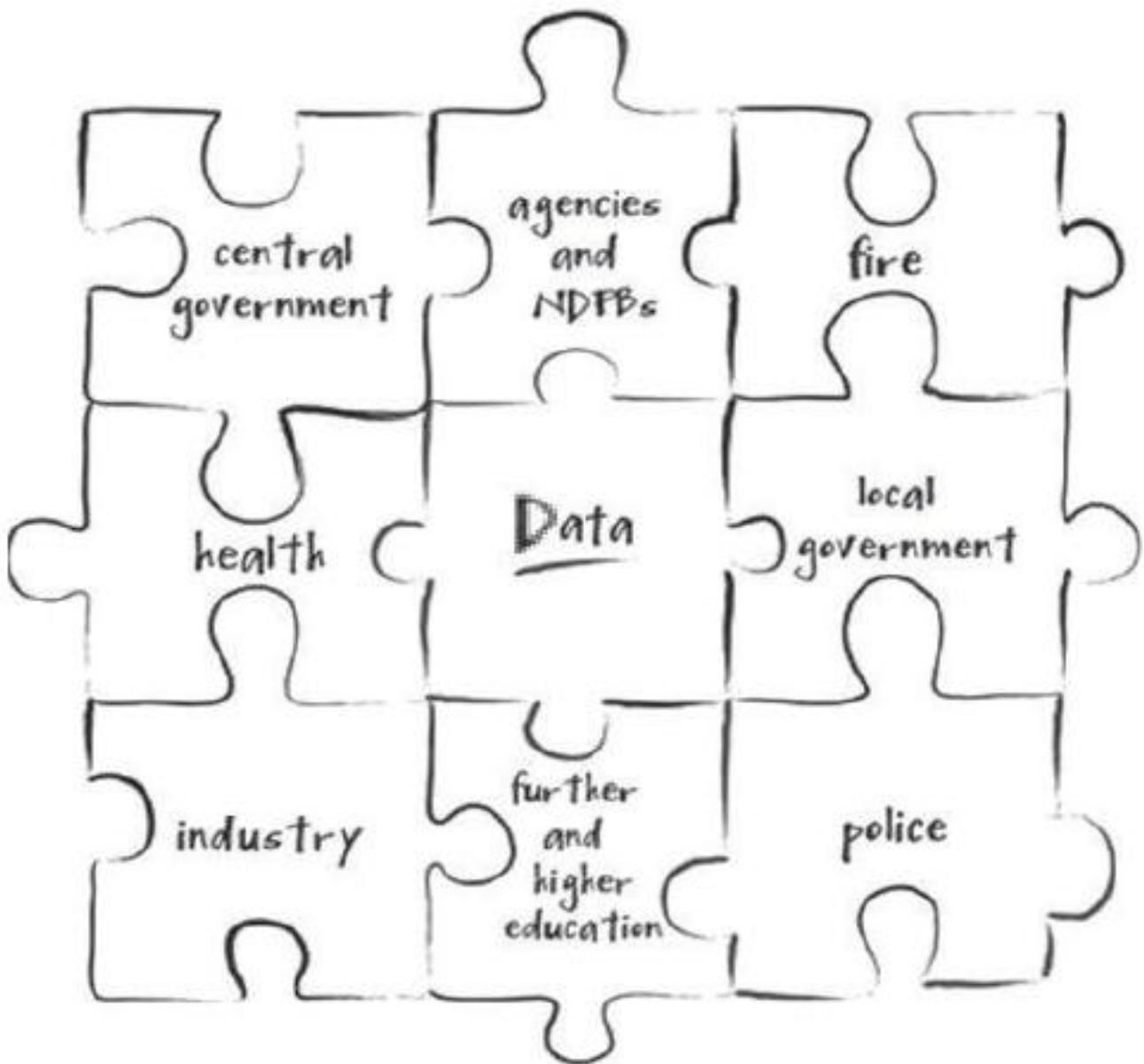# Open Data Resource Pack

Version 2.0

Scottish Government
Riaghaltas na h-Alba
gov.scot

Why and who should use this guide?

In 2015 the Scottish Government launched its Open Data Strategy. The Scottish Government has designed this Resource Pack to support the strategy. This Resource Pack is intended to help public authorities throughout Scotland develop and implement their own plans for open data. The pack has been developed with support and input from public sector colleagues from across Scotland. It has been specifically designed to assist Scottish public authorities who want to make their data open, but it can be used by anyone with an interest in open data. It is intended for those with little or no knowledge of the topic.

This Resource Pack is a living document and will be updated as open data work in Scotland progresses. This first version was published in August 2015 and this is now the second . Additional case studies have been added, and sections have been updated to reflect both relevant developments over the past year and user feedback. We welcome any further feedback or suggestions and will incorporate this in future editions. If you would like to get involved in discussions or make suggestions please join our Open Data Network on Knowledge Hub. You can also get in touch by emailing OpenDataPolicy@gov.scot.

The Open Data Resource Pack has been divided into 11 sections and 3 annexes. If you are new to Open Data you may wish to start with section 4.

1. Scotland's Open Data Strategy

The Scottish Government launched its Open Data Strategy in February 2015, adopting the G8 Open Data Principles. The strategy seeks to create a Scotland where non-personal and non-commercially sensitive data is recognised as a valuable resource which is made available openly for use by all.

Scottish public services are expected to release data in a way which aligns to the G8 Open Data Principles. The principles are shown in the diagram below.
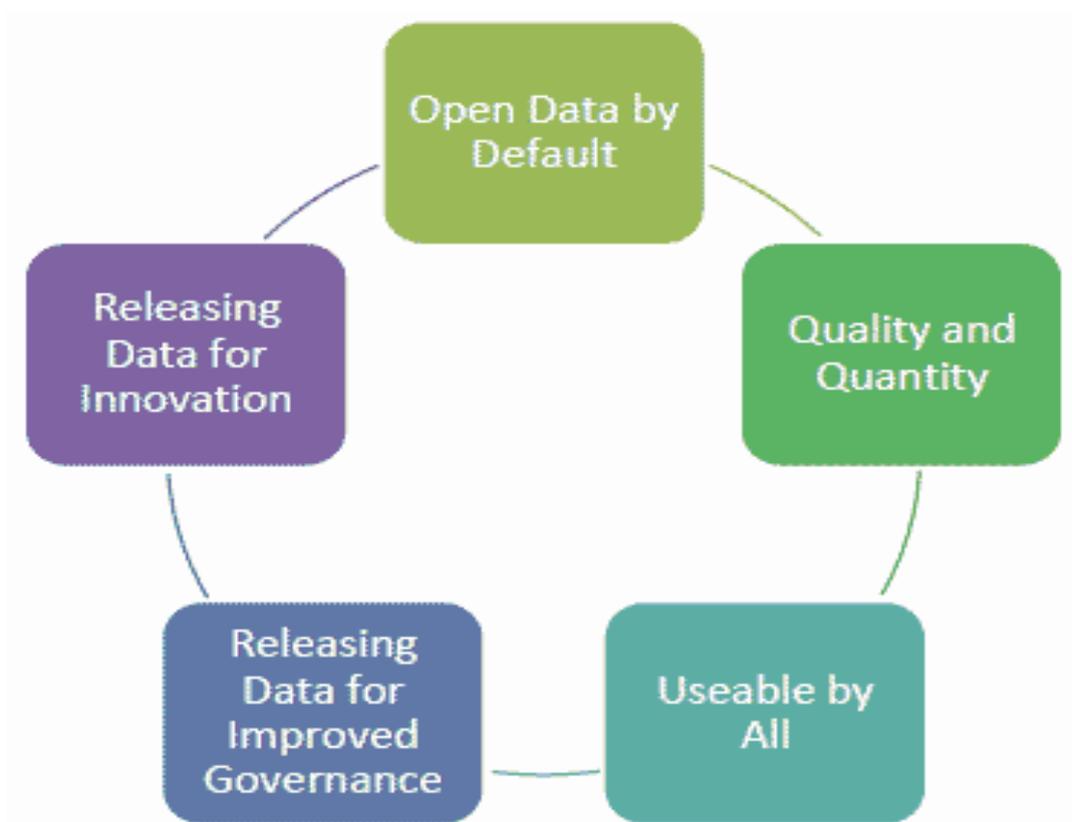
- **Open Data by Default** - Those holding public data should make it open and available for others to re-use. Those collecting new data should make sure that releasing data for re-use is built into the process. Over time releasing data openly should become the default business practice.

- **Quality and Quantity** - The amount of public data available is huge but the data quality varies. Published data must be supported by metadata. Metadata provides information about the data itself. Good metadata allows re-users to understand the data and its limitations.

- **Useable by All** - Data should be published in a manner which supports both easy discovery and easy re-use of the data. This includes making sure it is in a format which supports re-use and it has an open licence. Data will be made available free, with defined exceptions.

- **Releasing Data for Improved Governance** – Public authorities will release data which supports delivery of better public services. They will use the data to improve the services and policies they deliver. Public authorities should aim to engage and inform the public through the release of open data.

- **Releasing Data for Innovation** – Release of data will create wider economic and societal benefits. Others will be encouraged to make use of the data and develop new products or services for non-commercial and commercial use.

By 2017, all public authorities in Scotland should be publishing their data in a format of 3 star or above. 3 star data is data which is made available online, with an open license, in an open and machine-readable format. Section 8 explains the steps to achieve 3 star data release.

2. Existing Legislation

The [Open Data Strategy](#) supports a wider legislative initiative which has been designed to get public sector organisations across Europe releasing their data on a regular basis.

The strategy recognises and supports the duties placed on organisations under access to information legislation in Scotland and the UK. Whilst the Open Data Strategy encourages the proactive release of data in a manner currently beyond that required by legislation, legislation in this area can change over time. It is critical you understand your obligations under five key pieces of legislation. It is also important that you understand what powers the legislation gives you to support open data.

- [Freedom of Information (Scotland) Act 2002](#)
- [Environmental Information (Scotland) Regulations 2004](#)
- [Data Protection Act 1998](#)
- [INSPIRE (Scotland) Regulations 2009](#)
- [Re-use of Public Sector Information Regulations 2015](#)

You should also be aware of any other specific legislation which may govern your organisation's activities or obligations to disclose or withhold information.

Freedom of Information (Scotland) Act 2002 (FOISA) and

Environmental Information (Scotland) Regulations 2004 (EIRs)

FOISA and EIRs are enforced by the Scottish Information Commissioner.
The legislation gives any person the right to ask for recorded information held by a Scottish public authority and the right to receive it. Requests for information under FOISA must be in writing. Requests for environmental information under the EIRs may also be oral.

The Scottish Information Commissioner has created a [short guide](#) to help authorities decide which law applies to any particular request. More [detailed guidance is](#) available to help public authorities understand their obligations and deal with requests. There will also be a FOISA contact within your organisation who you should speak to if you have any questions.

The legislation imposes two broad duties on Scottish public authorities:

1. The duty to respond to information requests within statutory time limits**.**

There is a presumption that information will be disclosed but the legislation makes provision for it to be withheld if exemptions or exceptions apply. These exemptions and exceptions are also relevant when releasing open data. This should be taken into account when developing your Open Data Publication Plan.

2. The duty to proactively publish information.

Section 23 of FOISA requires authorities to "adopt and maintain a scheme […] which relates to the publication of information by the authority…". This is called a 'publication scheme'. Open Data is one class of information which the Scottish Information Commissioner has included in the Model Publication Scheme which she has prepared under section 24 of FOISA.

Data Protection Act 1998 (DPA)

The DPA is the main piece of legislation which protects personal data in the UK. Anyone who processes personal data is obligated to comply with the Act. The Act defines 8 data protection principles. If you have questions about your obligations under the DPA, you should speak to your Information Asset Owner (IAO) or Data Protection Office (DPO). Alternatively, you can consult the UK Information Commissioner's Office (ICO) guidance.

The DPA is based on the 1995 European Union Data Protection Directive. The EU has just passed a new General Data Protection Regulation (GDPR) which increases individuals' rights and imposes further obligations on organisations processing personal data. The GDPR will come into force in May 2018. Even though the UK is leaving the EU it is likely that we will have to comply with the GDPR to ensure we have an adequate level of protection for processing EU citizens' personal data. The Information Commissioner's Office (ICO) has guidance on the GDPR available here.

INSPIRE (Scotland) Regulations 2009

The INSPIRE Regulations cover the release of spatial data by Scottish public authorities and organisations which carry out duties on behalf of public authorities. The regulations brought into force the EU Inspire Directive which aims to create consistency across all member states so that all spatial datasets can be easily shared, modified and combined. INSPIRE defines common technical standards for publishing spatial datasets which fall within 34 data themes.

All spatial data which falls under INSPIRE must be published. You should ensure that you publish your data as both view and download services, to the required standards, as well as its associated metadata. The latest information about INSPIRE can be found on the UK Location website.

In Scotland, the Scottish SDI Metadata Catalogue (SSDI) supports the publishing of Scottish public sector spatial data metadata to the INSPIRE and UK Location metadata standards. It provides the discovery component for a set of on-line services that will allow users to evaluate and use public sector spatial data. Guidance on how to create INSPIRE metadata on the SSDI can be found here.

Re-use of Public Sector Information Regulations 2015

The Public Sector Information Regulations cover any information a public sector body:

    a) produces, holds or disseminates within its public task, and
    b) holds the copyright for.

Under the regulations, public sector bodies should make their information available for re-use under an open licence at marginal cost. Marginal cost in most cases will be nil. The legislation specifies exceptions to the marginal cost default, for example trading funds. In such cases, public sector bodies may charge re-users to cover the costs of collection, production, reproduction and dissemination of information, together with a reasonable return on their investment.

The National Archives have produced a checklist to help public sector bodies make sure they satisfy their obligations under the regulations. Detailed guidance, including information around charging, has been produced for public sector and cultural sector bodies.

3. Scotland's Open Data in an International Context

The Open Government Partnership

The Open Government Partnership (OGP) was launched in 2011 to provide an international platform for those committed to making governments more open, accountable and responsive to citizens. The UK was a founding member of the OGP. It now includes 69 participating countries, and works to develop and implement open government reforms which ensure people can see, understand, participate in and influence the workings of government. It brings together governments, civil society and business to tackle significant societal reforms, through transparency, accountability and increased citizen participation.

In April 2016, Scotland was selected to join the Open Government Partnership's Pioneer Programme. This programme includes devolved administrations, regional and local governments, municipalities and major cities. Under the Programme, each Pioneer will work together and directly with the Partnership to develop a national action plan by December 2016, which will then be implemented and independently reported on in 2017. Scotland is one of fifteen governments selected to join the Programme. The national action plan themes are Open Data; access to information; technology and innovation; civic participation; public accountability; and anti-corruption.

Pioneer status provides an opportunity for Scotland to work and share learning with other Pioneer governments around the world, and to showcase open data reforms at the Open Government international summit in Paris in December of this year.

Since 2015, Scotland has also been working as part of the intergovernmental working group developing the 2016-18 Open Government national action plan for the UK. The plan has been developed by the four governments in the UK, working with civil society. A draft plan with 30 commitments under the national action plan themes has now been developed. The national action plan 2016-18 is available to view here.

Organisation for Economic Co-operation and Development (OECD) Ministerial Declaration on the Digital Economy

In June 2016, the OECD invited Ministers from member and partner countries – including the UK – to attend a high-level meeting under the heading The Digital Economy: Innovation, Growth and Social Prosperity. The meeting was arranged to discuss and develop new approaches to digital economy policy.

An outcome of this meeting was a declaration which commits OECD members and partners to stimulate digital innovation and creativity through policies that "encourage availability and use of data, including open public sector data". While this declaration is not legally binding, it does provide an indication of the growing socio-economic relevance of open data and OECD members' commitment to it.

4. What is open data and why should you bother?

What is Open Data?

Data is information. For the purpose of this guide open data is the release of non-commercially sensitive and non-personal public sector information. Open data does not contain personal information relating to individuals or information which could be used to identify individuals. If you have any questions about dealing with personal information you should speak to the relevant Information Asset Owner in your organisation. You may also find it helpful to read the guidance issued by:

- the UK Information Commissioner
- the Scottish Information Commissioner about how to apply the personal data exemptions under FOISA and the EIRs.

Additionally, information which could cause economic harm if released is not within the scope of open data. There is no precise definition of 'non-commercially sensitive information', organisations will need to use discretion and balance the public interest of transparency against the right to confidentiality. The default position should be to release the information and you should not attempt to prevent its release unless there is a good reason.

