A TOOLKIT TO OPEN GOVERNMENT

GUIDE TO OPEN DATA

The digital age allows us to access vast amounts of information, including government data.

By making official data available for everyone to use and share, governments put power into the hands of citizens.

This guide explains what open government data is and how it can help to drive improvements in how we live and are governed.

CONTENTS

TASC
Think-tank for Action on Social Change

WHAT IT IS

Data and open data

Governments and public bodies routinely collect and hold vast amounts of information, including personal data, survey responses, administrative information and scientific measurements.

They regularly use this data to help decide how best to allocate public resources – for example, where to build new roads, hospitals, houses and schools.

The sort of official data that is routinely collected includes:

- Finance – government revenue and spending figures.

- Health – hospital waiting times or infection rates.

- Transport – timetables, real-time traffic updates and passenger statistics.

- Education – statistics on early school leavers and educational achievements.

- Population – census information and household poverty statistics.

- Environmental – bathing water quality or air pollution levels.

Traditionally, public bodies and government departments have kept much of their data to themselves. They have generally decided which data to release, when and in what formats.

The conventional formats for releasing data include reports, strategies, press releases and statements. In all of these formats, the data is already interpreted and packaged.

Why open up data?

Increasingly, governments and public bodies worldwide are sharing the underlying or 'raw' data that they gather to make their reports and strategies etc.

The idea is that official data released in its raw form can have many uses and applications including for social, economic, democratic or cultural purposes.

Modern technology means that even very large amounts of raw data – in tables, databases or spreadsheets – can be released at little or no additional cost to the public bodies that routinely collect and hold it.

For example, a government health plan might include a table showing surgery waiting times in all the country's hospitals. If the raw data used to make the table was published in a digital spreadsheet, then it could be reused and presented in different ways, for example to highlight different waiting times on a regional or county basis, or by surgery type.

Open data is not about personal information

Open data refers to information that is non-personal and non-sensitive. Personal data that is collected by public bodies should only be released as open data after it has been anonymised or aggregated so that individuals' privacy rights are respected.

What makes data 'open'?

These are main the technical and legal elements that make data 'open':

Data must be accessible: The data must be publicly available for anyone to use. This means downloadable via the internet, for free or at minimal cost.

Data must be machine-readable: This means that the data must be in a format that can be read by a computer. A PDF document is an example of data that is not machine-readable – it is more like a scanned image. One step beyond machine-readable data is structured data. This format allows computers to search and sort the data.

Data must be in an open format: An open format is one which places no restrictions, monetary or otherwise, upon its use. At the very least, data should be in a format which can be processed with at least one free software tool. For example, Microsoft Word or Excel are not open formats as you have to buy them. CSV (Comma Separated Values) format is preferable to Excel for tabular or spreadsheet data because it is free for anyone to use.

Data must be legally open for reuse: Many electronic files come with licences or restrictions attached to them. For example, if you search online for a photograph to use on a website, you may find that its licence means that you have to pay a fee, or you are not allowed to use it at all. For it to be open, data must have a licence that allows for its legal reuse, including for both non-commercial and commercial purposes.

🧩 The Open Definition

For more detail on what makes data open, see the Open Definition produced by the Open Knowledge Foundation.

There is also a five-star rating system for open data developed by Sir Tim Berners-Lee, the inventor of the world wide web.

What open data can do for us

In today's digital age, open data has been described as the raw material of the new industrial revolution.

The opening up data held by governments is not an end in itself – the release is only the start of the story.

It is what data users – from activists and journalists to analysts and entrepreneurs – actually do with it that gives open data its value in terms of economic, social and democratic benefits.

The benefits of open data include:

• **Government transparency and accountability:** When official data used to make public decisions is released, it allows citizens to know what the government is doing. The data can be analysed and interpreted by specialists who can identify trends, spot gaps, query policy choices and suggest alternatives.

• **Social and commercial development:** Open government data can be put to use for civic and commercial ventures. Examples include a mobile phone application or

'app' that allows commuters to easily navigate public transport networks in cities worldwide, or a website that helps people to track stolen bikes.

