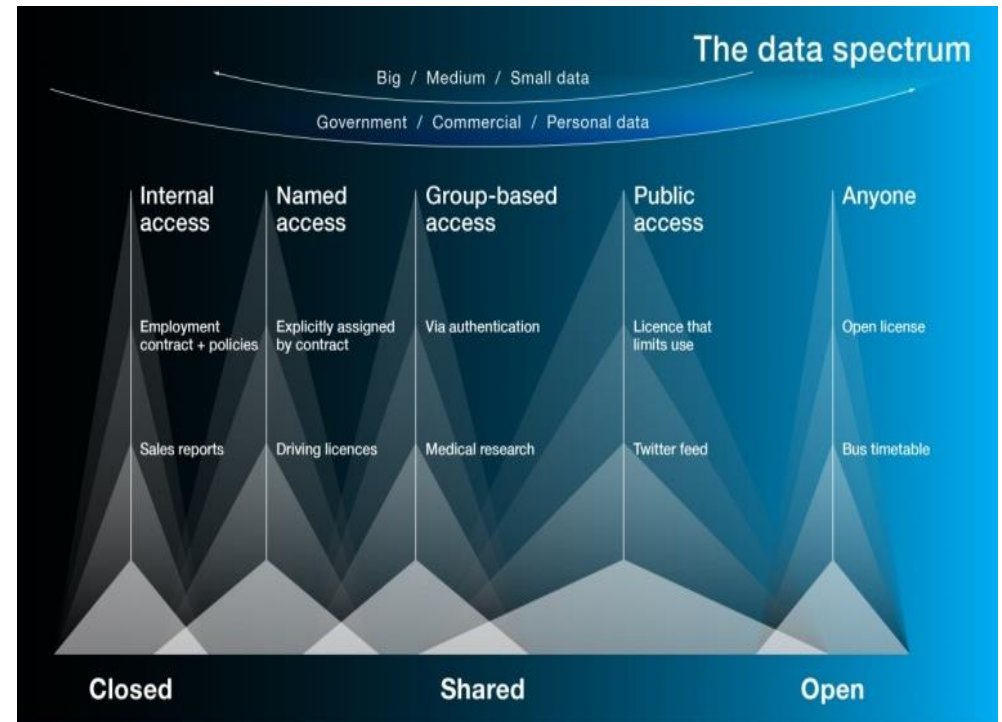
3. Committed to Open Data

There are a number of ongoing commitments and legal obligations that support the opening up of data, both locally and internationally.

Global

An Open Data Charter was agreed at the G8 Summit in June 2013. The G8 agreed that "*open data are an untapped resource with huge potential to encourage the building of stronger, more interconnected societies that better meet the needs of our citizens and allow innovation and prosperity to flourish*". Members agreed to the following five principles to be the foundation for access to and the release & re-use of data made available by G8 governments:

- Open data by default
- Quality and quantity
- Useable by all
- Releasing data for improved governance
- Releasing data for innovation

European

The EU INSPIRE Directive which we need to implement by 2019 aims to create an infrastructure that will enable the sharing of environmental spatial information among public sector organisations and better facilitate public access to spatial information across Europe.

The Directive on the re-use of public sector information, more commonly known as the Public Sector Information Directive (PSI), provides a common legal framework for a European market for government-held data (public sector information). It focuses on the economic aspects of re-use of information rather than on the access of citizens to information. It encourages the Member States to make as much information available for re-use as possible.

UK

Following publication of its Open Data White Paper in June 2012, the UK Governments commitment to open data has been reinforced by the creation of data.gov.uk, which brings together data in one searchable website and shares data from all central government departments and some public bodies.

The Code of Practice for Official Statistics, published in 2009, requires producers of official statistics to ensure statistics are disseminated in forms that maximise accessibility and encourages analysis and re-use.

Wales

The Well-being of Future Generations (Wales) Act requires Public bodies, listed in the Act, including the Welsh Government, to think more about the long-term, to work better with people, communities and each other, to try to prevent problems and take a more joined-up approach. The legislation

introduces seven well-being Goals for Wales, accompanied by five 'ways of working'. Together, they provide a clear framework for decision-making and delivery that will underpin everything we, and the wider public service, do in future. Opening up and sharing data is key to enabling us all to work together to contribute to the goals and see what progress we are making.

Having access to and using our data effectively is a key enabler of delivering good digital services. The benefits associated with open data are therefore highlighted in both Digital First, the digital strategy for the public sector in Wales, and the Welsh Government Digital Action Plan. The latter of which contains a specific action to develop and implement a co-ordinated approach for making available our key information assets to help drive economic growth.