You may find it helpful to read the Scottish Information Commissioner's guidance on responding to information requests as the same questions and principles apply to your open data.

- Commercial Interests and the Economy
- Confidentiality
- The Public Interest Test

5 Star Schema

Releasing your data isn't enough. There are other features which must exist if the information is to be considered open data. Open data should be:

- available at no cost to the user
- freely available to be used, redistributed and reused by anyone for any purpose, including commercial, without restriction. Aka, an open license
- available online in machine-readable formats
- easily discoverable through use of relevant metadata

> "Open data and content can be **freely used**, **modified**, **and shared** by **anyone** for **any purpose**"
>
> <div align="right">

Open Knowledge – The Open Definition
</div>

Tim Berners-Lee, founder of the world wide web, suggested a 5 Star Open Data Model which organisations can aspire to.

Summary of the 5 Star Open Data Model

★	Data available online with open license permitting re-use. Examples – Tables and charts in PDF document or scanned images
★★	Data available online in a machine readable format, with open license permitting re-use. Examples – Excel tables and charts
★★★	Data is available online, in non-proprietary machine readable format, with open license permitting re-use. Examples – Comma Separated Values (CSV) Extensible Mark-up Language (XML)
★★★★	Data is available online, in non-proprietary machine readable format, with open license permitting re-use. Data is described in a standard way and uses unique reference indicators, so that people can point to your data.
★★★★★	Data is available online, in non-proprietary machine readable format, with open license permitting re-use. Your data uses unique references and links to other data to provide context.

Under the strategy all public authorities in Scotland should be aiming to release all data in a 3 star format or above by 2017. In order to achieve this standard you should be building capability and capacity in your organisation now. Section 8 outlines the steps required to achieve 3 star release.

The 5 Star Schema model is an additive process. By this, we mean that the ambition to publish to a 3 star standard does not negate the need to also publish in 1 and 2 star formats. Many data users will continue to appreciate Excel and PDF document publication, and it is important to consult with your data users to determine which formats they will find most useful.

Why should you bother?

Uncertainty around the benefits and costs of open data often leads to organisations to ask why should we bother? There are many reasons why the public sector should be keen to release open data, both practical and ideological.

The volume of information available is increasing rapidly. Public sector organisations are large producers and collectors of information. As part of their public tasks, public sector organisations collect a wide range of non-commercially sensitive and non-personal data. This data is a valuable public resource, which in the past has been underused. Making the data available to the public helps realise the full potential of the data and creates many benefits, including:

- increased transparency and democratic accountability
- greater civic engagement
- improved efficiency and effectiveness of public services
- innovation and economic growth

UK Prescription Savings Worth Millions

Using publicly available prescription data, innovative start-up companies working with NHS doctors identified potential savings estimated to be worth approximately £200 million. The low cost project identified potentially huge savings in the prescription of statins, by doing simple analysis over a period of 8 weeks on publicly available data. Tools are now being developed to find savings in the prescriptions of other drugs, increasing the potential for significant savings.

Detailed analysis and results of the project can be found here: http://www.prescribinganalytics.com/

Showing the public how taxes are spent

Wheredoesmymoneygo.org is one of the many popular sites which have been built using publicly available data. Developed by Open Knowledge the site aims to show people, graphically, where public money in the UK is spent. The site always tracks historical spending so users can see where spending has risen or fallen and is a great example of open data being utilised to increase public transparency.

Open Knowledge hopes the information will "help citizens discover their own part in government economic activity — thereby encouraging them to take a more active interest in, and a more thoroughly informed engagement with, the official institutions around the UK".

Examples of how open data is benefitting the public sector and wider public in Scotland directly can be found in our case studies section.

Scotland's Environment Web - Land Information Search

SEPA, Forestry Commission Scotland and Scottish Natural Heritage have recently worked collaboratively under the umbrella of Scotland's Environment Web to develop the Land Information Search. The improved service, which collates many different datasets within one web portal, provides landowners and practitioners - such as farmers, moorland managers, developers and foresters - with a fast and convenient way to access a vast amount of information about their land and neighbouring areas. This includes native woodland surveys, forestry boundaries, Sites of Special Scientific Interest, historic sites and groundwater reports and much more.

Anyone looking to assess whether land that they manage is suitable for planting trees can now find out more easily thanks to the Land Information Search. The new 'one stop land information shop' harvests over 40 different open data sets providing landowners with detailed up-to-date information that will assist in the swifter preparation of high-quality applications for Forestry permits and Scottish Government grants

This is an excellent example of collaborative working between the Scottish Public Sector and demonstrates how making data openly and easily available benefits the wider community. In this case, making it easier for landowners to apply for forestry grants to plant and manage woodlands, providing value to their businesses, local communities and the environment.

A 2013 McKinsey report also recognised the potential of open data to generate wider economic growth. This was also highlighted in research conducted by Deloitte on behalf of the Department for Business, Innovation and Skills, which calculated the total social value and value to consumers, business and the public sector to be between £6.2 billion and £7.2 billion.

Cost of opening data

Open data uses existing internal data so the costs of preparing it for release should be low. However there will costs such as:

- web hosting and creation of portal
- promotion and advertising

- converting data into open formats
- time to update and maintain data
- time to promote open data both internally and externally

Costs will vary depending on the size of your organisation, your plans for open data and the level of open data maturity already existing in your organisation. The costs involved should not stop public authorities making their data open. In the vast majority of cases the data was captured or created using public funds and should be made accessible to all for re-use.

Open data is data which is available for free. This allows equal access to the data and allows it to be widely used and re-used. Any data which requires a fee to access cannot be considered true open data.

There are legislative exceptions which allow some public bodies to charge for their data in certain circumstances. If you are considering charging for your data, you should make sure you are entitled to do so under the existing access to information legislation.

Remember: Open data has the potential to help transform society, business and the public sector – why wouldn't you want to do it?

Useful links
8 Principles of Open Data
Open Data Handbook
Open Knowledge – What is open data?
Socrata – Why does my organization need open data?
The ODI – How will open data affect me?
The ODI – What is open data?
UK Government Open Data White Paper
The Local Government Association case studies

---

5. Overview of the Open Data Process
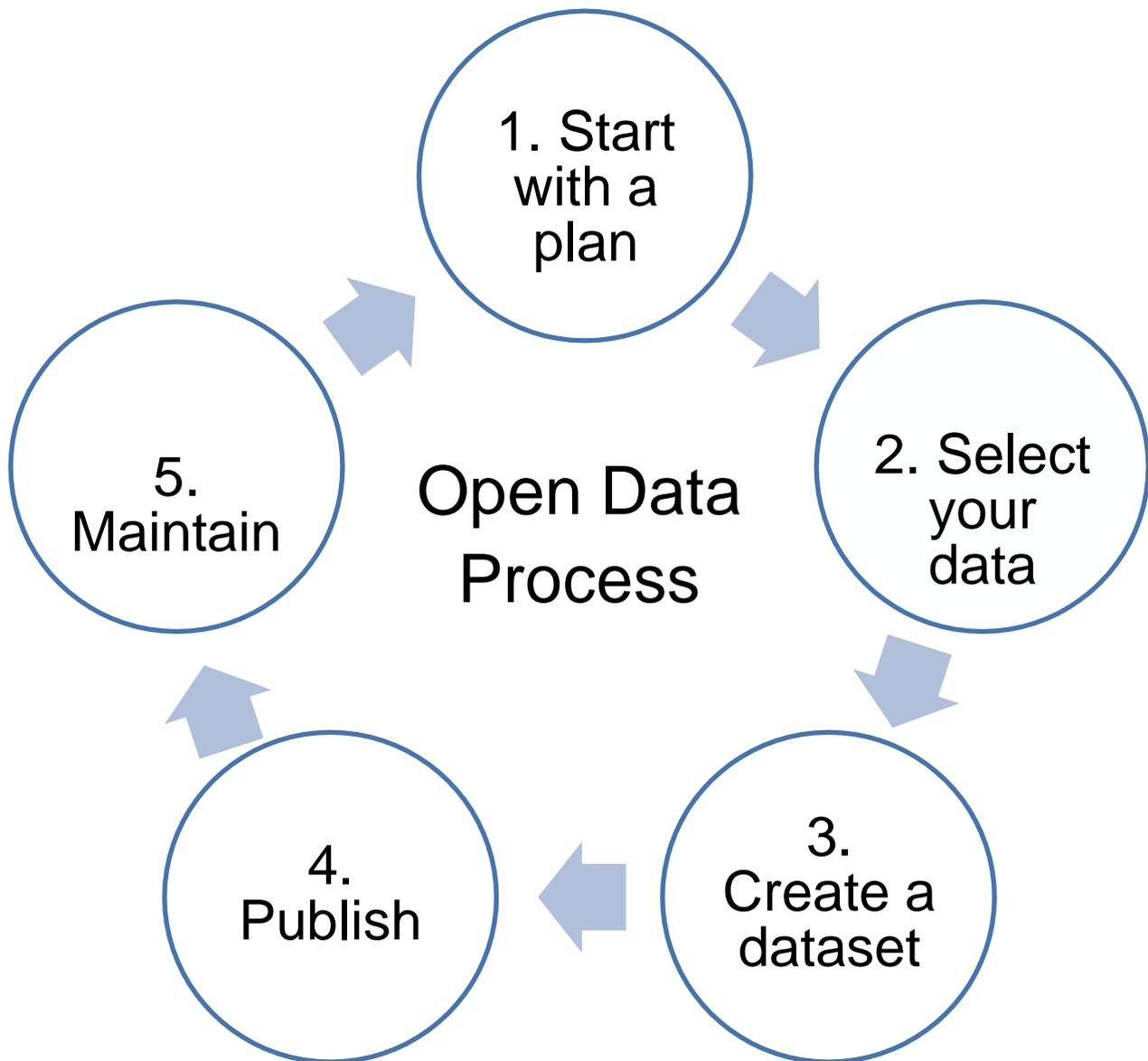
The open data process is not static. It is a continuous process that you will go through many times as your open data work develops. For ease the process has been described in 5 steps. In reality the process is not linear and you will likely do a few of these steps in combination.

Step 1: Start with a plan
Step 2: Select your data
Step 3: Create a dataset
Step 4: Publish
Step 5: Maintain

6. Start with a plan

Now you know what [open data is all about](#), you should be starting to ask how can your organisation get involved? To get the most out of any open data initiative, organisations must seek to embed the practice within its existing processes. The simplest way to begin establishing open data in your organisation is to design a plan.

This section provides practical guidance about how to create a successful open data plan in 4 steps. The order of steps is a guide and you could choose to do many of them simultaneously.

Have a goal	Assess current publishing and use	Get support	Demonstrate value
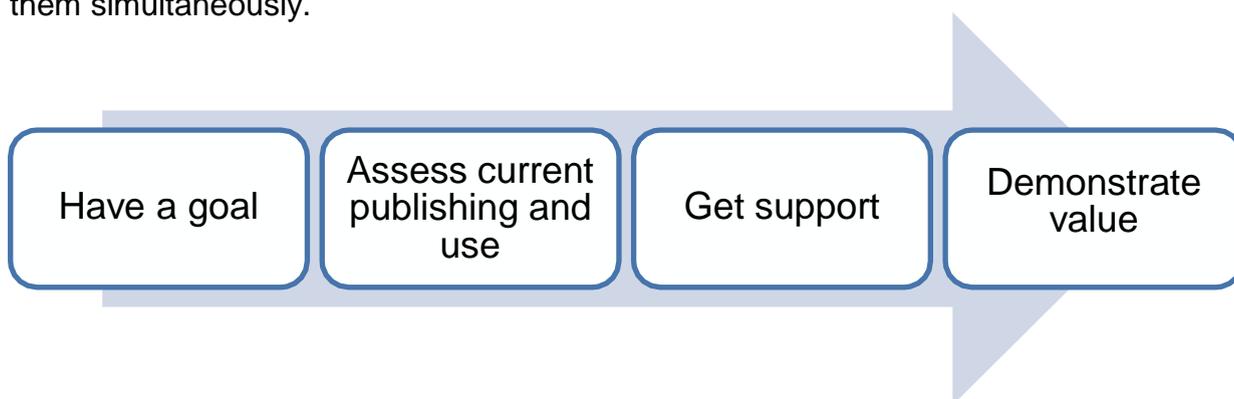
Step 1: Have a goal

Before launching your open data project, you should be clear about what you want to achieve. Your organisation will have its own strategic aims and goals, open data is another tool to support these. Does your organisation have a particular problem which open data could help solve or a target which open data could help achieve?

You will need to develop a [publication plan](#) which will require you to prioritise your data release. You may find it difficult to demonstrate value and get support. Having a clear goal for your open data and showing how opening data can help with organisational issues will encourage senior management support and can be the catalyst for establishing a longer term open data initiative.

Step 2: Assess your open data publishing and use

A useful tool to identify how your organisation currently uses and publishes data is the [ODI Pathways](#) self-assessment questionnaire. The questionnaire will help identify how engaged your organisation currently is with open data and will provide practical recommendations to improve your score. The tool could be used to help shape your open data goal as it helps identify gaps in your processes and areas for improvement.

Step 3: Get support

Feedback from colleagues across the public sector, has shown that for open data to be successful in the long term it requires support from across the organisation at all levels. Start with a small group of supporters who can then help you get support from others. Every organisation has a different structure and culture but generally you will want:

- **Senior management** – Having senior management support ensures your work is given weight. Senior management should be willing to help drive the project forward and help resolve disputes if they arise.

- **Open data champions** – These are people throughout the organisation who are interested in making the open data project a success. Open data champions will be the point of contact for questions and will help keep the wider organisation updated on progress of work.

- **IT and GIS specialist** – It would be wrong to make your open data project an IT only initiative, it requires support from all areas of the business. However, as open data will likely require working with new technologies and formats, their support will be essential.

- **FOI and/ or publication scheme contacts** – Colleagues in this area will have vital experience around dealing with requests, signposting people to information and an understanding of the FOISA/EIRs exemptions and exceptions.

Tips for getting support

- **Be specific** - explain how open data will help them specifically e.g. proactively releasing data may reduce FOI requests saving time and resources
- **Use case studies** – case studies help you show people why open data is worthwhile. Use the case studies in this resource pack to help persuade your stakeholders
- **Address fears** – be clear about the purpose of your project and address any concerns your stakeholders may have.

Step 4: Demonstrate value

Embedding open data within your organisation is an iterative process. You should continually evaluate your plan as your open data work progresses. Does your goal need to change because your plans have moved on? Are your aims still realistic and achievable?

You should also be looking to capture evidence which will show stakeholders how open data is bringing value to your organisation. A good way to demonstrate value would be to show how your project is helping support the initial goal. Demonstrating the success of the project can help persuade others to get on board and help embed open data into your organisations business practices. Other examples which may help demonstrate value:

- decrease in freedom of information/general requests
- efficiency savings e.g. decrease in processing times, financial savings
- greater public engagement with your organisations
- economic benefits e.g. development of apps and innovations using your data

---

EdinburghApps

EdinburghApps began as an annual once a year competition with the Council providing challenges and teams taking part over 6/7 weeks to develop strong concepts or/and prototypes which are then judged in a final event. EdinburghApps supports Edinburgh's open data strategy, challenges must all be supported by the sharing and release of data sets. The winners then have the opportunity to work with the Council to develop their ideas, and deliver products.

In two cases the Council helped winners to start their business from scratch, and also supporting participants to find other business opportunities.

Read the case study to find out more about the benefits to participants and the City of Edinburgh Council.

Useful Links

Code for America Open Data Guide
Socrata – Define clear and measurable goals
The ODI – How to make a business case for open data
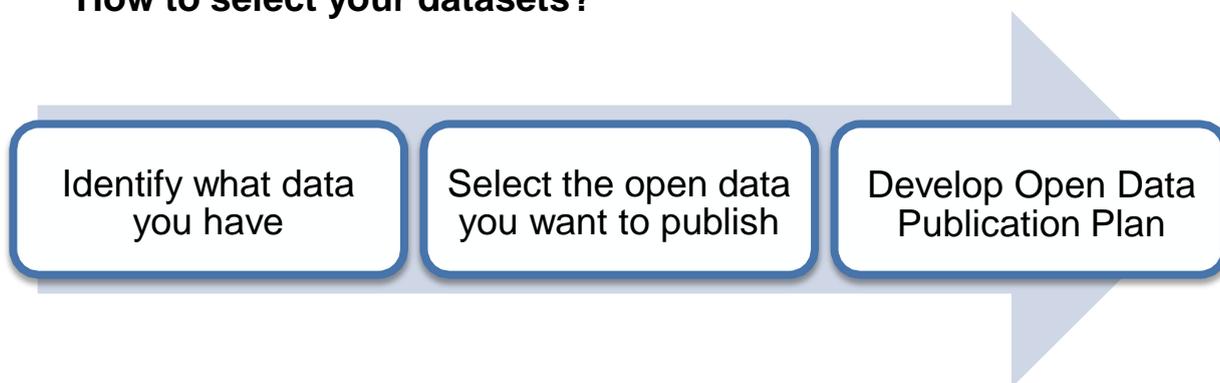W3C Best Practice – Organisational internal engagement

7. Select your data

How do you decide which data to publish first? Prioritisation of data release is necessary as it is impractical and potentially costly to release all your data at once. There is no definitive guidance on data prioritisation, there are many ways an organisation can choose to select its data depending on its goals.