- **Citizen participation in democratic decision-making:** By opening up data, governments, public bodies and local authorities can enable people to make better informed choices about the services they receive and the standards they should expect. Budget data released by local authorities, for example, can allow citizens to better analyse the quality of local services and flag areas of waste.

In a nutshell: data, open data and big data

- *Data is a collection of facts and figures such as statistics or measurements that can be analysed.*

- *Data is considered 'open' if it is made available for anyone to use for any purpose at no cost.*

- *Big data refers to information that is bulky, like national statistics on health care provision or education. This kind of data usually consists of hundreds of individual bits of information held in digital formats such as spreadsheets.*

Open data and Freedom of Information

Open Data and Freedom of Information (FOI) are related. Yet they often serve very different goals and have different objectives.

Open data is about public bodies automatically and routinely releasing 'raw' data without anyone asking for it. This could include geographic data, postcodes, transport data, or data about corporations.

FOI on the other hand is mostly about data that public bodies release because someone asks for it. This can take the form of raw data like spreadsheets. But it is often qualitative information, like emails, internal memos or guidance documents, or minutes of meetings.

While anyone can use their rights under FOI to ask for their own personal records, open data is only ever about the release of non-sensitive and non-personal information.

For more on how our right to know under FOI works, read the Toolkit Guide to FOI.

HOW IT WORKS

Open data in Ireland

While open data is catching on in Ireland, its benefits have yet to be fully exploited.

Not all public sector bodies routinely publish raw data sets for others to freely reuse. Some public bodies store data in formats that are not easily published as open data because they are not fully machine-readable.

As part of the government's Open Data Initiative, government departments and public bodies throughout the country are being encouraged to identify and release key data sets in open data formats.

Citizens, activists and entrepreneurs are also increasingly tuning in to the potential ways to reuse public sector open data.

Public sector open data licence

Governments can create their own licences that they attach to official data to make it 'open data' – that is, data that is legally available for anyone to reuse.

The Open Data Licence used for public sector open data in Ireland is called Creative Commons Attribution (CC-BY 4.0). This licence allows other users to distribute, remix, tweak, and build upon data, including for commercial purposes, as long as they credit the originator.

The introduction of this licence in 2015 allows data users to be legally certain about what they can and can't do with the data.

The licence that comes closest to open data is the Public Sector Information (PSI) Licence. This is widely used by many public bodies and government departments. However, PSI is not considered to be a truly open licence as it imposes legal restrictions on how data can be re-used or distributed.

Ireland's open data portal

The website data.gov.ie is a searchable directory of public sector data that is available for free re-use. The data is published by government departments, State agencies and local authorities.

For example, the portal includes statistical data from the Census as well as environmental, housing and heritage data. People can suggest new datasets that they would like to see published as open data.

Dublin councils' open data

The four local authorities in the Dublin region have taken the lead in releasing useful datasets in open data formats on a dedicated website, www.Dublinked.ie.

This site allows the councils to work closely with developers and analysts to trial new products and services that use their open data. The four local authorities are: Dublin City Council; Dún Laoghaire-Rathdown Council; South Dublin Council; and Fingal County Council.

The website contains constantly updated data on:

- land use and planning

- real-time bus and rail passenger information

- real-time parking space availability

- fire brigade and ambulance call outs

Dublin data on public services

Some of the data published on Dublinked.ie has been mixed with separate datasets from government departments and various other official sources to produce detailed data visualisations and interactive maps on the website www.dublindashboard.ie.

Dashboard users can access detailed, up-to-date information about many aspects of the city and public services. For example, the site allows users to monitor traffic flows in real time, to view charts tracking crime levels, and to find out the number of patients waiting on trollies in Dublin's main hospitals. The raw data used on this website can also be reused by anyone.

 Open Data world leaders

The amount of government data that is available in Ireland as open data is limited in comparison with other western European countries, according to its ranking in the 2015 Global Open Data Index.