A new Welsh Government Information Strategy has been developed, as recommended by the National Archives Information Management Assessment, which aims to provide a framework around how we manage our information. It also explains how we are addressing the governments Information Principles, several of which can be addressed through open data.

Via its current Programme for Government report, Welsh Government has put transparency and accountability at the heart of what we do as a Government. This is to enable the people of Wales to know what their Government is doing on their behalf and how it's managing and spending public money.

**Making use of open data:
Bathing Water Quality App**

The Bathing Water Quality bilingual app, developed by Natural Resources Wales (NRW), enables people to check the quality of the beaches across Wales allowing them to make informed choices. The app contains public sector information, on all bathing waters in Wales, under the Open Government License (OGL).

NRW has also created the Bathing Water Widget Designer that allows people to customise a live feed of water quality data for a given bathing water site, or all of the sites in a given area, and display that feed onto their website.

Cyfoeth Naturiol Cymru
Natural Resources Wales

4. What are the benefits of Open Data?

Data is an extremely useful resource which if made accessible can provide a number of benefits and opportunities, including presentation of data in easy to understand formats.

Providing access to data can empower the people of Wales, the media, civil society, and businesses to create improvements in public services and inform improved decision making. It can be used to provide individuals and businesses with a wealth of information in order to carry out their personal and work business.

➢ **Improved government openness, transparency and accountability**
Open data can help provide the people of Wales with a better understanding of how Welsh Government and public services in Wales are performing and how budgets are being utilised.

➢ **Improved public services**
Opening up public sector data enables public services to be transformed through better planning and targeting of services. It also provides opportunities for collaboration and allows public bodies to become more efficient and make savings. For example, by making data available we provide the platform for others to make use of digital technology to develop apps that transform public services and give the citizen a better experience.

➢ **Innovation and economic growth**
Open data can help drive innovation through the development of applications that enhance people's daily lives and consequently economic growth. It also provides the opportunity for further innovation through predictive analysis.

The Open Data Institute (ODI), has identified, via its research 'Open data means business', 270 UK companies that use, produce or invest in open data. These companies have a combined annual turnover of over £92bn, employing over 500,000 people.

Data-driven support for Mental Health

Digital design studio M/A work have worked with mental health charity Mind to build an online hub and web app that pulls together open data from a whole range of sources with the goal of helping people with mental health conditions find opportunities for support and, eventually, employment.
http://theodi.org/summer-showcase-plexus

➢ **Public empowerment and participation**
Providing access to data can help inform people about the availability and accessibility of services. It enables people to make informed decisions, both in using and challenging

services provided. It also empowers people and allows them to participate more actively within their communities and wider.

> ➤ **Reducing burden**
> Making data openly accessible can reduce the need for Freedom of information (FOI) requests. Furthermore, if data providers open up their data this may also reduce the need to aggregate data to respond to data collection requests.

Whilst making data openly available can result in numerous benefits, it's not always easy to quantify the cost benefit. However, in 2013 an independent review of public sector information estimated the direct economic benefits of public sector information to the UK economy to be around £1.8 billion a year. It's also worth noting that the value of open data can appreciate economically through its re-use.

Prescription savings worth millions identified by Mastodon C

In 2012, Mastodon C worked with the ODI, Open Health Care UK and Dr Ben Goldacre to analyse open data around prescribing patterns of statins in GP practices, CCGs, and PCTs across the UK.

The current evidence shows that for most patients, all drugs from this class are equally safe and effective, so doctors are usually advised to use the cheapest. The analysis examined how much money was spent in each area on the more expensive drugs. If the research had been conducted a year before, potentially over £200m of unnecessary expenditure could have been saved.

http://theodi.org/stories

5. Risks and limitations

Whilst open data offers up many benefits, it is not without its risks and limitations both in terms of making data open and the use of it. However many of these risks and limitations can be mitigated against if careful consideration is given to what data can be made open and the means in which this is done.

Risk to privacy
Certain types of data should not be opened up as doing so could affect the privacy of an individual, business or organisation. However we can alleviate this potential risk by carrying out a Privacy Impact Assessment (PIA) as well as considering whether aggregation or anonymisation of data could prevent disclosive information being made available.

Incorrect interpretation of data
Some users may, through lack of understanding of the data and its limitations, misuse data. However, by providing advice on how the data should be used, the limitations associated with the data, as well as accurate and up to date information about the data (metadata), data providers can help reduce this potential risk.