This section will present a list of practical guidelines based on best practice from around the world. Annex A has simple downloadable checklist which has been developed to help you navigate each of the steps.

How to select your datasets?

| Identify what data you have | Select the open data you want to publish | Develop Open Data Publication Plan |

Step 1: Identify your data and create asset register

Before you know which datasets to release, you must identify what data you hold. If you do not know what data your organisation has, then you may miss out valuable data.

This may seem like a daunting task as your organisation will likely hold its data in various places and across multiple platforms, for example databases, spreadsheets, folders, documents and websites. Do not let the size and scale of the task put you off. This is an important step in your open data journey and will also be useful for other work related to Freedom of Information and Re-Use of Public Sector Information. Beginning with the identification of high level datasets and adding granularity over time will make the task more manageable.

It may be useful to ask colleagues across your organisation to help with this step, people are likely to have a good understanding of the data held within their own department or division.

Capture metadata

When you are identifying your data you should begin to capture metadata. Metadata is descriptive information about the data. It can describe elements such as the content, format, currency and limitations. More guidance on metadata can be found elsewhere in this pack.

At this stage you should attempt to capture as much metadata as possible as it will make things simpler in later stages. You should begin with what you would like to include in your asset register. The checklist in Annex A provides a short list of key metadata elements which you should begin capturing. For more on capturing metadata, please see Section 8.

Asset Register

You now have enough information to create a comprehensive list of all the data you hold. Your asset register will be used to create an Open Data Publication Plan which will inform the public about the data you hold and intend to release as open data.

Example asset registers

Department for Transport Information Asset Register

DCLG Data Inventory

Home Office Information Asset Register

This asset register does not need to be published and can be kept as an internal resource. However, it would be possible to combine an open data asset register with your organisations PSI asset register. The 2015 PSI Regulations require your organisation to publish a register of both published and unpublished information assets which fall within its public task. The potential open data that your organisation holds may fall out-with its public task. The PSI asset register could be extended to cover all of the data your organisation holds.

Your register is not static, the information you hold will change over time. Your asset register(s) should be reviewed and updated at regular intervals.

Useful Links

National Archives Asset List Guidance

National Archives Identifying Information Assets and Business Requirements

National Archives Information Asset Register Guidance

National Archives Public Sector Information Guidance

W3C Best Practice – Discover published information by site scraping

W3C Best Practice – Identifying what you already publish

Value Assessment

During the initial stages of identifying data and capturing metadata, you should make an initial assessment about the data's value and priority for release. An initial value assessment can help identify potential priority releases. Each organisation will assess the value of their data differently, depending on their priorities and quality of available data. The checklist in Annex A has a handy list of areas which should be considered in order to assess value.

Step 2: Select the open data you want to publish

When it comes to selecting data to publish, there is no right way. The important thing is to begin putting data out there. We recommend **starting small and building up**. Focusing on a few key datasets will help you create a maintainable publishing process. You should then add more data over time.

You will have to consider dataset prioritisation. Which datasets should you release first? If you have identified a few priority releases, should these be released together or separately? When prioritising your data you will begin shaping a plan for future releases. This plan or schedule will be helpful when compiling your Open Data Publication Plan.

There are number of easy ways to begin prioritising your data.

Start with your goal

You should return to the goals of your open data project and identify the datasets which support the realisation of those goals.

Quick wins

Sometimes an organisation may choose to release data which is easy to publish openly. Examples include upgrading data already published online (PDFs, Excel files, Word documents or other formats) into an open format. As this data is already released to the public, converting it to an open format should be easy and non-contentious. Another example would be publishing raw data alongside any analysis, if this is not done already.

Small, easy releases help get the project off the ground and build momentum, but organisations should be careful not to rely on easy releases for too long as the public may lose faith in the initiative if valuable datasets remain closed.

Demand driven release

Release the data that users want. All organisations have at their disposal a really **valuable asset** in the requests they get from the public, partner organisations, and other bodies for their data. These requests can be examined to see trends and consistency in type and extent of data required. This intelligence can then help inform the types of datasets that users want to see and the format and frequency that they want them released in.

Examine informal (e-mail/calls) and formal requests (FOISA requests) for data. Does your organisation have a twitter or Facebook page? Check the comments to see if there are suggestions about possible data you could release. By making the most requested datasets available in a discoverable, open format you can satisfy public demand and help reduce administrative burdens on departments e.g. fewer enquiries or requests.

Another way is to ask the data user what they want. As they are the people who will be using the data, they will likely have a good understanding about which data would be useful. Invite users to suggest ideas via social media, surveys or on your own website. Hack events are also another great way to generate interest in your open data and find out what people want or need to make their ideas a reality.

Perth and Kinross Council: Open Data Workshop

In October 2015, Perth and Perth and Kinross Council ran an "open data identification" workshop with community planning partners, regional organisations and council officers from a range of services. This ensured that a wide spectrum of individuals were able to give insight into which data sets would be most useful for Perth and Kinross to publish.

The datasets identified by the stakeholder workshop gave Perth and Kinross' team a solid basis from which to start creating a Publication Plan. You can read more about this in Annex B.

Scottish Government Dialogue App

Between 8th June and 14th July 2015 the Scottish Government held an open data discussion on the [Dialogue App](#). The Dialogue App is a crowdsourcing software designed for government. It allows the public to suggest, rate and comment on ideas in a collaborative way. The most popular and important ideas can then be easily identified and viewed.

As part of the Open Data Strategy, the Scottish Government made a commitment to engage with the public about which datasets they would like to see released from public sector organisations. The Dialogue App was chosen to hold the discussion as the format enabled everyone to participate in an open discussion.

Over the course of 5 weeks a total of 18 ideas were posted from 9 individuals. Nearly a quarter of the ideas submitted related to the release of information about public sector assets, physical (buildings, land) and non-physical (information, asset registers).

Follow best practice

Open data is growing and there are many public sector organisations both in Scotland and worldwide that are beginning to release their data openly. **Don't reinvent the wheel**, **copy what has worked** for others and **build upon their success**.

Cities and departments all over the world are beginning to release their own open data catalogues. Spend some time browsing their sites, see which datasets are popular and which ones your own organisation could release.

Examples of Open Data Portals		
Scottish Official Statistics	Leeds Data Mill	The City of Edinburgh Council
SEWeb	UK Data.gov.uk	EU Open Data Portal
New York Open Data	Open Glasgow	US Data.gov

This list is a very small snapshot of the portals available!

Public services are also reliant on each other for learning and sharing good practice. There are a number of places online you can go to find peers from the wider Scottish public sector to learn from and work with as you start thinking about your own open data plan. In order to facilitate a productive exchange of ideas, we have opened a Knowledge Hub group. Digital Public Services - Open Data Network is a knowledge exchange and collaboration space. There is also the existing Open Knowledge Scotland Group, and an active network of practitioners and forward thinkers on Twitter. Stakeholders have noted that the practicalities around opening data are difficult and by strengthening these existing networks organisations will help each other progress.

The G8 Open Data Charter, Open Data Barometer and the Open Data Census have all published works detailing what should be considered high value datasets and considered consider priority releases. Of course, some of the datasets may not be relevant to your organisation and you may not be ready to release them just yet, but it is a good starting point if you don't know where to begin.

The following table lists the 14 categories which the G8 considers high value, priority releases. Examples of the types of data which fall under each category are also listed.

**G8 High Value, Priority
Releases**

G8 Category	Example datasets
Companies	Company/business register
Crime and Justice	Crime statistics, safety
Earth observation	Meteorological/weather, agriculture, forestry, fishing, and hunting
Education	List of schools; performance of schools, digital skills
Energy and Environment	Pollution levels, energy consumption
Finance and contracts	Transaction spend, contracts let, call for tender, future tenders, local budget, national budget (planned and spent), international trade data
Geospatial	Topography, postcodes, national maps, local maps

Global Development	Aid, food security, extractives, land
Government Accountability and Democracy	Government contact points, election results (national and local), legislation and statutes, salaries (pay scales), hospitality/gifts
Health	Prescription data, performance data, doctor surgery locations
Science and Research	Genome data, research and educational activity, experiment results
Statistics	Data used to produce Official Statistics including the Census, sample surveys and administrative data. E.g. Datasets would include GDP, skills, unemployment
Social mobility and welfare	Housing, health insurance and unemployment benefits
Transport and Infrastructure	Public transport timetables, access points broadband penetration

Useful reading

Socrata – The data plan
W3C Best Practice – Discover published information by site scraping
W3C Best Practice – Identifying what you already publish
W3C Best Practice – Understand demand for data
Sunlight Foundation Open Data Guidelines 1 - 7

Step 3: Develop an Open Data Publication Plan

Once you have decided which data you want to publish as open data you should develop a publication plan. The benefit of an Open Data Publication Plan is the public will have a comprehensive list of the datasets you will be publishing open data and when they will be released.

The publication plan does not replace the publication scheme you are required to have under section 23 of FOISA. It should be part of your publication scheme which should:

- signpost your publication plan in your Guide to Information
- explain briefly how your open data will be published

Contact the Scottish Information Commissioner for more information about the Freedom of Information (Scotland) Act and publication schemes.

The publication plan shows the authority's commitment to open data and demonstrates its understanding of the benefits which releasing data openly can bring. As a guide, it is recommended that any Open Data Publication Plan should:

- tell users what information is available as open data
- explain when the information will be available, if it is not already
- tell users the currency of the data, available formats and licensing conditions
- provide contact details should someone want to get in touch about the dataset
- provide details about how users can make recommendations for future

An ambition of the Open Data Strategy was for all Scottish public authorities to have published their Open Data Publication Plans by December 2015. Given the range of public authorities in Scotland, it is recognised that not all will have been able to achieve this. Some have already published their plans, others are currently developing their own plans and some more are currently in conversation with their own governance structures around how to approach a publication plan. Annex A has a link to the template which has been designed to help organisations develop a relevant plan.

The template uses much of the information captured in the dataset asset register. The main difference between the asset register and the publication plan is that the publication plan will only identify the datasets that your organisation has released as open data, or intends to release as open data in the future.

Examples of organisations that have already published an Open Data Publication Plan include:

- Renfrewshire Council
- Scottish Natural Heritage
- Scottish Environment Protection Agency
- The Scottish Information Commissioner

8. Create a dataset

After selecting the information you wish to publish you need to organise it so it can be made available for download in bulk and in machine readable formats. This is called creating a dataset. Creating a dataset is a quick and easy process. A dataset is a structured presentation of data, such as a spreadsheet or table.

The steps to for creation of your dataset are set out below. Annex A has a [checklist](#) to help you make sure you cover each of the steps.

| Apply an open format | Capture metadata | Apply an open license | Review datasets |
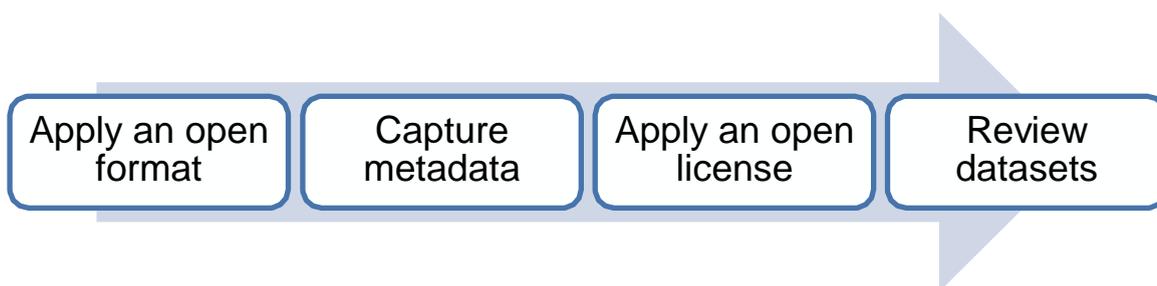
Step 1: Apply an open format

One of the most common questions asked is 'what format should I use?' Open data should be in an open format and machine readable.

- **Open Formats** are non-proprietary and platform independent. They can be accessed by anyone and do not require access to licensed software. E.g. Microsoft formats are not open as they use proprietary software.

- **Machine Readable** formats allow a computer to read the data. Machine readable data is structured and easy to query using code.

The most appropriate format will depend on the type of data. Any type of data can be stored in an open format, but it is likely you will have to transform the data from its original format. Open, machine readable formats allow the data to be used and edited easily. It also allows for interoperability between different datasets. For example, a PDF publication may look nice but it severely limits the user's ability to re-use the information.

You should be aiming to select a format which satisfies 3 star publication requirements. Below is a table of common open data formats which satisfy 3 star release.

Examples of Common Open Formats

Format Name	Definition	Type of data to use this for
Comma Separated Values (CSV)	Comma Separated Values (CSV) is a great way of storing large amounts of data with just commas separating the data values. Often the CSV file will contain a header with names describing what data is populating the file.	Tabular data e.g. Use instead of Excel
Tab-Separated Values (TSV)	TSV is a very common form of text file format for sharing tabular data and is highly machine readable.	Tabular data Use instead of Excel
JavaScript Object Notation (JSON)	JSON uses human-readable text to transmit data objects consisting of attribute–value pairs. It is used primarily to transmit data between a server and web application, as an alternative to XML. The file size will be more compact or smaller than XML.	Complex structured data Multidimensional data Tabular
Extensible Markup Language (XML)	XML is a widely known markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. Users create and define their own tags.	Complex Structured data Multidimensional data Tabular data
Rich Site Summary (RSS)	RSS (originally RDF Site Summary), often dubbed Really Simple Syndication, uses a family of standard web feed formats to publish frequently updated information: blog entries, news headlines, audio, video. An RSS document (called "feed", "web feed" or "channel") includes full or summarised text, and metadata, like publishing date and author's name.	Use for announcements or events e.g. on websites
ATOM	The Atom Syndication Format is an XML language used for web feeds. The Atom format was developed as an alternative to RSS. Note RSS is the preferred standard.	Use for announcements or events e.g. on websites

Open Document Format for Office Applications (ODF)	The Open Document Format for Office Applications (ODF), also known as OpenDocument, is an XML-based file format for spreadsheets, charts, presentations and word processing documents. It was developed with the aim of providing an open XML-based file format specification for office applications.	Non-system generated metadata or additional information you release with your dataset. (replaces Excel, Word, PDF)
HTML	Used for formatting information on the web	Non-system generated metadata or additional information you release (replaces PDF, Word)
Keyhole Markup Language (KML)	KML is an XML language focused on geographic visualization, including annotation of maps and images.	Spatial/location data
Geography Markup Language (GML)	GML is the XML grammar defined by the Open Geospatial Consortium (OGC) to express geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the internet.	Spatial/location data
GeoJson	GeoJSON is an open standard format for encoding collections of simple geographical features along with their non-spatial attributes using JavaScript Object Notation.	Spatial/location data

Table taken from Government of South Australia Open Data Process Guide

Useful reading
Government Service Design Manual – Choosing appropriate formats
Open Knowledge - Open Format Definition
Sunlight Foundation Open Data Guidelines
W3C Best Practice – Make the data available in a language people want

Step 2: Capture Metadata

Your data can only be used effectively if you also provide some metadata. Metadata is descriptive information about the data. It can describe the content, format, currency, limitations and frequency of updates. Metadata provides the user context about the data and good metadata will allow interoperability with other datasets.

Your publishing portal may allow metadata to be displayed below your data, or you could create an accompanying file for the metadata.