The United Kingdom is considered a world leader when it comes to releasing public sector data as open data. Up to mid-2016, the UK had released around 30,000 data sets from government departments and agencies. Ireland had released more than 4,000 official data sets for re-use by this date, although not all of these are in true open data format.

For more information on how Ireland compares with other countries when it comes to open government data, see the Open Data Barometer.

What open data should governments publish?

Different users of official open data have different priorities for the kinds of data they want to see released. There is general agreement worldwide that some types of data that governments hold have a 'high-value' in terms of their impact and potential for reuse.

Ireland has signed up to the primary international standard for open data, the G8 Open Data Charter. This defines high-value data as data that improves democracy and encourages innovative re-use.

Examples of high-value data sets include:

- Budget/spending data

- Companies data, like company registers

- Crime statistics

- Meteorological data

- Pollution levels data

- Geospatial data, like national maps

- Global aid spending data

- National statistics and census data

- Election data

- Roads networks and public transport usage data

Open data's impact

Here are some ways in which open data is having an impact in Ireland and throughout the world:

Opening up government spending

Health care costs: Open data helped identify a £200 million saving in the UK's National Health Service budget. A 2012 study used publicly open GP prescription data to show how the government was wasting money buying branded cholesterol-reducing drugs, rather than using cheaper generic versions.

Where Does My Money Go is a free online visualisation tool that allows users to understand where public money in the UK is spent. It uses spending data from government departments and local authorities.

OpenSpending is a global open database of public financial information, including budgets and spending data. It exists to track and analyse public financial information globally. It relies on a community of users and contributors to upload official datasets to its open database.

Local Authority Finances is a website that allows you to compare budget data on 29 out of 31 local authorities in Ireland. You can:

- investigate your local council's spending and revenue

- compare your council to the national average

- compare your council to a selection of other councils

You can also download the data yourself as an open data set.

Tracking the work of public representatives

Websites like KildareStreet.com in Ireland and TheyWorkForYou.com in the UK allow people to monitor the issues their elected representatives are active on. Both of these websites are run by non-governmental organsiations, and both reuse official parliamentary debates data.

WhoIsMyTD.com also uses official Irish data to allow users to insert their postal address and find out who their TDs, MEPs and local councillors are.

Improving public services

The UK's Crime Map allows people to explore and compare crime and outcomes of crime in their neighbourhoods. It uses police open data.

Climate Corp in America uses government open data on crop yields and weather patterns to provide services that help farmers make business decisions.

Locatable helps people in the UK decide where to live. It uses various open datasets about public transport, property prices, crime rates and public services.

ParkYa is one of several mobile phone applications or 'apps' that help commuters find and pay for public parking spaces in Dublin. It uses real-time public transport open data from Dublin's four local authorities.

Fixyourstreet.ie allows users to report nuisance issues like graffiti, broken street lighting or litter to any local authority in Ireland. The local authorities post replies and updates on the website.

Highlighting corporate networks

Open data can be used to decipher the complex corporate networks used by big companies worldwide. For example, the Open Corporates website helps people find out more about what corporations are doing worldwide. It is the largest open database of companies in the world, with data for 49 million companies which it takes from publically available company registers.

Other open data projects which investigate business include projects on:

- The use of anonymous shell companies to avoid tax

- The corporate structures of Big Oil

DO-IT-YOURSELF

Working with open data

Here are some of the types of users who can put public sector open data to good use:

- Researchers and civil society activists can use official statistics to analyse trends, gain fresh insights into how policies are working in practice and propose evidence-based reforms.

- Public bodies themselves can use public sector open data to identify emerging needs and promote more effective and targeted service delivery.

- Entrepreneurial and technical users can use raw data files to develop new products or improve existing ones.

Things to consider when deciding to use open data

If you think you or your organisation could benefit from working with open data, it is worth asking yourself the following questions to begin with:

What kind of datasets are relevant to your interest or your work?

Are they already available as open data? Are they suitable for your needs? For example, the Central Statistics Office might have data that is relevant to the work of many different organisations or community groups.

Given the lack of available open datasets in Ireland, you may not be able to find the data that you need. But you could request the data as open data via https://data.gov.ie/data and/or lobby for it to be made available.