Inaccurate data
It may not always be clear what, if any, validation data has gone through. Where no information is provided about the quality of the data, users need to be mindful of the possible limitations of the data they are using and consider whether the data is fit for purpose.

Inconsistent with aggregated data
In some instances inconsistencies may occur between open data and published aggregated data. This may be due to different levels of validation being carried out or data referring to different time periods. This in part can be addressed by data providers sharing accurate metadata about their data. We can also help mitigate this risk by publishing official statistics in as timely as fashion as possible to ensure there is a clear set of authoritative data available on nationally important data series.

6. Accessibility and usability of Open Data

Accessibility

For data to be classed as open it needs to be published in a format that is easy to access and be re-used. The 5 star rating scheme below was developed by Sir Tim Berners-Lee, creator of the web and President of ODI, and provides a simple indication of how open data is. Welsh Government will use this rating scheme to both assess the openness of our data and communicate it with others.

5 Star rating scheme

★ — Make your data available online (whatever the format) under an open license

★★ — Make your data available as structured data but proprietary format (e.g., Excel instead of image scan of a table)

★★★ — Make your data available in a non-proprietary open format (e.g. CSV as well as Excel)

★★★★ — Use URIs to denote things, so that people can point at your data (e.g. RDF)

★★★★★ — Link your data to other data to provide context

Quality

Furthermore Open Data Certificates, which Welsh Government is aiming to obtain, have been developed by ODI to measure how effectively someone is sharing a dataset for ease of re-use. Covering technical issues, rights and licensing, documentation, and guarantees about availability, the certificates offer a rounded assessment of the quality of publication of a dataset, providing a clear quality mark for re-users.

Knowing when data can be used and re-used

Data owners need to provide clarity to users on whether their data is open and can be re-used freely. Within the Public sector this is provided in the form of the Open Government Licence (OGL). The OGL is a simple set of terms and conditions under which information providers in the public sector can license the use and re-use of their data. If data is not made available under the OGL users will need to contact the relevant rights owner of the data.

Whilst an increasing proportion of the information Welsh Government publishes is available under the OGL, work is ongoing to make all appropriate information available under the OGL. However there will always be data that cannot be offered for use under OGL such as data protected by the Freedom of Information (FOI) Act and the Data Protection Act (DPA). Also some data used by Welsh Government cannot be made freely available due to license restrictions such as those imposed by the Ordinance Survey (OS) or other commercial data vendors.

7. What are we currently doing?

Making data available:

Whilst this is the first Open Data Plan published by the Welsh Government, work is already ongoing in the field of open data. Whether it is the publication of structured data or the dissemination of spatial data, the Welsh Government is committed to meeting the differing needs of our data users. As such it's not a case of one size fits all, but using the most appropriate tool for the data and its users.

Two of the main means of sharing data openly are via Lle, our open platform for spatial data, and StatsWales, the main site for accessing and downloading structured or aggregated statistical data such as Official Statistics.

> **Lle**
>
> Lle (which is Welsh for 'Place') is an open platform for spatial data which has been developed in partnership between the Welsh Government and Natural Resources Wales (NRW). Lle serves as a hub for data and information covering a wide spectrum of topics, though the original focus was on environmental data.

Lle was originally created to deliver the requirements of both Welsh Government and Natural Resources Wales (NRW) to comply with the EU INSPIRE Directive.

Built on a number of APIs (Application Programming Interface), or web services as they are sometimes known, most of the APIs or web services that serve the spatial datasets on Lle are compliant with Open Geospatial Consortium (OGC) standards. The use of these 'open standards' enable datasets to be easily discovered and re-used by digital tools and systems.

In order to enable users to view or download data to use in a wide variety of ways, all the data, where appropriate, published on Lle is licensed in accordance with the Open Government Licence (OGL), which means it is free to share and re-use for personal, research or commercial purposes.

Future developments planned for Lle include providing a web based access point to search, view and download data relating to the full range of data across Welsh Government and wider. These developments, alongside those planned for StatsWales, will help create an Open Data Service for Wales.

StatsWales

StatsWales is the Welsh Government's online repository for detailed statistical data. It was an early example of opening up statistics for re-use and manipulation by the user, having been launched over a decade ago. StatsWales allows users to view and manipulate datasets, produce charts and download data in a variety of formats. StatsWales includes data on Wales' population, economy, the environment, education and health.

As well as enabling users to manipulate data tables and create charts, StatsWales also allows users to download data in a variety of formats including CSV and Excel.