What standard should be used

| Internal standard | Dublin core | DCAT |
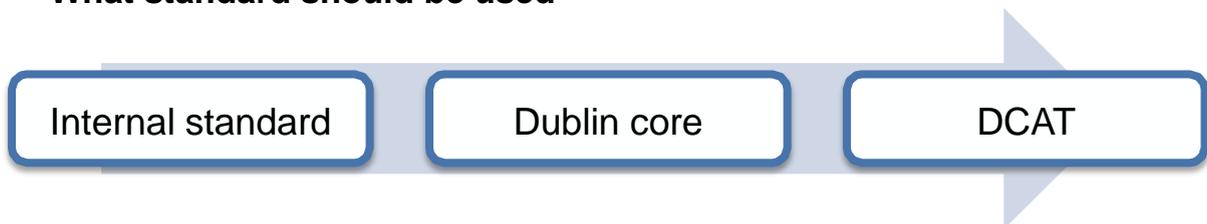
Suggested progression

Over time it is expected that all public authorities will progress towards the Data Catalog Vocabulary (DCAT) standard. DCAT will be used to describe all public data in Europe. Its use will make public data searchable across borders and sectors thus enabling discoverability by automated systems including aggregators and search engines. Progressing towards DCAT is an ambitious aim. DCAT is a high standard and captures much more than your organisation may have considered. It is recommended that you build your metadata catalogue slowly, embedding it within processes and ensuring that the metadata is recorded consistently. The use of intermediate standards, such as Dublin Core, is recommended. Dublin Core provides a robust metadata standard which can then be built upon as your organisation progress towards DCAT.

Current metadata themes

Feedback has indicated that public authorities are capturing limited metadata and that it is often being captured in an ad-hoc manner. In order to build capability and increase metadata maturity, you should decide within your organisation what metadata should be captured and begin to record it in a consistent manner for all data.

It is useful to consider embedding metadata collection into your data governance processes from the outset. Increased open data publication should be a long term organisational goal and there is a benefit in adopting consistent metadata standards early to facilitate further re-use and demand for publication.

As sharing data and making it open becomes the norm, the adoption of metadata use is also expected to grow.

Marine Scotland: Marine Portal Metadata Collection

Marine Scotland have now mandated the adoption of MEDIN Discovery Metadata Standard. This means that any dataset to be published on the Marine Scotland portal must meet this standard prior to publication. MEDIN satisfies and exceeds the suggested Dublin Core standard.

While Marine Scotland acknowledge that this requirement can occasionally slow down dataset publication, they are confident that this approach offers significant benefits. In particular, it ensures that consistent metadata collection will become a routine business practice.

You can read more about the ongoing work in Marine Scotland in Annex B.

Useful Reading

DCMI Metadata Basics

Dublin Core Elements

NISO Understanding Metadata

The ODI – Marking up your dataset with DCAT

W3C Best Practice - Metadata

Step 3: Apply an open license

A key requirement of making data open is applying an open licence. Prior to publication, all datasets should have an open license. The following applies to data and information which you hold the intellectual property rights for, if the data you wish to release includes third party IP rights please read the [Licensing and Third Party rights section.](#)

Why is licensing important?

Licensing data is essential to provide potential users with clarity and certainty. When you create something, original works or photographs for example, you automatically obtain rights over the work and can determine how the work is used. Applying a licence to your work explicitly tells users what they can and cannot do with it.

Applying an open licence to your content or data should allow people and organisation to re-use, modify and share content in any way. It should allow others to use the data for commercial purposes. It is generally accepted that only two restrictions may be attached to an open licence:

- **attribution** – users must acknowledge the source of the data
- **share-alike** – users must publish any derived data under the same licence

Open licenses can have no restrictions (public domain – all rights waived), attribution or attribution and share-alike.

How to select a licence

The chosen licence should support your organisations open data strategy. You need to think about what you want to achieve by releasing your data. Requiring attribution will normally help promote your open data initiative as users have to link back to your original work. Share-alike restrictions will require users of the data to publish their work openly. This may deter commercial businesses and people who want to make profit from their use of the data, resulting in reduced innovation and use.

Whilst possible to create your own unique licence, it is advisable to use a standard re-usable licence as they provide greater recognition amongst users, increased interoperability due to the use of standard terms and increased user compliance.

There are two instances when you cannot choose your own licence –

- Crown Bodies - if your organisation is a [Crown Body](#), which covers most government departments and arms-length bodies, then any information you have gathered or created is owned by the Crown. This information must be published under the Open Government Licence.

- Publishing data that has been derived from data published under a share-alike licence. You must publish that data under the same licence as the original data.

Open Government Licence 3.0

The [Open Government Licence 3.0](#) (OGL) allows anyone to publish, distribute, transmit and adapt the licensed work, and to exploit it both commercially and non-commercially. The user must acknowledge the source of the work and where possible provide a link to the OGL. The OGL was developed to be used by public sector bodies.

There can be no charge for data licensed under the OGL. The OGL is compatible with the latest versions of Creative Commons Attribution Licence (CC-by) and the Open Data Commons Attribution Licence (ODC-by).

Other popular licences

Open Knowledge provides an [extensive list](#) of the licences which conform to the open definition. The most popular are:

Level of Licence	Creative Commons Licence	Open Data Commons Licence
Public domain (all rights waived)	CC0	PDDL
Attribution	CC-by	ODC-by
Attribution & share-alike	CC-by-sa	ODbL

How to write your attribution statement

If your license requires attribution, you must state how your work should be attributed. As your work may be combined with others who also require attribution you should keep the statement to a minimum. For example, your organisation's link to the data that is covered by the licence and link to licence.

You can also prescribe how the attribution should be presented (size, location, format etc.). You will need to consider the users of the data and make sure any requirements are not too onerous.

Successfully apply your licence

You must signpost users to your licence by using both human-readable and machine-read-able descriptions. Your descriptions should be displayed prominently with your data so that users know they can use the data you are licensing.

The common standard licences – OGL, CC and ODC, all provide machine and human readable descriptions and logos that you should use.

- [Open Government Licence](#)
- [Creative Commons licence chooser](#)
- [Open Data Commons licences](#)

Licensing and Third Party Rights

You can only apply a licence to data which:

- you own the copyright/and or database right for;
- or the owner has given permission for it to be licensed.

If you do not own the intellectual property or do not have the owner's authority, you cannot release the data openly.

Public sector organisations engage with and contract with many third parties in the course of its daily activities. Many of those contracts will grant third party rights. It would be an inefficient and costly use of public resource if all of those contracts were to be renegotiated. In the future we want to limit the existence of third party rights in public data. We expect when future contracts are negotiated and put out to tender that it will be made explicit in the contract that any data resulting from the contract will be subject to open data principles and may be release for free to the public for onward use.

If you require further guidance on any matters relating to third party rights you should speak to your organisation's legal department.

Useful Reading

[The ODI Publishers Guide to Open Data Licensing](#)
[Licensing Open Data: A Practical Guide](#)
[National Archives Guidance](#)

Step 4 - Review datasets

Every dataset will vary in completeness and quality, before releasing the data you should strive to ensure the data is as complete, accurate and up to date as possible. However, there is no such thing as "perfect" data. Scotland's Open Data Strategy emphasises both quality and quantity of data.

Imperfections should not deter you from releasing data. When you publish your datasets, be explicit about any limitations and add caveats which will help any re-user understand the limitations of the data. The clearer you are about the limitations, the more usable your data will be as re-users will have a greater understanding about what the data represents. Additionally, re-users will provide feedback on data quality and mistakes, which will help improve the quality of your data.

ODI certificates are a great tool to assess how open your dataset is. The tool also provides tips and information about how to improve the openness of your dataset.

9. Make your Data Available

The next step is to publish your data. This section will take you through the steps you should consider when making your data available to others.

Choose where to publish

Existing website

You can choose to use your existing website to host your open data. This may be a time and cost efficient option when you have limited datasets available. Your website team will already be familiar with making other data files available for download.

Depending on the layout of your site and the search facilities available, it may be difficult for outsiders to find your open data. This limits the effectiveness and reach of your open data. Thought would have to be given to whether the layout of your website would need to be altered to enhance the discoverability of your open data. You could have some really interesting datasets available, but if people cannot find them, then they cannot use them.

Data Catalogues or Portals

It is becoming the de-facto standard for open data portals or open data catalogues to host an entire organisation's open data. These are essentially websites where the host can upload and update datasets and the public can search for and download datasets.

If you go down this route, your options will be:

- build your own platform in-house
- buy an existing platform
- use an existing open source platform

Where possible you should resist the urge to invest money and resources to build a system from scratch if there is a suitable existing model. For example, CKAN is the most popular open source platform in the UK and has been used successfully by many organisations. It is the UK Government's chosen platform for data.gov.uk. DKAN is a drupal based open source system that was based on CKAN's key features. Both platforms allow for the easy publication and visualisation of data.

Statistics.gov.scot

In late 2013 the Office of the Chief Statistician and Strategic Analysis (OCSSA) commissioned a consultancy report and a pilot website to demonstrate how open data could be applied in practice to improve access to Scottish Government statistical data. In 2014 OCSSA commissioned the development of an open data platform to replace the Scottish Neighbourhood Statistics website that was used to disseminate small area statistics. A Beta version of the platform was launched in June 2015 for testing and to seek public feedback to inform the development.

In February 2016 the Scottish Government launched an open data platform at www.statistics.gov.scot to publish the data behind Scottish Official Statistics so that it is discoverable, accessible and reusable. It enables users of official statistics in Scotland to access data covering all areas of life in Scotland at national and local level – 100 million statistical observations in total. This data is published in open formats that exceed the ambition of the Open Data Strategy. During 2016 and early 2017 OCSSA will promote and market the site through social media and events, and engage with users to make enhancements, add new functionality and publish more data.

Regardless of the route chosen it is essential that your platform:

- allows authorised users to make uploads and updates
- supports your chosen metadata standard
- supports the uploading and downloading of your data in bulk
- supports common open formats
- allows the public to search metadata and download datasets

When choosing your option other factors you should consider are:

- **Total cost** – not just initial cost of building or buying system, but on-going costs such as hosting support, maintenance and update costs

- **Satisfaction with features** – in addition to the essentials your organisation may value other tools such as data visualisation tool and API compatibility

- **Sustainability of option** – is there enough support? Is there enough in-house resource or will you need to contract out? Will you need to invest in training?

Examples of Open Data Portals		
Statistics.gov.scot	Leeds Data Mill	The City of Edinburgh Council
New York Open Data	US Data.gov	EU Open Data Portal
SEWeb	Open Glasgow	UK Data.gov.uk

This list is a very small snapshot of the portals available!

Useful Reading

Open Data Handbook – How to open up data
Sunlight Foundation Open Data Guidelines
Open Government Data Toolkit – Technology options

10. You've published, now what?

You have published open data, congratulations! This should not be considered the end of the journey – improving your data and ensuring that it reaches the widest audience are important to maximise the opportunities presented.

Improve your data

Continuous improvement should be a key feature of your open data initiative. As well as acting upon external feedback about your data, you should also have an internal review to see what areas you can improve.

ODI certificates are a great tool to assess your published data. They ask useful questions and give you clear targets to aim for in order to improve your data.

Promotion

Tell everybody! Let people know that you have opened up some datasets. Consider your audience. If there is a certain group of people you want to target, what would be the best way to get the message to them? A few simple promotion ideas are:

- press release
- announcement on website
- inclusion in your Guide to Information
- social media promotion
- post on third party sites - contact popular sites/blogs with an interest in this area and offer to write an article or post
- contact leading organisations who have an interest/work in this area – they will likely be happy to spread the word
- use any relevant internal mailing lists/contacts to directly contact people

Generate interest

Making people aware your open data exists is an essential step, but it isn't enough. You need to encourage them to use your data. It is worthwhile investing time and money in this stage because it means success: people actually using your datasets is the outcome you should be aiming for. It also brings may other benefits including:

- Greater civic participation in your organisation – your relationship with your customers and the wider public will become more open

- Potential positive impact on society and economic benefits if a new tool or app is created from the data

- Efficiencies – time and cost savings. Users will also be happy to give feedback to help improve data quality.

- You will know what to release next – as users start to understand what data you have available and what they need, they will begin requesting more datasets

Hack Events

Hackdays are a great way to engage with the community and bring together developers, open data enthusiasts, the local community and your organisation. These events allow developers and citizens to work creatively to solve issues. Promotion of your event is key to generate excitement and interest. Apps developed at hackdays can be very successful (see EdinburghApps case study), though the main aim should be to increase interest and participation.

Socrata has developed a comprehensive guide to planning a hackathon

SEWeb Hackathon

SEWeb organised a hackathon event. Students from universities throughout Scotland were invited to come up with fresh new innovative ideas to make better use of available data, and to collect new local environmental data that can help further our understanding, and encourage people to get interested and get involved in Scotland's Environment. Feedback from all who attended was overwhelmingly positive and the next step is to see how ideas can be supported beyond the prototype stage.

More information about the event can be found in the case studies section.

Competitions

Competitions give developers a slightly longer time frame to showcase what can be developed from your data. A prize will normally be awarded to the winner.

> **NYC Big Apps**
>
> Over five months, BigApps challenges developers, designers, and entrepreneurs to create functioning, marketable technology tools that help solve pressing civic challenges. In 2015 the 4 key areas are affordable housing, zero waste, connecting cities and civic engagement. Over $125,000 in cash prizes and product development support is available to finalists and winning teams. More information can be found on BigApps.

Conferences and Meet-Ups

Face-to-face events are a great way to encourage use of data. Conferences are a traditional way of engaging with others where you set the agenda and arrange presentations on set topics. It is a great way to show what you have been doing and share your future plans and developments. Due to the topic and audience you are trying to attract, it is best you try and include interactive elements where possible. This will allow you to find out more about them, discover interest level and enable re- users to connect with each other.

Instead of the traditional conference, you could host a more informal, participant driven event. A meet-up could allow participants to form the agenda, give presentations and lead sessions on their areas of interest.

Any event will increase exposure and potentially increase likely users.

> Renfrewshire Open Data 'Enlightenment' Event
>
> In January 2016, Renfrewshire's Community Planning Partnership held its first Open Data event titled 'Enlightenment'. Its aim was to introduce partners, councils and council staff to open data and encourage them to participate in influencing the vision of Open Data in Renfrewshire.
>
> Guest speakers were invited from a variety of organisations, to provide their perspective on the open data journey. Break-out sessions were then conducted where the audience divided into groups and considered three questions:
> - How can Open Data be of use in Renfrewshire
> - What datasets would your organisation like to have access to?
> - How can your organisation play a part in Renfrewshire's Open Data future?
>
> Top priorities were identified for each question, and a number of attendees volunteered to identify candidate open datasets within their own service area or organisation. It also led to the creation of a Renfrewshire Open Data Advocates Network. More information about Renfrewshire's event can be found here.

Develop an App

Get creative and make something yourself! One of the best ways to show what your data can do is for you to make something yourself. If you have the time and technical resources you could develop a user-friendly app which addresses an issue in your organisation or community, or even just demonstrates what your data can do.

Clacks Kids

As part of the Open Data Scotland programme run by Nesta and Code for Europe, Clackmannishire Council developed an app, Clacks Kids, which would help parents and carers access personalised childcare resources. The app has been developed to support the councils Making Clackmannanshire Better Programme.

More information of Clackmannanshire Councils experience developing the app can be found in our case study section.

Run the City App

Run the City is a guided tour for runners and winner of the 'wild card' challenge for **EdinburghApps** 2014. Runners will always get their run in, even when away on business, but running in a strange city is difficult when you don't know where to go. Run the City solves this challenge as the app, through audio messages, not only gives runners directions but also highlights their attention to the city sights and makes their run in Edinburgh more engaging with anecdotes about the areas they are running through.

It utilises the Council's open data as content for the app and will also create data we can make open. More information can be found in the case study.

Make it sustainable

Holding hack events and competitions are ideal in the short term, but are not sustainable. You need to make open data become the norm in your organisation. To do this you need to appoint someone, whether existing or additional, to be responsible for your open data efforts. You will then require others to support this role by being responsible for open data in their own division or departments. This does not need to be a new person, but could be delegated to existing employees.

These individuals should work together to co-ordinate open data efforts in your organisation, publish data, engage with public and address feedback.

Helsinki is a great example of how open data can transform public services for the better.

Useful Reading

Code for America Planning for sustainability

Open Data Handbook

5 Stars of Open Data Engagement

Open Government Data Toolkit – Demand and engagement

The ODI – Engaging with reusers

---

11. Training and Resources

To help get you started we have pulled together all of the useful resources referenced throughout this pack, plus a few extra.