Can you work with the data on your own, or do you need help?

You don't need to have advanced computer skills to take a piece of open data and explore it. Small open datasets can easily be analysed using familiar desktop tools like spreadsheet software Excel.

However, data analysis usually requires someone with analytical skills. More advanced analysis, or analysis of complex datasets and interpretation of results, can require specialist skills and knowledge. For data to be visualised in ways that make it more accessible and understandable may also require the services of computer programmers and graphic artists.

Groups or individuals who are interested in open data analysis could consider pooling resources with others to commission technical experts to help analyse and make use of data.

There are many open data enthusiasts in Ireland who hold regular 'meet ups,' hackathons or data 'dives' where they take data sets and explore them. These voluntary groups are often on the lookout for good causes and good datasets to work with (see below).

Groups working on open data in Ireland

Dublinked Innovation Network is a hub for businesses and other users who are interested in using public sector open data to build new products and services.

Dublinked regularly hosts workshops and other events to encourage open data-based collaboration and innovation. Membership of the network is open to individuals, companies and universities who have a demonstrated interested in the Dublin region, or are undertaking activities relevant to the Dublin region.

Datakind Dublin is a Dublin-based voluntary group which arranges for data scientists to team up with non-profit organisations and charities to help them 'mine' or explore data. They offer advice and ideas on how to use, interpret and visualise data, including open data. They are part of the worldwide DataKind movement.

Open Data Ireland is a loose collection of open data enthusiasts who use a Google Group mailing list to share news and information. The community organises evening or weekend 'meet-ups' and workshops, mostly in Dublin, to discuss different aspects of open data and open government. These are open to anyone to attend.

Open Knowledge Foundation Ireland is a non-profit group which specialises in open data and the use of online tools to increase citizen participation in public decision-making. It organises hackathons, training events and meet-ups that are open to all. It is part of the global Open Knowledge Foundation (OKFN) non-profit network.

Code for Ireland is a voluntary organisation of computer developers, citizens and government workers who work together to develop applications using open data. The group uses 'crowdsourcing' techniques to build websites or mobile phone applications by inviting people to come together to collaborate on them. Its focus is on services that would improve community services. It holds monthly events.

PEOPLE'S STORIES

KildareStreet.com - helping you keep in touch with the work of your elected representatives

KildareStreet.com is Ireland's largest open data project to date. The website provides a readable and searchable copy of the parliamentary record – what was said in the Dáil, Seanad and Oireachtas Committees.

KildareStreet.com takes official data from the Oireachtas website and presents it in a user-friendly format and with extra features. For example, users can set up dedicated email alerts to find out when certain topics or words are mentioned in the proceedings of the Dáil and Seanad. KildareStreet.com also has a strong search function.

It also produces statistical information on every TD and Senator, including:

• their voting records

• their declarations of private interest, such as shares that they hold or land that they own

• the number of times they have spoken in debates

• the number of written parliamentary questions they have submitted to Ministers

The open data

KildareStreet.com uses data that comes from the debates section of the official Oireachtas website.

The developers

KildareStreet.com was set up in April 2009 by John Handelaar and fellow journalist Gavin Sheridan from thestory.ie

Handelaar said he developed the site after moving to Ireland from the UK and trying unsuccessfully to find the identities of his local public representatives on the internet.

"I was living in an electoral division that kept changing and I as a citizen wanted to find out who my local councillors and TDs were," explains Handelaar.

"That search ended up in KildareStreet.com getting built, because it needed doing. The Oireachtas was publishing structured open data in a place where it could be predictably found. However, when I started using the data, it wasn't linked to from anywhere online. I just started guessing urls [internet addresses] and found that the Oireachtas was publishing the data, and started using it."

KildareStreet.com is based on the UK website TheyWorkForYou, which is part of a suite of online political and civic 'self-help' sites developed by the UK charity MySociety.

RESOURCES

Open data portals

National organisations publish significant amount of data. Much of this data is in open data formats, such as machine-readable spreadsheets with an appropriate licence allowing it to be freely reused.