However developments are underway using OData, the open data standard, to continue to improve the openness of the data held on StatsWales. In addition to making data on StatsWales available in a machine readable format, an API (Application Programming Interface) will be introduced that will allow users to define the data they require and extract the data automatically. Not only will these developments help move StatsWales to 4 stars on the 5 star open data rating scheme, but they should also improve the ability to link data.

Improvements have also been made to improve the completeness, quality and structure of the metadata on StatsWales. Work will continue to ensure that the standard of metadata established is maintained and that the metadata is made available via the API once in place.

Figure 1: Data flows between Welsh Government websites
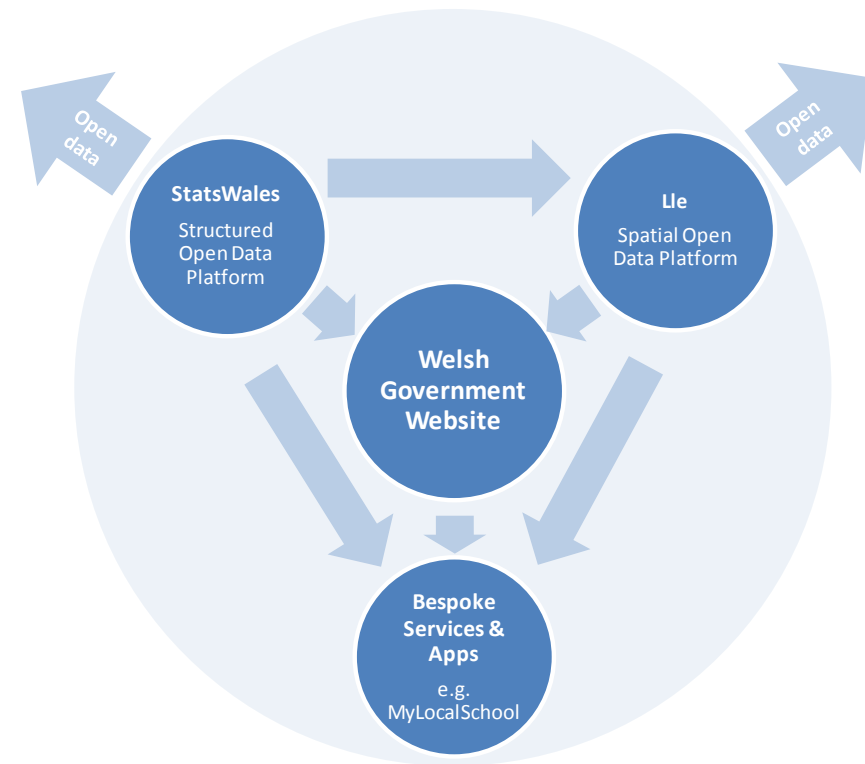
The previous diagram outlines how data flows between the various Welsh Government websites/platforms, with both Lle and StatsWales acting as the main data stores and means of sharing open data. In addition to opening up data, both platforms also feed data into various bespoke services and applications, current examples include MyLocalSchool and the WIMD interactive dissemination tool. They also both provide data for newly developed dissemination tools sitting on the main Welsh Government website, such as Welsh economy: in numbers and Local Authority Service performance.

Opening up data: Programme for Government

The current plan of action for Welsh Government puts transparency and accountability at the heart of what we do as a Government.

In addition to using graphical representation of performance indicators to openly show progress, the underlying data behind the performance indicators can also be accessed in a number of formats using our open API.

Making use of open data

To date much of the focus has been around the role of data owners and encouraging them to publish data openly. However this is only part the story. To be able to fully realise the benefits of open data, the data needs to be made use of. Therefore we want to work with other public bodies using the developments around StatsWales and Lle to re-use data in an automated fashion via other websites.

Whilst opening up data allows anyone to re-use the data; the Welsh Government is also committed to making the most of open data itself. Consideration is currently being given to a number of investigative pieces of work looking at how open data can help inform policy development. One example will draw together locations of grassland fires, the time of year and other data, such as environmental characteristics and areas of deprivation within Wales. Mapping this information together should provide an insight into the sources and characteristics of deliberate grassland fires that could in turn help identify areas of high risk and inform deterrence work. Making this available publically will allow these insights to be shared with other agencies such as local authorities and the third sector.

In July 2015 the Welsh Government announced a new Digital Innovation Fund that will explore how digital technology can innovate public services in Wales. The Digital Innovation Fund should also provide opportunities in the re-use of open data.