Scottish Government Training

In July 2015 the Scottish Government ran a procurement process for the provision of open data training across Scotland. Urban Tide were selected as the preferred training partners to progress this work with the Scottish public sector. The format of the training has been in both one day and two day workshops:

- The one day workshop was designed with the needs of public sector leaders, senior managers and data owners in mind. It provides an introduction to Open Data and explores the potential benefits to an organisation
- The two day workshop extends the learning of the one day session to help an organisation implement the principles and practice of Open Data in line with the Open Data Strategy. This is designed with the needs of Open Data practitioners in mind, and includes detailed knowledge on developing Open Data publication processes and standards.

The training will run until end September 2016 and feedback received has been very positive. There are still some sessions available, and availability can be found on Eventbrite. Participants have particularly welcomed the opportunity to share knowledge and exchange ideas with other attendees.

SEPA - Practical use of the Resource Pack

SEPA, following guidance in the Resource Pack, is looking at the demand for SEPA data:
- The data needs to be made easily accessible to the public using SEPA's business-as-usual processes
- The user's needs have to be understood so that effort is concentrated in the right place and a priority list of datasets is created
- The user's needs should be balanced against the needs of SEPA, creating a publication plan which produces tangible benefits for all parties
- The demand for SEPA data comes through a number of channels – data and diligence requests, FOI/EIR requests, science helpdesk calls and calls and emails to Floodline and the SEPA Contact Centre

- By analysing the requests that come in through these channels and the methods employed in servicing them, SEPA will be able to **identify the most requested data** and construct efficient business processes to deliver them resulting in a system which gives the user immediate access to the data he needs while alleviating demand on SEPA staff time – a win-win solution
- An analysis of data and diligence requests for Q1/2016 showed the type of information that was asked for and the amount of staff time and resources that is being consumed

To this end, SEPA designed a survey that they sent to relevant contacts to ask for feedback.

SEPA hope that this work will enable them to create a **SEPA Data Service** that will allow users to serve themselves from a SEPA Data Download and discovery site. This Service should give a better experience to our users but also reduce the work required by SEPA staff to process requests for data and allow them to concentrate their resources on making more data open.

The James Hutton Institute – Practical Use of the Training

Staff from the James Hutton Institute attended the two day open data training workshops. Open data was recognised as a major opportunity for The James Hutton Institute - particularly in terms of organisational profile and enabling access to research - and the workshops were considered timely. Prior to the sessions, colleagues were aware of open data and aware that it was gaining momentum. However, they also acknowledged there was a steep learning curve to this process.

The workshops provided James Hutton Institute colleagues with an opportunity to hear from other organisations and to learn that similar challenges were encountered across the public sector. In the words of one participant, it allowed them to realise that 'we're all singing from the same hymn sheet'. It also enabled them to recognise areas of best practice in their own organisation.

James Hutton Institute colleagues felt the training equipped them with an understanding that data is an asset and it helped set the scene in terms of where their own organisation would go. They are now in the process of preparing a paper for their senior management team, to ensure open data plans gain senior management buy-in.

Online Resources

The open data guides cover most elements of open data and you may find it helpful to refer to them first before using more specific guidance.

Open data guides – For those new to open data
Open Data Handbook
Open Data White Paper
8 Principles of Open Data
Open data guides – Further information
Open Data Field Guide
Open data value framework
Open Government Toolkit
Open Data Playbook
W3C Best Practice
Asset register
National Archives Asset List Guidance
National Archives Identifying Information Assets and Business Requirements
National Archives Information Asset Register Guidance
National Archives Public Sector Information Guidance
Engagement
How to run a hackathon
How to run datapalooza or data jam
5 Stars of Open Data Engagement
The ODI – Engaging with reusers
Formats
5 Star Schema
Choosing formats
Licensing
Guide to Open Data Licensing
Publishers Guide to Open Data Licensing
The Open Government License
Open Data Commons Licenses
National Archives Guidance
Creative Commons licence chooser
Licensing Open Data: A Practical Guide
Metadata
The ODI – Marking up your dataset with DCAT
Data Catalog Vocabulary (DCAT)
Dublin Core Elements
DCMI Metadata Basics NISO Understanding Metadata
Selecting Data
Socrata Top Datasets

Annex A Templates and Checklists

3 supporting resources have been developed to help authorities develop and implement their own open data plans.

Selecting Data Checklist – can be used to make sure that all of the steps for selecting data are covered.

Open Data Publication Plan Template – simple template designed to help authorities list what open data they plan to release and when, along with some other key metadata.

Creating a Dataset Checklist – can be used to make sure that all key steps are considered when creating a dataset.

[These resources are all available here.](#)

Annex B Identifiers

Although organisations are only advised to aim for 3 star data format as their initial goal in open data publishing, using identifiers in the data can make it easier to move towards 5 star data. In order to see what is involved, it's helpful to first look more closely at the kind of information that is typically expressed in tabular data. As an example, consider the number of births in Glasgow City during a given year. In a spreadsheet, we might see something like this:

Year	Area	Number of births
2014	Glasgow	7,311

This row of data can be read as a statement: "In 2014, the number of births in Glasgow was 7,311". In order to make a clear statement, we need to be able to refer to things (or more generally, to resources) in a non-ambiguous way. But the term 'Glasgow' *is* ambiguous. We could be using it to refer to Glasgow City (pop. about 600,000) or Greater Glasgow (pop. about 1.2 million). Of course, a careful data publisher could try to consistently distinguish between these two terms, but it is hard to eliminate the risk of confusion.

In the context of open data, terms for referring unambiguously to things (both concrete and abstract) are called unique identifiers. The most common approach uses Uniform Resource Identifiers (URIs for short). These look just like familiar web addresses, but are intended primarily for naming resources rather than as web pages to visit in a browser. An example URI is http://statistics.gov.scot/id/statistical-geography/S12000046, which in fact is a unique identifier for Glasgow City. Although the standard framework of 5 star data goes beyond the limitations of tabular data, it is perfectly possible to use URIs in tabular data. For example, we could produce a data record like this:

Year	Area URI	Area	Number of births
2014	http://statistics.gov.scot/id/statistical-geography/S12000046	Glasgow	7,311

This preserves the informal label for Glasgow while also including a URI

URIs are useful as a formal mechanism to clarify what a data record is talking about. They are useful in another, perhaps more important way, in that they give a precise way for two different datasets to give information about a

common set of resources. For example, dataset A could contain a wide range of administrative data about Glasgow City, while dataset B contains information about health resources, including maternity wards and antenatal classes. As long as the two datasets use the same identifier for Glasgow City, it becomes much easier to automatically combine them so as to derive a much richer and more complete set of statements.

URIs are designed to be processed by machine rather than by humans. Although a URI does not have to correspond to a page that can be viewed in a web browser, it is best practice to have a way of getting from a URI to a human-readable web page, so that any questions about usage and interpretation can be clarified. This step is usually accomplished by something called Content Negotiation, which redirects a web browser from a URI to an associated page; for example, if you point your browser to the URI

- http://statistics.gov.scot/id/statistical-geography/S12000046

you will automatically get redirected to

- http://statistics.gov.scot/doc/statistical-geography/S12000046

These addresses look exactly the same except that 'id' (for identifier) in the first has been replaced by 'doc' (for document) in the second. The second address takes you to a web page with lots of useful information about the resource.

Unfortunately, there is another level of complexity in using URIs. A URI is unambiguous because it names only one resource. However, there might be a second URI (or indeed many) which also names that resource. As an example, http://dbpedia.org/resource/Glasgow is the URI for Glasgow City provided by DBpedia (and http://dbpedia.org/page/Glasgow is the associated human-readable page). This multiplication of synonymous URIs can be dealt with by adding a piece of data to your datasets which says that two URIs refer to the same thing.

In summary, including URIs in tabular data is a useful step in increasing the interoperability between multiple datasets and allowing more useful inferences to be drawn from the data. Identifying the right URIs to use, however, is not always straightforward. Although best practice is to re-use existing identifiers, it requires a mixture of experience, effort and luck to search these out. As the open data ecosystem in Scotland evolves, it should become increasingly possible to establish standards for public sector URIs.

Useful Reading

[Statistics.gov.scot – Linked Data Vocabularies](#)
[Cabinet Office - Persistent Resolvable Identifiers](#)
[Chief Technology Officer Council – Designing URI Sets for the UK Public Sector](#)
[Thomson Reuters Lab – Unlock Your Data](#)
[Phil Archer – Study on Persistent URIs](#)

Annex C Case Studies

The following is a collection of case studies which demonstrate the value that open data is bringing to individuals, companies and public authorities in Scotland.

1. [Clackmannanshire Council: Open Data Scotland and Code for Europe](#)
2. [Crichton Institute: Regional Observatory](#)
3. [The City of Edinburgh Council: ARC-E App](#)
4. [The City of Edinburgh Council: EdinburghApps](#)
5. [The City of Edinburgh Council: Run the City App](#)
6. [Royal Commission on the Ancient and Historical Monuments of Scotland – SENESCHAL](#)
7. [Registers of Scotland: Cadastral data for the INSPIRE directive](#)
8. [Scotland's Environmental Web: Ecohack](#)
9. [Perth and Kinross Council: Open Data Workshop](#)
10. [NHS National Services Scotland: ISD Prescribing Data](#)
11. [Marine Laboratory: Marine Data Portal](#)
12. [Aberdeen City Council: Leisure App](#)

If you are interested in case studies which cover a particular theme:

Theme	Case Study
Improving public services through the use of open data	▪ 1. Clackmannanshire ▪ 3. Edinburgh Council ▪ 12. Aberdeen Council
Increasing transparency and accountability through the use of open data	▪ 6. RCAHMS ▪ 11. Marine Laboratory
Generating improved civic engagement through the use of open data	▪ 2. Crichton Institute
Engaging data users in the process	▪ 8. SEWeb: Ecohack ▪ 9. Perth and Kinross
Building a business case around open data	▪ 10. NHS NSS: ISD
Improving data governance through the use of open data	▪ 7. Registers of Scotland
Generating innovation and economic growth through the use of open data	▪ 4. Edinburgh Council ▪ 5. Edinburgh Council

If you have a case study you would like to share or you would like to be put in touch with the case study subjects, then get in touch – [OpenDataPolicy@gov.scot.](mailto:OpenDataPolicy@gov.scot)

# 1. Clackmannanshire Council: Open Data Scotland and Code for Europe

Many of the problems which Open Data is typically used to solve don't exist in a small Council. Mass transportation isn't an issue with only 3 bus routes. There isn't a developer community taking part in hack events and generating innovative applications. Why then should a small Council pursue Open Data?

Open Data affords opportunities to be more efficient, whether through being nimble by adopting freely available Civic Apps to improve service delivery or by reducing the time spent responding to information requests from the public or partners. In time, it is likely that we will be required by statute to share more data anyway.

Location based services will become increasingly important. In the near future citizens will expect to be able to use their personal device and using the tools of their choice, see and interact with services which are nearby. In order for Council services to be part of this world, data about those services must be published openly.

Background

'Open Data Scotland' is a programme which has involved over the last year, four of Scotland's local authorities - Edinburgh, Aberdeen, East Lothian and Clackmannanshire. Aberdeen and Edinburgh City Councils have been at the leading edge of nascent open data work in Scotland and can be seen as 'mature' players, willing to share their knowledge and expertise with others. East Lothian and Clackmannanshire came to the programme with little or no experience of open data, but with an ambitious attitude and a willingness to experiment and embrace innovation.
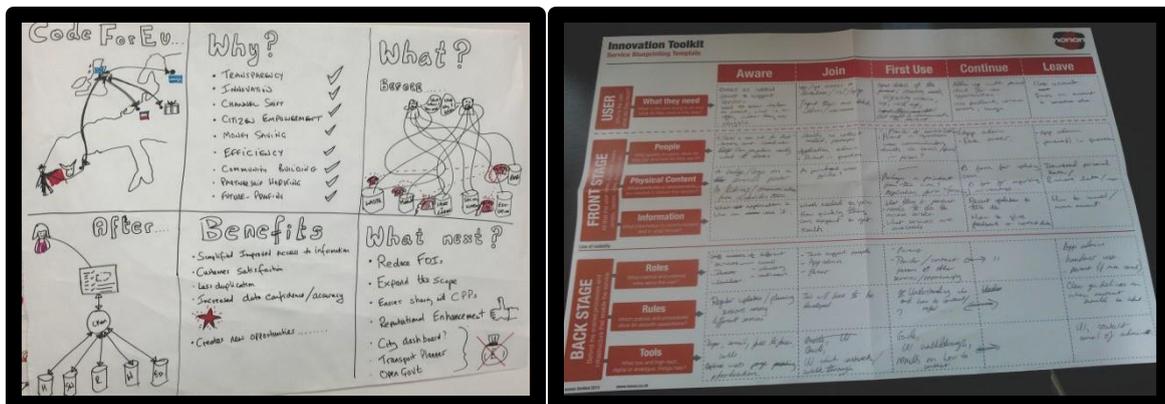
Each local authority was appointed a 'Code Fellow in Residence' (a technologist) who has worked intensively with the local authority staff over 12 months to open up data sets, publish these on a portal so they can be re-used and created new digital public services - apps and web content to enhance both citizens and visitors experiences of the local authority. A 'Designer in Residence' also worked with the technologists and local authority staff across the four authorities.

We have been part of the wider 'Code for Europe' programme which has involved designers and technologists across Europe working with civic authorities to increase the use of open data sets to enhance civic transparency and improve decision making.

Clackmannanshire Council is Scotland's smallest mainland local authority and their learning from the programme below demonstrates that this is not beyond the reach of any government agency or public body in Scotland, with the right culture and access to skills.

Approach

With little prior knowledge of Open Data, our initial ambition for this project was to develop a mobile app which would provide personalised access to childcare resources as part of the early intervention strand in our 'Making Clackmannanshire Better' change programme.

As the project evolved we focused on three main areas: Knowledge Transfer, Developing a Portal and our Childcare Application.

Knowledge transfer provided Council officers with information about standards and systems used in Open Data, the ecosystem of agencies involved in Open Data and the sources of existing Open Data applications which were available for re-use.

Outcome

We successfully built a CKAN Open Data portal and developed an app called Clacks Kids which is a location based service directory. Spin-off activities have lead us to develop an open GIS mapping portal which is likely to inform our future GIS Strategy and a reporting platform based on the Open311 standard.

Lessons learned

- The biggest lesson is that even a very small Council can engage with Open Data. The key components are having access to people with the right skills and attitudes, easy access to servers and software with which to tinker, and permission to experiment.

- Developing an Open Data infrastructure is important if the project is to be sustainable. While the "app" may be the most high impact product, without the infrastructure there will be no data to use in the app. Apps are also transient, they will be replaced by other apps in future.

- Follow your nose! We have revamped our GIS infrastructure opening up the opportunity of significant future cost savings as a direct consequence of our need to provide mapping tools for this project.

- Civic Apps are not as easily transferrable from one Council to another as you might expect.

- In a shrinking Council, persuading others to prioritise your project can be difficult especially when there is nothing concrete to demonstrate. You need to have something to show people. Once we had a working app, we then found services coming on board as they could see how it could be used.

2. Crichton Institute: Regional Observatory

In promoting this project both within and beyond the region, both local and global issues have collided. It is clear that new technology has created an accelerating hunger for information and we have observed with interest the parallel dialogue around Open Government and the 'Smart City' agenda. It seems to us that there is something of gap in strategic thinking and policy and we have been asking the question: '…if there is such a thing as the Smart City, what would the Smart Countryside look like…?'. So there is a dialogue that needs to take place about rural-specific opportunities in the open sharing of data and service improvement and provision which we feel we should follow.

The above issue is compounded by the overall capacity constraints which rural agencies face. While an obvious plea would be for more resources for rural areas in this field, there are perhaps more immediate advances that can be made by better sharing of experience and existing resources currently being directed to urban areas/solutions.

Raising this issue has gained us some exposure. We have, in addition, been cited as an example of regional-level data innovation in the recently published SG Open Data Strategy. We have also been encouraged by the fact that others, including those much better resourced than us, have had to grapple with the same issues and that we are seeing emerge a community of like-minded people who are prepared to provide advice and support.

Background

Crichton Institute is a Scottish Funding Council (SFC) funded collaboration involving the academic partners based on the multi-institutional Crichton Campus in Dumfries, south-west Scotland (University of Glasgow, University of West of Scotland, SRUC, Open University, Dumfries & Galloway College). The Regional Observatory is one of the strategic arms of the Institute and has been in development since the start of CI in January 2013. CI's work has very much concentrated on Dumfries & Galloway. The SFC funding period comes to an end in December 2016.