Here are some of the main sources for public sector open data:

Public sector data

data.gov.ie, is a searchable directory of official datasets that are available for re-use by anyone. These are published by government departments, state agencies and local authorities

Official statistics portal

www.statcentral.ie is a portal that provides information about and links to all official statistics produced by government departments and state organisations. These cover the economy, business and the labour market, as well as statistics on people and trends in society. Some of the data that this site links to is in open data format. The portal is maintained by the Central Statistics Office.

Public finance databanks

There are two government databanks that contain data on public spending and revenue raising.

Department of Public Expenditure and Reform Databank: http://databank.per.gov.ie/ contains data on public expenditure from 1994 to the present, and the numbers of people employed by the state since 1980.

Department of Finance Databank: http://databank.finance.gov.ie/ shows Exchequer Tax Receipts on a monthly basis from January 1984 to present.

Spatial data

The All-Island Research Observatory (AIRO): AIRO, based at Maynooth University, is the leading spatial analysis and planning unit within the National Institute for Regional and Spatial Analysis. The AIRO data store contains many data-sets from a wide range of sources and sectors, as well as data maps or visualisations.

Irish Spatial Data Exchange (ISDE): ISDE is a data discovery tool which allows users to find spatial data and services that are produced by government and academic organisations.

Local authority data

Dublinked: Dublinked is the open data portal of the four local authorities in the Dublin region – Dublin City Council, Dun Laoghaire-Rathdown County Council, Fingal County Council and South Dublin Council. These authorities release their datasets on this website for developers and others to reuse to build new services.

Councils which also publish open data include:

- Fingal County Council

- Clare County Council

- Galway County Council

- Limerick City and County Council

- Roscommon County Council

Information about open data in Ireland

Key publications detailing the open data landscape in Ireland, produced by the Insight Centre for Data Analytics, were published in May 2014. They are:

Open Data Ireland: Best Practice Handbook: This handbook draws together existing best practice standards for the publication and re-use of open data. The handbook aims to assist in establishing best practice standards in Ireland.

Open Data Ireland: Data Audit Report: This is an audit of the Irish public sector datasets available online as of May 2014. It shows how many 'high-value datasets' are available as open data.

Open Data Ireland: Roadmap: This sets out a detailed 3 year plan for the progress on open data in Ireland.

Open Data Ireland: Evaluation Framework: This is a set of recommendations to government on how to evaluate its efforts as part of its Open Data Initiative.

Open Data Ireland: Publication Handbook: This is a step-by-step guide for public bodies publishing open data on the national open data portal, www.data.gov.ie

General information about open data

Open Data Handbook is an introduction to the legal, social and technical aspects of open data, produced by the Open Knowledge Foundation. It can be used by anyone but is especially useful for those working with government data. It discusses the why, what and how of open data – why to go open, what open means, and 'how to do' open data.

Data Journalism Handbook is a resource book for anyone who thinks that they might be interested in becoming a data journalist, or dabbling in data journalism.

Open Data Institute is an independent, non-profit organisation set up by the creator of the world wide web, Sir Tim Berners-Lee, and Professor Nigel Shadbolt. It carries out research, training and mentoring and works to promote innovation around open data globally. It has a series of guides and case studies that provide detail on open data.

The G8 Open Data Charter: The G8, or Group of Eight, refers to the eight highly industrialised nations — France, Germany, Italy, the United Kingdom, Japan, the United States, Canada, and Russia — that hold an annual meeting to foster consensus on global issues like economic growth and crisis. The G8 pledged in its first Open Data Charter in 2013 to make government data 'open by default' and 'usable by all'. Ireland, while not a G8 member, has endorsed this charter.

Listen to our podcasts

For more information on topics covered in the Toolkit guides, you can listen to our six-part podcast series – How to be an Active Citizen. These educational programmes aim to help listeners become active citizens in their own lives. The series was broadcast on RTÉ Radio 1 Extra in May/June 2016. The podcasts are available on www.tasc.ie/activecitizen