8. What do we plan to do?

In order to develop further as both a publisher of open data and user, the Welsh Government aims to focus on the following commitments which take into consideration the G8 open data principles:

Open data by default:
(i)   Develop StatsWales to increase its openness rating to 4* by May 2016.
(ii)   Publish all new information on Welsh Government website in open formats from 1 April 2016.
(iii) Publish all Welsh Government owned information, where appropriate, under the Open Government Licence (OGL).
(iv) Provide support to public sector bodies in making use of open data through sharing of information and encourage them to publish their own open data plans.

Quality and quantity:
(v)   Obtain ODI open data certification for StatsWales and Lle.
(vi) Develop and publish a catalogue of datasets published openly by the end of 2016.

Useable by all:
(vii) Develop an Open Data Service for Wales, building on the existing Lle and StatsWales infrastructure, with a focus on helping to improve public services by the end 2016-17.

Releasing data for improved governance and innovation:
(viii) Identify the open data needs of users to prioritise the datasets to publish openly through ongoing engagement with internal and external users.

Making use of open data:
Food Hygiene Ratings App

Hygienic, the food hygiene ratings app, is a mobile application created to help people check official food hygiene ratings and search for the most hygienic places to eat, thereby helping them make informed decisions about where they eat. The app, developed by Hookshot Ltd., draws on the UK food hygiene rating data published by Food Standards Agency under an Open Government Licence (OGL).

SGÔR HYLENDID BWYD
FOOD HYGIENE RATING
0 1 2 3 4 **5**

© Crown Copyright.
Food Standards Agency Food Hygiene Rating Scheme 2014

9. What does this mean for you?

Achievement of the Welsh Government commitments in section 8 should provide more clarity to officials on what data is openly available and how it can be accessed. It should also mean that our stakeholders and citizens also have greater access and are therefore more informed.

If you are planning to collect data or are already a data owner you may want to give consideration to whether there are any reasons why your data shouldn't be published openly. If there's not and you want advice on how best to share and publish your data, there is plenty of support available both within your departments (Departmental Knowledge and Information Managers (DKIMs) and Information Asset Owners (IAOs)) as well as centrally (Knowledge and Analytical Services (KAS) and the Office of the Chief Digital Officer (CDO)).

Monitoring progress
Progress made on achieving the Welsh Government commitments outlined within this open data plan will be continually monitored and reviewed. Further to this, the Office of the Chief Digital Officer (CDO) within Welsh Government will report on progress on an annual basis.

Open data can help you get where you need to be for work or pleasure

Moovit, launched in the UK in January 2015, is a free public transport app that draws on open data and live inputs from users to provide a variety of journey options and estimated arrival times across many cities.

Helping users make decisions about their journeys, Moovit incorporates a journey planner, real-time travel information (including Tube line statuses and bus arrival and departure times), a favourite routes dashboard, narrated and a live map that shows nearby stops, the lines that run through them and upcoming arrival times.

Source: Moovit – Logo by Kateaz777 (Own work) [CC BY-SA 4.0 (http://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons

10. Glossary

Aggregated data
A combination of unit records created with the objective that individual details are not disclosed.

API
A specification intended to be used as an interface by software components to communicate with each other. An API may include specifications for routines, data structures, object classes, and variables.

Closed data
Data that is personal and can identify an individual or data that for security reasons cannot be made open. Can also include data which is closed as no decision has been made as to whether it should be open, shared or closed.

Dataset
A collection of data, usually presented in tabular form, presented either electronically or in other formats

Linked data
The term used to describe the recommended best practice for exposing, sharing and connecting items of data on the semantic web using unique resource identifiers (URIs) and resource description framework (RDF).

Machine readable
Data which is in a format that can be understood by a computer.

Metadata
Data that describes or defines other data. Anything that users need to know to make proper and correct use of the real data, in terms of reading, processing, interpreting, analysing and presenting the information. Thus metadata includes file descriptions, codebooks, processing details, sample designs, fieldwork reports, conceptual motivations, etc., in other words, anything that might influence the way in which the information is used.

Open data
Data which can be used, re-used and re-distributed freely by anyone - subject only at most to the requirement to attribute and share-alike. There may be some charge, usually no more than the cost of reproduction.

Shared data
Data that is shared only with named people or organisations or under terms (licences) and conditions that are not 'open'.

Spatial data
Spatial data refers to all types of data objects or elements that are present in a geographical space or horizon.

Star rating scheme
In UK Linked Data, a system of ranking data sources that indicates ease of machine readability. It is not a measure of the quality of the data content.