The objective of RO is to provide an information and knowledge portal that acts as a one-stop open access service for open data, information and intelligence on a wide range of social, economic and environmental factors across Dumfries and Galloway and the South of Scotland.

Approach

Rural areas have in general been poorly resourced in terms of data gathering, access and usage and D&G and the South of Scotland are no exception. In many cases, public and 3rd sector agencies have had to resource external consultancy to assist with even the most basic of regional data gathering and interpretation (the exception being the local NHS Board which has a well - resourced public health intelligence unit). Effective data sharing has, as a consequence, been somewhat the exception.

With a lack of capacity and no consistent track record of high level collaboration, the benefits to be derived from sound data management and data sharing have not been fully understood or exploited. While individual agencies are striving to take advantage of new information and communications technologies, the absence of effective data management is inhibiting the genuine desire to move to a more 'open government/open data' culture. Change in recent developments in Community Planning, a move towards better understanding the needs of service users and service integration (within and between agencies) is however supporting the drive towards a culture of open government/open data in the region. In fact, over the last few months Dumfries & Galloway Council (DGC) have auditing their records management and are working on an open data strategy. At the same time Third Sector Interface in D&G have been working with *Think Data Scotland* around the issues of data gathering and sharing within the Third Sector community http://www.thinkdata.org.uk.

Outcome

The setting up of the RO was approached as much as an organisational development and trust-building issue as simply an exercise in data management/sharing. Hence the RO has been developed in close collaboration with the Dumfries and Galloway's Community Planning Partnership. The objective from the outset has been to ensure buy-in to the concepts of Open Government, collaboration, service improvement and data sharing across institutions, communities and businesses across the region.

The first step was to engage in discussions with the local Community Planning Partnership to ensure that there was a view that such a thing as a data observatory was needed, but also that there was high level cross-agency support for its development. The proposal was greeted with enthusiasm and two tranches of support funding for the early development stages of the Observatory.

With support secured, an initial Technical Group was established which included representatives from DGC, NHS Dumfries and Galloway (NHSDG), Scottish Centre for Enabling Technologies (SCET) and CI.

This group worked together to agree on the vision, look and feel of the online portal, the process of populating it, maintaining it and promoting it. As part of this some desk research was done to look into what other observatories and open data portals offer. Some of these were approached directly to inform us whether we were heading in the right direction. Armed with this background, a Project Initiation Document was agreed and specialist part-time consultancy engaged to convert the vision into reality.

One issue we struggled with was '…when do we go public…?' We were confident with the basic functionality/feel of the portal, but less assured on content issues. We have made life difficult with the notion that our customers would not just be the usual professional data-users. Our vision also included, for example, local P6 pupils using the portal for a project on "Jobs in our Region", complete with map-building and visualisation tools.

We decided to opt for a 'soft launch' of the RO website in June 2014. The Technical Group became the Data Suppliers Group. With the key technical issues addressed and the portal functioning it was felt that the 'harvesting' of data and documents should now widen to include others, for instance Third Sector Interface Dumfries & Galloway. Approaches have been made to Police Scotland and other local community organisations. These discussions have served a dual purpose. As well as serving to promote the RO and secure further data access, they have opened up dialogue about the promotion of data sharing, better understanding, and more effective service delivery. This was supported through a workshop in the autumn of 2014, which was also attended by Scottish Government providing us with confirmation that we are moving in the right direction and that we have the same issues as other open data projects.

However, even though feedback has been overwhelmingly positive and encouraging as people like our vision and the objective, non-partisan approach, since RO does not have any authority over the various stakeholders, progression around data 'harvesting' has been extremely slow.

Lessons Learned

With Phase 1 of the project complete, we have a well- functioning and attractive portal, though it is still fairly one-dimensional. However, local stakeholders and the discussions as part of the autumn workshop showed a great interest in RO providing access to a regional economic dashboard, providing instant, up-to-date access to key economic performance data for D&G, as well as access to

interactive mapping and other visualisation tools. With not having technical expertise within RO and the Data Suppliers Group we were looking at how best to enhance the portal. This led us to the School of Computing Science at the University of Glasgow. After detailed conversations we ended up with 140 3rd year Computing Science students who worked on a number of IT related projects with RO and other D&G based organisations: Crichton Carbon Centre, NHS Dumfries & Galloway, Third Sector Interface D&G and The Stove Network. A total of 11 projects presented a range of real-life challenges to the students helping to find IT solutions to questions such as:

- How can regional economic indicators be visualised and made more accessible to everybody?
- How can we sell waste products from our industrial process, rather than land-filling?
- How can individuals find out about local volunteering events and sign up to participate?
- How can I discover where people go and what they do in my visitor centre?

Prototypes were created and in some cases the solutions offered are of such quality that the intention is to have them go live. One of the projects also looked at a redesign of the RO website, using WordPress for the actual website, while bolting on CKan open source application to deal with the data.
This collaboration provided a great platform to show what is possible, but it also left us again in a situation that, unless we find resources to secure technical support we are unable to migrate the projects, implement them, maintain and further develop the solutions given to us.

3. The City of Edinburgh Council: ARC-E App

The app enables the service area (Health & Social Care) to open up data that was not previously available to the public in this format. Previously service information for addition recovery support groups in Edinburgh had only been available through leaflet and PDF formats. An API was created with this data and can be shared with the app and other applications.

The app has been built so that it is scalable and more features can be added. The framework of the app can also be redeveloped to suit other groups with similar needs.

Background

ARC-E App was created by TM&R Ltd (Anne-Marie McMann and Ella Robbins) and developed by David Morrison as part of the **EdinburghApps** 2014 civic challenge programmed.

The application, the Addiction Recovery Companion - Edinburgh app (ARC-E app), is aimed to support those in the process of recovery from an addiction. The app allows users to document and reflect on their progress, becoming a constant companion and supporting them as they help themselves. It will make it possible for people who already have some support from Council services to use their mobile device to help them in their recovery from addiction.

Approach

The approach taken to develop the app was an agile, co-creative approach. Through working this way the team have been able to develop the app with the Council service area and services users to ensure that the deliverables are being met and a worthwhile product is created.

Design & Build

The approach of the design was to put the needs, wants and limitations of the users at the heart of the design process. From the start, the project ran focus groups with potential users of the app to inform on how to move forward. First the team made sure that the objectives reflected problems that impacted recovering addicts and then tested and iterated on potential solutions using low-cost prototypes before implementing them.

To build the product, a version of the Scrum agile development approach was adopted and adapted to fit the small and distributed team. This approach recognises that requirements often change during a project and the team has to be in a position where it can quickly adapt to these changes.

The project was divided into objective themes. Each objective theme contained a collection of user stories and at the end of every iteration the team produced a build of the app. This build was tested against the user stories for the iteration and used as an artifact for user testing. This allowed the Council to assess the current build for milestone acceptance and potential users to test and feedback on its value to them. The outcomes of testing influenced the planning of future iterations e.g. new user stories maybe added to the backlog or remove ones that have been shown to be invalid. This ensures that a meaningful product is being built at every stage.

Outcome

ARC-E App was developed in order to:

- improve access to appropriate local support services and information about the service.
- make it straightforward for users to reach out for immediate support in times of crisis.
- help users keep track their appointments and commitments, related to managing their recovery.
- keep users up-to-date on events organised by the council or by members of the recovery community that might be relevant to their recovery.
- allow users to look back at daily messages to support motivation to stay on track.
- allow users to access mindfulness activities, particular during a crisis/emergency.

The app is available for both Apple and Android devices and is generating interest as this solution is reusable for other authorities providing this service, as well as offering other re-use options.

Lessons Learned

The main lessons learned have been around working co-creatively. The client (Health and Social Care) and the end user (people in recovery from addiction) have been involved at every step of the process. Working this way has ensured that milestones have been hit on time and on budget whilst creating an app that meets the user's needs and achieves the objectives set out by Health and Social Care.

4. The City of Edinburgh Council: Edinburgh Apps

This programme is completely transferrable to any other organisation and sector. Edinburgh did not create something that was untried – civic challenge competitions take place all over the world, and are very successful. Supporting events, hack weekends, data days etc. are also happening widely, and are not expensive to produce. All of these events add to learning and increases awareness of the power of open data. It is a new way of working, but it is already the way many companies work, and something the public sector needs to do to find efficient and cost effective solutions.

Edinburgh's track record speaks for itself – its agile approach to development meant that quality products could be built quickly and were known to meet customer needs. Most of the products are shareable which means the public sector can use them right now.

Background

Launched in 2013, EdinburghApps was the first event of its kind in the UK, a civic challenge programme that works with the Council and other partners, encouraging developers, designers, creatives and small businesses to take part, and offering winners business support and the potential opportunity to work with the Council to develop their concepts further. Participants choose from challenges set by the Council around a number of key themes.

At its core is a vision to change the city through encouraging innovation with technology, design and user-centric development. Edinburgh has exceptional design and tech communities and a large number of young companies in these areas whose fresh thinking mean that Edinburgh has great opportunities to produce original and cutting edge solutions to city challenges.

The programme of challenge events:

- supports growth of and partnership with new IT, design and other related businesses and partners in the city
- encourages a digital culture change internally, supporting skill development for council officers
- delivers innovative and efficient solutions for the Council's customers, in line with the Council's priorities and the ICT and Digital Strategy objectives.

Edinburgh Apps was developed to support the Council's Open Data strategy. For each challenge, data sets are shared, increasing the Council's delivery of open data and opportunities for innovation EdinburghApps wants to change the city by providing creative, customer driven solutions to city challenges. It aims to work with everyone interested in making this change happen.

Approach

EdinburghApps began as an annual once a year competition with the Council providing challenges and teams taking part over 6/7 weeks to develop strong concepts or/and prototypes which are then judged in a final event. The winners then have the opportunity to work with the Council to develop their ideas, and deliver products.

EdinburghApps now runs a range of events to encourage solution finding working with key partners, Council officers and customers

- Annual challenge competition
- Subject hackathons
- Service area mini events

All of these events aim to support partners in finding innovative solutions to business and city challenges. Data is a core requirement in all of this, and is published as open data whenever possible.

Winners of these events have the opportunity to take forward their proposal for development with the appropriate area.

The benefits of this approach are:

- Delivery of new digital products which meet a clearly defined need
- Delivery using an agile approach, and at a far reduced cost to working with larger companies
- Customised solutions, co-created, which are built directly to meet requirements
- Building longer term relationships with local IT & Digital companies
- Opportunities to support the growth of the city's business economy

The competition event is now in its third year and has been very successful encouraging a wide range of entries and the delivery of a number of products. These include:

- **Tend** – routing tool which optimises planning and deliveries for Health and Social Care's Equipment Delivery Service
- Recycling Edinburgh – a location app case study, sharing recycling facilities in the   city
- Run The City – an app aimed at visitors looking to explore the city using running routes, offering a commentary on places on interest
- ARC-Edinburgh – a buddy app to support those in recovery programmes for addiction

The Council has helped winners to start their business from scratch, and also supported participants to find other business opportunities. Edinburgh Up Close was also recently launched, working with technology developed by a winner from EdinburghApps 2013.

The event runs in three events – a kick off weekend, a midway workshop and the judging final.

Outcomes

When EdinburghApps was first launched it was intended to bring about a number of benefits, including:

- new thinking to solve city challenges
- innovation in technology and design
- the sharing of civic data
- stimulate the city economy through working with SMEs
- social change for the city

The programme has achieved these outcomes, but it has also brought about much more:

- innovative and cheaper solutions
- improved sharing and publishing of open data
- ongoing relationships with new businesses
- change to ways of working
- awareness raising for open data
- new business thinking
- benefits to customers

Lessons Learned

EdinburghApps has now taken place three times, and hack weekends have also been run in between to generate wider awareness and explore specific subject based challenges. In 2016, Edinburgh expanded their approach to work with partners such as NHS Lothian and Sustrans Scotland which meant that challenges could be more strategic in nature, and provide service improvements that had a holistic impact. There are synergies for organisations both in terms of challenges and for product and data usage. Longer term, Edinburgh will continue to build relationships both with partners and businesses in the city and with the tech community to achieve sustainable outcomes.

Edinburgh Council will continue to expand the sharing of data as well as the sharing of ideas.

- **Build support**: it is important to have a suitable sponsor in your organisation (and some funding) to do this. Edinburgh could demonstrate it was achievable because it had been done elsewhere and this helped them find supporters. Align with relevant strategies in your organisation, this will also build support.

- **Changing business thinking**: when Edinburgh started this programme they didn't realise the impact of bringing business change into the Council. Inviting developers and designers to work with them brought fresh thinking and new ways of working. This has had an interesting internal ripple, and they now find service areas keen to see what can be achieved, not just with a product development, but for their service generally.

- **Data**: this is the central component and takes time to find, cleanse and publish. This can be challenging and service areas may need help with this work. Ideally a data resource should be available to do this.

- **For the competition itself**: Edinburgh have discovered that a mix of skills works better for teams and builds better prototypes, so they now advertise across a number of sectors. You need developers to support the whole event, provide mentoring and knowledge sharing, so build relationships with your local tech and design communities. Some teams have no idea how to deliver their idea
  - Edinburgh Council are now offering a midway workshop to help teams learn how to build a proposal, cost and plan their concept. They have developed their own processes around the event which include the use of design thinking, customer experience and business planning.

- **Challenges**: Whoever submits a challenge must now take part in the whole event, providing further information and advice for teams. This means that, whoever wins, the challenge owner is already engaged with them and it makes it easier to take the project forward.

- **Funding**: funding for the competition is not straightforward, and should be bundled into a larger business case for open data and innovation. For those taking part, there is an expectation that the Council will fund development. In 2016 Edinburgh put in place a co-production approach. Winners continue to work with challenge owners, who will work to identify funding once a concept has reached that stage of maturity. A lot of entrants have full-time occupations, so there would be a risk in going straight into a contractual arrangement with timelines that are difficult to meet. Both parties have time instead to reflect following the event, and can decide how they want to move forward. For funding, identifying appropriate sources is very useful; and for this involving funding officers is really worthwhile.

- **Sponsors**: a range of sponsors and types of sponsorship are required and this is time consuming to achieve. It is never too early to start working on this.

- **Communications**: communication has to be regular and continue throughout the year, not just around events. This requires resource and should not be under-estimated. Engagement is essential to keep your audience interested and encourage them to come back each year. Use appropriate platforms such as Twitter and Instagram and set up a blog so that you keep everyone up to date on activities.

Finally, and most importantly, look for any opportunity to work with others in this area. This is one approach but there many other methods for engaging and changing thinking, and developing open data. Build partnerships to make it easier to accomplish more. Edinburgh work with individual developers and creatives as well as companies, and we do look for those who have the same goals for data and innovation and want to see change happen.

5. The City of Edinburgh Council: Run the City App

The Run the City App solves a challenge faced by running enthusiasts who are new to the city by providing routes around the city of different lengths. It increases visitors' engagement with the city by highlighting city sights and providing engaging anecdotes.

The app has been built so it is scalable and with the intention that other cities and routes will be added.

Background

Run the City is a guided tour for runners and winner of the 'wild card' challenge for **EdinburghApps** 2014. Runners will always get their run in, even when away on business, but running in a strange city is difficult when you don't know where to go. Run the City solves this challenge as the app, through audio messages, not only gives runners directions but also highlights their attention to the city sights and makes their run in Edinburgh more engaging with anecdotes about the areas they are running through.

It utilises the Council's open data as content for the app and will also create data the Council can make open.

Approach

The project was undertaken over two main stages the Build Phase, and the Beta Phase. The initial build phase allowed us to deliver a minimum value product which could be tested to ensure that on the project was on track to deliver the objectives before full build was complete. This also ensures that a valuable product was being delivered that people wanted to use.

The team that developed the app worked co-creatively with the Council service area to ensure that the app met customer expectations and also aligned with Council outcomes.

Build Phase

The build phase developed the main components of the app (login, cities, routes, tracker, activity, settings, activity timer and location tracker pages.) In this phase, before building the user-interface of the app Edinburgh Council created a route planning functionality, which allowed us to design and record routes to be uploaded

into our app. The milestone for this phase will be the delivery of the MVP (Minimum Value Product)

Beta Phase

This phase involves both production of the audio for the tour and device testing and user testing. The test app was shared with runners/walkers around the city and they were asked to test the runs and report back any ideas or issues they had. Edinburgh anticipated that user testing of that app would take three to four weeks but it actually took longer and was carried out in two phases due to the changes required after feedback was received.

Outcome

An engaging running app, which considers routes which would appeal to walkers, has been created using Council open data. The app has been aligned to Edinburgh Outdoors and has created the additional benefit of creating data which can be shared.

The app is now live for Apple and Android users, and is gathering very positive feedback. This is encouraging the company who developed it to have conversations with a number of other cities for future development.

Lessons Learned

The idea behind the app was sound, and user research demonstrated that runners thought it would be a great way to experience the city. It turned out to be quite a challenging development, with issues around design and build. A major issue was calibrating the app for runners and walkers (due to differing speeds) and getting the geo-location to match. This problem was not easy to predict – the developer was very experienced; it was just an issue that had to be understood and solved.

Both members of the new company were in full-time employment which meant development was elongated. Ideally it would be better to work with an existing company who could commit specific time to the project. This would have increased costs and EdinburghApps is clear that it wants to support new business development.

This development has sparked a lot of thinking around digital tourism. Cities and towns will have a number of apps in the market place for their visitors. It is worth giving some consideration as to the audience you want to target. Our experience suggests that specific audiences may be more likely to download and use an app on their visit than general tourists.

6. Royal Commission on the Ancient and Historical Monuments of Scotland - SENESCHAL: Semantic ENrichment Enabling Sustainability of arCHAeological Links

Adopting Linked Open Data can benefit the wider heritage community through improving standards and introducing efficiencies. The benefits of publishing controlled vocabularies are starting to be realised. Simply by adding a SENESCHAL RESTful service into their Collections Management System, the Archaeology Data Service, University of York were able to access the authoritative controlled vocabularies remotely. This not only eliminates errors that inevitably creep in through free text typing but improves the consistency of indexing records.

Background

Controlled vocabularies are key to both the storage of information in the database and its discovery online. In particular, we use thesauri to help classify the types of monument, object and maritime craft associated with each site record. We encourage the use of thesauri standards amongst local Historic Environment Records (HERs), who maintain databases about the historic environment for local authority areas across Scotland, and more widely amongst the profession.
For cultural heritage, demand for Linked Open Data came from the research community. They saw the absence of controlled vocabularies as limiting opportunities for combining data from different providers through semantic links.

Major controlled vocabularies should act as hubs for the Web of Data, but publication as free text strings limits opportunities for connecting to data published elsewhere. Although we publish our controlled vocabularies online as thesauri, they are not particularly visible. The thesaurus for architecture, implemented in 2005, limits the potential of the terminology as the terms lack the persistent Uniform Resource Identifiers (URIs) that would allow our resources to act as hubs for the Web of Data. Adopting a Simple Knowledge Organisation System, or SKOS, using the Resource Description Framework (RDF) provides a more flexible approach enabling the vocabulary owner to define a concept rather than the term. Each concept is expressed as a URI. The concept may then be expressed in any number of ways including alternate labels, dialect terms or in different languages.

The development of Linked Open Data for cultural heritage is part of good practice, helping to deliver Government policy towards transparency and Open Data. Scotland's Open Data Strategy encourages Public Data to be published in reusable,

machine readable form under an open licence which enables free reuse, including commercial reuse to open standards following relevant recommendations of the World Wide Web Consortium. Moreover, Public Data from different departments about the same subject will be published in the same, standard formats and with the same definitions. Defining the concepts used to index records about cultural heritage is a first step towards meeting that goal. It introduces the standards and machine-readable formats necessary for interoperability. However, before becoming operational, it requires acceptance of the standards, investment in research and development time beyond the day-to-day operations of many organisations.

Approach

The solution was to find partners who understood the Linked Open Data requirements and to secure funding to enable the research and publication of Linked Open Data. We were fortunate that colleagues at English Heritage already had an established relationship with the Hypermedia Research Unit at the University of South Wales and that there was a shared recognition of the need to publish our vocabularies as Linked Open Data.

The partnership approach between a university research department and public bodies enabled a successful application to the Arts and Humanities Research Council for a one year Knowledge Exchange project. This made it significantly easier for vocabulary providers, such as RCAHMS, to make their vocabularies available as Linked Data and for users to index their data with uniquely identified (machine readable) controlled terminology that is semantically enriched and compatible with Linked Data.

The resultant SENESCHAL project (Semantic ENrichment Enabling Sustainability of arCHAeological Links) brought together vocabulary providers from English Heritage, RCAHMS and RCAHMW, together with the Archaeology Data Service University of York with the domain experts, Doug Tudhope as Principal Investigator and Ceri Binding as Research Fellow, at the University of South Wales.

Outcomes

Intended Outcome	Actual Outcome
Freely accessible and reusable persistent vocabulary resources as linked data, the techniques to achieve this being made freely available.	Achieved: http://heritagedata.org established as the home for Cultural heritage reference vocabularies and concept schemes published for RCAHMS Monument Type, Archaeological Objects and Maritime Craft Type Each concept has its own unique reference indicator.
Web Services to SKOS representations of the vocabularies and semantic enrichment services, along with web application components	Achieved: Downloads, Services and Widgets published. Users are able to download the vocabularies in various flavours of RDF (N-Triples, Turtle, JSON or XML). A series of REST URI calls have been developed for the vocabularies with results returned in a JSON structured string which permit AJAX callbacks for use in browser based applications.
Knowledge exchange tools to facilitate semantic enrichment (via URIs) within data entry. Development of downloads, Services and Widgets.	The project has also developed a suite of predefined visual user interface tools, or widgets.
Mechanism for feedback of supplementary terms to augment existing vocabularies	Not Achieved: candidate terms are still submitted through RCAHMS own thesaurus management system and data periodically uploaded to heritagedata.org website
Raising the profile of Linked Open Data with Historic Environment data curators in Scotland	Achieved: through a workshop was held in Edinburgh at the end of the project for stakeholders and presentations on Linked Open Data to stakeholder groups.

	Additional outcome:
	Demonstrating application of approach to handle multi-lingual expressions of concepts: During the course of the project we were able to make use of Gaelic translations of the monument type vocabulary provided by Historic Scotland from a Bòrd na Gàidhlig funded project. So a concept may now be expressed in English or Gaelic, with a preferred or alternate label.

Lessons learned

Publishing the terminologies as Linked Open Data is the first tentative step toward delivering cultural heritage data as 5 star data. Maintenance and update of the terminologies is not seamless and requires periodic data uploads, so the vocabularies may not be up-to-date instantaneously.

Exposing controlled vocabularies is inevitably organisation-driven and there is a need, where appropriate, to align vocabularies by theme to deliver further efficiencies in maintaining and developing resources. Through our membership of MEDIN we are exploring opportunities to develop more marine and maritime-related Linked Data vocabularies with colleagues in Belfast and Dublin.

The benefits of Linked Open Data have still to be fully realised within the business and more widely across the heritage community. However, making the terminology more openly accessible as Linked Data should encourage wider adoption of standard terminology, develop interoperability with other related resources, and encourage community feedback on possible improvements to the vocabularies. Opportunities will continue as part of the new organisation Historic Environment Scotland when RCAHMS and Historic Scotland come together in October 2016 to form the new lead body for Scotland's historic environment.

7. Registers of Scotland (RoS): Cadastral data for the INSPIRE directive

Scotland has made significant progress in publishing spatial data in a prescribed format driven by the EU INSPIRE directive. The local government sector is also currently developing a project to support a more collective approach to the management and publication of spatial information, providing access to all spatial data created by local government in a consistent form.

RoS is proud to be at the forefront of the provision of spatial data. We would recommend any approach that supports collective management and publication of spatial information. The challenges faced internally are far outweighed by benefits realised in the short and longer term.

The experience we have gained over the last seven years with INSPIRE has led us to increasingly identify considerable benefits of a coherent, trusted, and consistent set of information on land and property in Scotland. Enabling access to core land and property information in one place where it can be made available to all is increasingly important for Scotland as a whole.

Background

European Directive 2007/2/EC, better known as 'INSPIRE,' was transposed into UK law in December 2009. The aim of the directive is to establish a spatial data infrastructure (SDI) for Europe. In general terms, this means providing an IT infrastructure which allows access to harmonised spatial data (data collected to the same standards and requirements) via the internet. In theory, an SDI should improve the access and use of data at local, regional, national and international levels, improve data sharing between public authorities, and improve public access to spatial data.

INSPIRE instructs EU member states to make spatial data available in a consistent format which come within the scope of the directive, as well as providing network services (mostly internet access) and metadata to support the data. You can read an informal consolidated text of the Scottish INSPIRE regulations here.

The Scottish government is responsible for the management of INSPIRE in Scotland. The management is coordinated by the Spatial Information Board and

work has been broken down into five main areas. The two areas of interest to RoS are land, property and addressing; and service delivery and technical implementation. All EU member states are required to submit a monitoring report with details of available datasets to the European Commission every May. You can read the UK's most recent monitoring report here.

Approach

Our first step was to establish a project and project team to handle the legal, commercial and technical aspects of INSPIRE.

During the lifetime of the project we consulted widely with other European and UK organisations, both within and outside our domain. Colleagues within the Scottish government and UK geographic community provided an excellent and knowledgeable resource. This enabled us to overcome a wide variety of challenges and allowed us to improve our own expertise in a number of crucial areas.
As well as the technical requirements, we had to consider the wider implications of INSPIRE on our business and staff. These included the effect on our commercial activity and the types of services we offered, our IT infrastructure, and any legal impact on our day-to-day activity. Each of these requirements was processed by a small multi-disciplinary team reporting back to the project board who led the overall INSPIRE strategy.

The nature of the legislation naturally broke the project into a number of phases, each of which required an increasing level of resource and budget.

- **Phase 1**: metadata – in May 2011, we complied delivering GEMINI 2.1 metadata describing our land register data and the future web mapping service (WMS).

- **Phase 2**: discovery and view services – in November 2011, RoS provided access to the metadata created in phase 1 to the Scottish Spatial Data Infrastructure. At the same time, RoS provided a view of its initial cadastral parcel data. For the deadline, RoS chose to use the services of a third party (ThinkWhere) to host the WMS element of the service.

- **Phase 3**: download – RoS delivered a service that will allow a customer to download all or part of our land register dataset. RoS again chose to use professional services of a Think Where to host the download service.   Licensing considerations on the reuse of data were investigated and led to the creation of an INSPIRE download license.

- **Phase 4**: fully compliant – this phase will deliver full inspire compliance by supplying parcelled cadastral data by November 2017.Outcome

RoS has delivered the first three phases and is on course to fully comply by November 2017. The service is being increasingly used by customers and has sparked wider thinking about our data within RoS.

Lessons learned

The majority of the challenges that RoS faced were based on technical and compliance issues as well as data re-engineering. Our recommendation for any organisation with an INSPIRE obligation would be to ensure your internal domain expertise is brought together to guarantee you have a firm grasp of the issues and technical requirements required. For RoS, this meant a multi-disciplinary team drawn from IT, Geographic Information Systems (GIS), senior management, legal, commercial and core business. RoS consulted widely with fellow organisations and took part in a number of UK and European working groups to make sure we had an understanding of our responsibilities, as well as having an opportunity to influence those discussions. We would recommend that organisations seek advice, support and best practice from professional bodies, as well as learning and investigating best practice from examples throughout the world, including RoS and the Scottish Government.

Although the investment in INSPIRE can be onerous, there are considerable benefits that can be accrued if your organisation is committed to INSPIRE. For RoS, this meant spatial data has been brought to the forefront of the business, improved our expertise, developed staff, and led to several customer-focused initiatives. The core aspect of INSPIRE, data, and access to it, led us to re-evaluating data and data quality, as well as influencing a wider digital transformation project.

8. Scotland's Environment Web: EcoHack

Scotland's Environment Web wants to help people discover and understand more about the environment. Environmental data is really important – to provide context to reports on the state and quality of the environment, to improve our understanding of the challenges and opportunities our environment faces, and encourage communities, school children and individuals to investigate their own local environment further, observing what is happening around them, collect their own data and take action to protect and improve their local environment.

Background

Putting our objectives into practice, a hackathon event was organised over the weekend of 30th and 31st May 2015. Students from universities throughout Scotland were invited to Edinburgh, to come up with fresh new innovative ideas to make better use of available data, and to collect new local environmental data that can help further our understanding, and encourage people to get interested and get involved in Scotland's Environment.
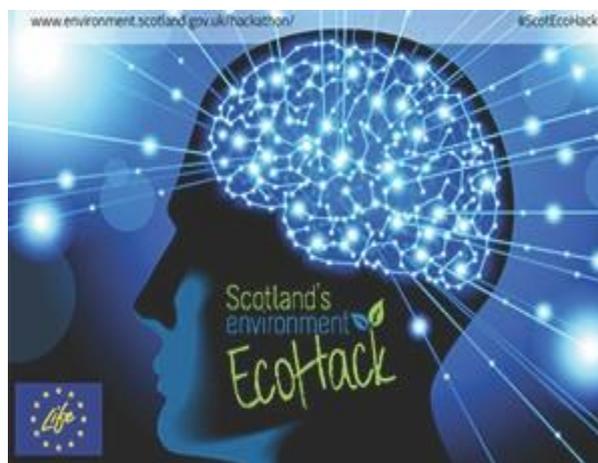
Interest was generated in the event via a number of routes:

- We had university lecturers and students on the steering group and who   also helped out as mentors so were able to help spread the word to their students and peers.

- A leaflet was sent to all universities and posted on their Facebook pages.

- For students one of the most accessible forms of quickly sharing information is on social media, with a lot of co-ordinated information sharing posts on Facebook and twitter (#ScotEcoHack) in the run up to and during the event that were shared and retweeted to an extended audience, bringing lots of new twitter followers to @ScotEnvironment following #ScotEcoHack

Examples of the interest generated on twitter can be found here.

Approach

The EcoHack challenge

During the weekend event we challenged teams of students and mentors to explore data and develop ideas that could make a real difference in helping people observe, monitor, educate and take action in the environment. Ideas were encouraged around exploring new data relationships to help analyse the state of our environment and the impact it has on us, develop apps that use and visualise data to help explain and view the environment, and provide new ways of collecting and viewing data.

A wide range of open source data was available to the teams - dataset list – and they were allowed to choose any platform and programming language and spent the weekend collaborating and being creative, innovative and inventive.

In the run up to the event, we provided links to information about a range of environmental issues to inspire new Ecohack ideas, covering topics such as Air Pollution, Water, Soil, Young People and Citizen Science, Environmental data, Nature, data visualisation, EcoSchools, Climate Change and communities, mobile apps, infographics.

EcoHack mentors

We couldn't have run the event without the help and support from our mentors. With a wide range of skills and experience, they were on hand to provide advice and guidance to the students throughout the development of their ideas from initial scoping and definition right through to the development and presentation of the prototypes. Some of the mentors saw some real opportunities in using some of their own data and tapping into the expertise of their mentor colleagues, and worked together to develop some of their own ideas to share with us at EcoHack.

Outcome

Feedback from all who attended was overwhelmingly positive and we hope to keep in touch with many of those who supported the event – judges, mentors and students. The standard of ideas was very high and in the end the judges selected 2 winning ideas and 1 runner up. More information on the winning ideas and a video of soundbites from the event on the EcoHack webpage.

9. Perth and Kinross Council: Open Data Workshop

It is important for an organisation to engage data users as it begins to consider making more of its data open and accessible. When it comes to how best to engage, there is no one right way. However, garnering views early will allow particular datasets to be prioritised for publication. It will allow you to understand what goals stakeholders and re-users have and what datasets they identify that would help them to achieve these goals. It will also offer the opportunity to advertise the release of key datasets to partner organisations and wider potential re-users. This will help to create a network of open data advocates.

Background

Perth and Kinross Council covers one of the largest areas of any council within Scotland and has been one of the fastest growing areas of Scotland in past decades. With a population of around 147,000 and predicted to grow by another 20% to 2035 they have a lot of challenges to face that a mature open data framework can help with. However, like most other public bodies the need to select the appropriate data sets which are related to city challenges was an issue. Therefore, Perth and Kinross Council embarked on a process to allow them to prioritise the datasets they can release by gaining feedback from key stakeholders.

Approach

Perth and Kinross Council ran an "open data identification" workshop with community planning partners, regional organisations and council officers from a range of services. This ensured that a wide spectrum of individuals were able to give insight into which data sets would be most useful. The workshop aimed to answer three main questions:

- What are the major challenges facing Perth and Kinross?
- What key datasets will help us understand these issues?
- How can this data be used to inform policy and service provision?

To answer these questions attendees were divided into four groups each with a facilitator. The first issue was to identify the major challenges facing Perth and Kinross under themes that included: Live, Work, Visit, Move, Learn and Inform. Attendees were then asked to come up with key datasets that were pertinent to these themes. A wide range of issues were identified including: costs of housing; GP

waiting times; the services available to an ageing population; low wages; access to childcare; and access to broadband. If attendees felt that the pre-identified themes where too restrictive they were encouraged to come up with their own themes which were added to the set.

Using the answers from the first question the groups then moved onto identifying potential datasets which could help alleviate or tackle these challenges. Many datasets were identified as priority, reflecting a wide range of concerns and challenges facing the area. Among these were destination of graduates of the local University; footfall in Perth City Centre; road traffic collisions; crime data; and statistics for the number of visitors to local events and attractions.

Further notes from the day included:

- There was agreement on the need to involve community groups in the process of data selection and the crowd sourcing of community data was also identified as playing an important role in the platform.
- The need to aggregate to anonymise data was stressed. As was having a robust mechanism in place to ensure data quality.
- People wanted to know not only about spending and access to services, but how it can be broken down by area.
- The economic issues identified revolved around the fact that while unemployment is relatively low, there is a lack of high-skill, high-wage jobs.

Outcome

The datasets identified by the stakeholder workshop gave Perth and Kinross' team a solid basis from which to start creating a Publication Plan, though the process also involved looking at Council priorities and the plans other local authorities have published. Along with specific datasets, there were several reoccurring issues. Perhaps the most prominent of these was the different challenges faced by urban and rural communities in Perth and Kinross in terms of issues such as transport, digital connectivity and access to services. Consideration was given to how the data platform can reflect these challenges.

In addition, it was the first real chance for the Council to "advertise" not only the benefits of open data, but also to announce the development of an open data platform to a wide range of organisations. This in particular has been vital as the team has moved towards gathering datasets; knowing that they are grounded in stakeholders' views of the challenges facing the area and bringing an understanding of the nature of the data being released.

Next Steps

Perth and Kinross are now considering running future workshops based around particular themes. While these are yet to be organised, they are looking to "integrate" open data in number of topics and major initiatives. For example, this includes areas like Health & Social Care integration or how the council reports and manages performance.

10. NHS National Services Scotland: Prescribing Activity Data

The Scottish Government Open Data Strategy sets out the aim that Public Sector organisations should aspire to reach at least 3-star standard open data by 2017. Information Services Division (ISD)'s *Transforming Information Programme* therefore asks us to look differently at the way we publish and present data and intelligence. With England, Wales and Northern Ireland already releasing prescribing activity data in an open format, this presented an opportunity to learn from and build on the successes of those who had gone before.

A demonstrably high public interest in prescribing activity data provided us with a good starting point for publishing and developing a prescribing Open Data file. This being very much a pilot for the wider organisation, evaluation of the process and outcomes became a key component. We have seen some success in reducing the number of and amount of time spent on Freedom of Information (FOI) Requests, and are now using this publication.

Background

Prescribing was identified as a priority area for developing new ways of presenting NHS Scotland data and intelligence. It enjoys an already extremely high public interest in its raw data which singles it out for benefits that can be gained both from reducing the time and effort spent on, and increasing the value of, any individual data releases. The natural solution is to place all the data into one file for download by many.

A public consultation raised the possibility of open data in the autumn of 2015. It received a record number of responses (27), with roughly half from outwith the NHS. A clear divide emerged between respondents who wanted as much data as possible to be released (non-NHS), and respondents who recognised the value of open data but were to varying levels concerned about disclosure of patient and prescriber information (NHS). A full summary of responses to this consultation is published on our website.

We decided on this basis to pilot the publication of a Monthly Prescribing Activity Data file, pending a disclosure risk assessment. An organisational aim, to have our publications compliant with UK Stats Authority requirements for Official Statistics,

prompted an agreement to label them "experimental statistics" in recognition of their pilot status.

Approach

Scotland's island and rural geography uniquely shape the delivery of primary care, and so our risk assessment focussed on disclosures which might arise in this context. NHS Scotland's data landscape presents some challenges shared by other Scottish public bodies, particularly those that are community-based. Organisations may find our disclosure risk assessment helpful when contemplating their own.

The core approach of the project was straightforward: replicate the other UK nation's prescribing activity files with NHS Scotland data. Although we were fortunate that other nations had led the way, the responses to our consultation still recommended a cautious approach and we were proactive in making our plans and our protections open to scrutiny by national prescriber and pharmacy leads.

Each file comprising one month of prescribing activity in Scotland contains over a million rows of data. Given that we were publishing only a quarter of the volume of the HSCIC (England) data we chose to stick with their convention. Later on we were to encounter several parties who found these files unmanageable.

Endeavouring to reach out to new audiences we developed a communications plan centred on social media, in order to garner the interest, support and feedback of the active Open Data community. We also felt that the nature of open data meant that we had a responsibility to make the information accessible to all. Our latest development has been to produce a visualisation of the data which allows users to intuitively explore the information contained within.

Outcomes

ISD Prescribing aimed to achieve the following through this pilot project:

1. Complete the UK picture of prescribing trends
2. Reduce volume of and amount of time spent on FOI requests
3. Explore the practicalities of publishing open data and disseminate lessons learned to the wider organisation
4. Engage new audiences with prescribing data

Although we thought objective 1 would be completed with the publication of the data, a data quality issue, detailed on page 8 of the FAQ document, led to suppression of one of the fields. The feedback we received in relation to this one issue was extensive. Many stated that it is essential for UK-wide analysis which, with eventual publication of the field, actually validated the outcome. In national stats,

acknowledgement that a data release is being used as intended is surprisingly rare, making this a genuine win.

Analysis of FOI workload so far has returned mixed results. By most measures, FOI workload has decreased. The exception to this is the workload from brief FOI requests, which would tend to comprise raw data requests subsumed by Open Data. Contrary to expectations workload on these appears to have increased as a proportion of the (lower) total. Qualitative investigation may shed light on this shift.

Practical issues that arose included how we deal with missing data, the sheer size of the file, and how the data and supplementary information is presented on the website. A lot of work has been done to institutionalise the learning from our venture. Feedback from Open Data Champions within and outwith the organisation have been invaluable in shaping this learning, as they can often confront us with the issues that organisationally we might be blind to.

As part of continuous improvement we have collected feedback and become aware of a subset of customers who wish to use the data but find the files unmanageable. With these in mind we have dovetailed this project with further pilot work in data visualisation. Users can now explore a high-level summary of each month's Open Data file at the click of a button.
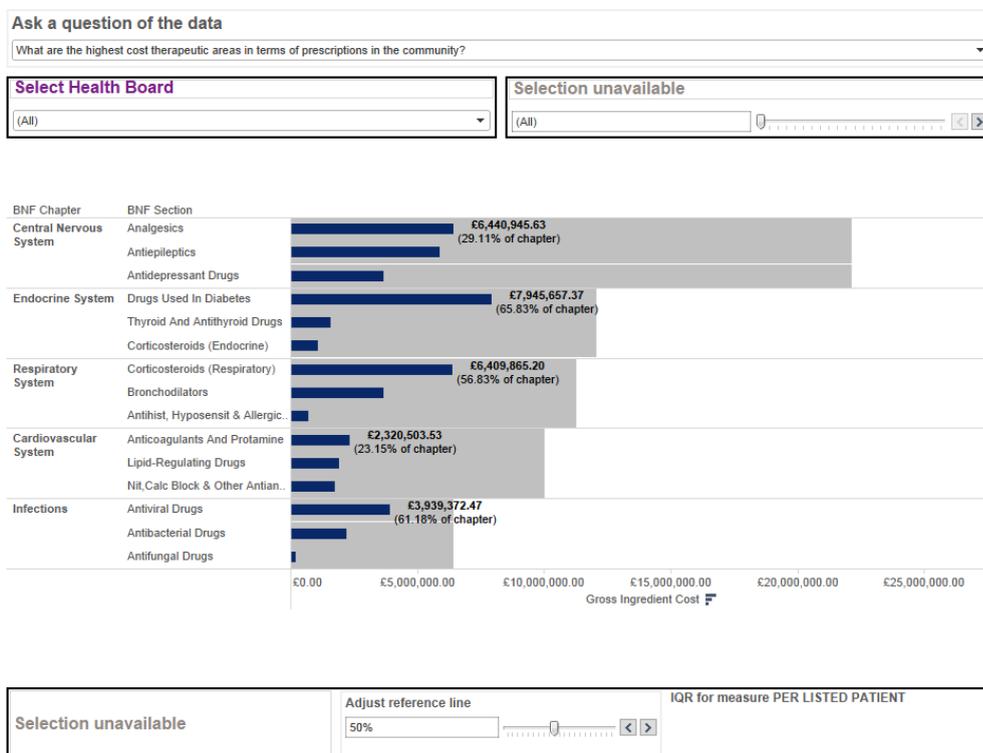
Lessons learned

- Open Data can reduce strain on services receiving relatively large numbers of FOI requests, particularly where these comprise releases of raw data.
- Organisations embarking on large data releases are likely to come across a wealth of locally-specific IT issues. These are rarely unassailable.
- Consider your audience: data and IT literacy may restrict engagement with new audiences.
- The Open Data community has established conventions by which it presents data and supplementary information. Although these are well-reasoned, largely stemming from ease of discoverability, they may not be immediately obvious to entrants in the arena. Fortunately they comprise an active and forthcoming community and their input can be invaluable when sought.

11. Marine Scotland: Open Data Network

An organisation's open data publication will likely be comprised of a number of different strands. Marine Scotland have embarked on a process to both align their open data approaches and expand the amount of data they were making open and accessible. This has culminated in the recent launch of a third open data portal to make their data more accessible.

Interactions with the new portal have been very positive and Marine Scotland are starting to promote engagement with it now it is live. With this open data network, Marine Scotland now have the right tools in place and are already making much data open and available.

Background

Marine Scotland cover many aspects of managing Scotland's seas; from policy development, planning and licenses to enforcement and science. Marine Scotland had already made progress in terms of increasing the amount of data they are making open and accessible. However, they had a challenge in terms of converging and aligning their already available open data tools and approaches.

Marine Scotland now have three sites in place that make data more accessible, enabling the public and colleagues to more easily locate and download data.

Approach

A clear business case was presented to the Marine Scotland Senior Management Board, and the necessity to align and expand open data publication efforts was recognised. After an initial review of available publishing portals, DKAN was selected as the most appropriate platform to progress.

Marine Scotland have also mandated particular metadata standards into the process of dataset publication. Adopting the MEDIN metadata standard, no dataset will be made available on either of the three publication portals unless it has appropriately high quality metadata attached. While this occasionally slows dataset publication, it also ensures that consistent metadata collection will become a routine business practice.

Outcome

Marine Scotland now have in place a 'trinity' of tools to make data more accessible:

•   The National Marine Plan interactive (http://maps.marine.gov.scot) - which delivers in depth access to GIS layers from Marine Scotland and a range of partners (Scottish Natural Heritage; Scottish Environment Protection Agency; Joint Nature Conservation Committee; Marine Alliance for Science and Technology for Scotland

and many more). From here you get access to view the maps, download the ones available, and access more information about the layers.

• Marine Scotland Data Publication Portal (http://data.marine.gov.scot). This is where Marine Scotland publish downloadable files and tabular data. The individual datasets published here are registered with Digital Object identifiers (DOI) to allow better citation of the datasets. Currently, the Data Portal predominantly hosts more scientific data and reports. Within peer reviewed journals, there are increasing demand that data mentioned in papers is made available. There is also demand around the ability to mint DOI for datasets which allows scientists to make data available without having to place it in orphan repositories.

• Marine Scotland Information (http://marine.gov.scot) – Marine Scotland's newest portal and where resources are tied together. On this site, Marine Scotland take in the service data from the other two platforms and provide information pages where the relevant maps, data sources and contextual information are put together. While both the mapping and data portal provide in-depth information and functionality for the specific jobs that they do, Marine Scotland Information is meant to be approachable for all. Users can search across all information types as well as more descriptive information that puts the maps and data sources into context. Marine Scotland Information merges content from several separate sections of Marine Scotland's web, which was previously used to supply datasets.

Portals:	Data	MSI	NMPI
Primary Audience	Analysts, academic, students, other data portals	General public, other departments, NGO	Marine planners, GIS analysts, other GIS portals
Delivers	Persistent Identifiers & Citations, Tabular data, visualisation, services	Context and background information, Quality information, linkage	Maps, mapping services, GIS Resources
System software	DKAN	Drupal	GeoServer + custom front end
Drivers	SG Open Data Strategy Improved data citation	Evidence based policy SG site revision Digital Public Services	Marine planning INSPIRE legislation
Subdomain	data.marine.gov.scot	marine.gov.scot	maps.marine.gov.scot

On the National Marine Plan Interactive, Marine Scotland now make approximately 800 spatial data layers/maps available to view. On the Data Publication Portal, they currently publish 100 data sets consisting of a mixture of data and reports. Finally, on the new Marine Scotland Information site, links to all of the above are available along with an additional 300 information pages crafted and maintained within Marine Scotland. It also includes hundreds of links to data sources and map layers from other providers.

Overall, the three sites take in between approximately 500 and 1,000 visitors per month each. While undoubtedly there will be a proportion of overlaps in these statistics, they are still serving up a considerable amount of data.

Lessons Learned

The work in Marine Scotland has taken a long time to get to the position where all of the appropriate tools are in place. Three sites were eventually chosen rather than one, due to a mixture of providing enough specialised tools for maps and data along with the gradual development of the approach.

Marine Scotland are now planning to do further stakeholder engagement and take in initial user feedback from visiting the sites. Now the new portal is live, they are also planning further promotion work. This is likely to be in the form of blog posts and tweets, along with an increased awareness within both Marine Scotland and the Open Data Community.

One of the key things will also be to integrate with other new exercises. So when new initiatives for data sharing in government and/or the marine data community pops up, Marine Scotland will seek to use the services now set up to be able to provide data to these exercises without additional work.

12. Aberdeen City Council: Leisure App

Many websites confuse the medium of delivery with the content that they seek to deliver. There is no separation of the information or data from the presentation, be it html, PDF, or print.

This lack of separation means that the data can only be used for the single reason for which it was published – and even that publishing is often poorly done.

By separating the data from delivery – and making it available as Open Data – allows users to consume it, and interact with it in the way that they need, potentially in much more sophisticated ways than were originally intended.

Background

Code The City, who now run ODI Aberdeen (Scotland's first node of the Open Data Institute) have run hack weekends using Open Data in Aberdeen for the last four years. In 2014, Aberdeen City Council commissioned Code The City to run a weekend-long session on the theme of Sports Fitness and Wellbeing. The council, as a service provider, was keen to understand from a service user's point of view what got in the way of their using the city's sports facilities to their full potential.

The weekend involved sports staff, service users, coders, designers, UX specialists, data wranglers, bloggers and others.

Several ideas were put forward, refined, and then teams formed to work on these. Some rapid prototyping followed. By tea time Sunday the attendees had developed a number of prototypes which were demonstrated to the council sports staff.

One prototype sought to address a specific problem which had been identified in relation to access to sport centre timetables. Its aim was to provide not only a solution to that but to show what might be possible if a better approach was taken to timetable data.

Approach

At the hack weekend it was shown that someone coming to the city council website, in the expectation of finding information on what sports activities were provided at what time, would face two hurdles: redirection to the websites of eight arm's length service providers; then 31 different timetables in a variety of formats. Many of these were poorly presented and were designed for print, not web.

Over the weekend, the data from these timetables was extracted using a combination of techniques including manual transcription and machine scraping of the data. A database was created and all timetable data was imported, and the data enhanced, which included adding geographic co-ordinates for all venues.

Finally a search mechanism was created allowing citizens to quickly find what they wanted through a single interface.

Outcome

Subsequent to the hack weekend, the Sports team commissioned the creation of a mobile app (which drew on the updated open data store) which allowed users to search, navigate and locate sports classes within geographic and time parameters; to store preferences, eliminate irrelevant classes; and for users rate classes for exertion levels. The app was designed after extensive user input and testing ensuring that specific, articulated needs were addressed.

Creating an Open Data store to hold timetable data allowed the data to be repurposed many times, for many purposes. While maintaining that data store raised issues which meant that the processes of multiple providers had to be changed to put the data (rather than the presentation) first and to ensure continuity of publication, the benefits to end users were clear and there was commitment to the continuity of data publication.