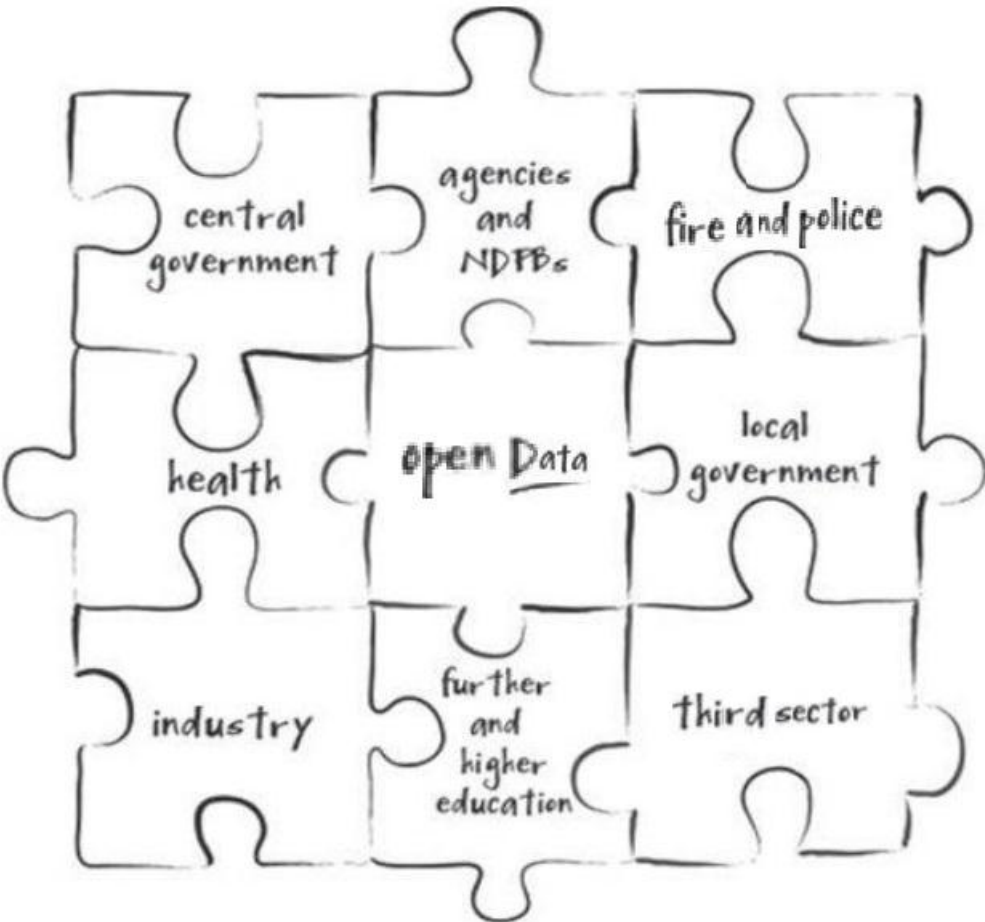# Open Data Strategy

Open Data Strategy

Introduction

The amount of data that the public sector holds is continually growing. The data ranges from sensitive personal information held to allow delivery of personalised services (e.g. health and social care) to non-personal information (e.g. environmental data). This data has value for the organisation collecting and holding it but there is potential to add value by making appropriate data available to others to re-use. We call this making data "open".

> Open data is non-personal and non-commercially sensitive.[1] Open data is easily discoverable, accessible to anyone and able to be freely used, re-used and redistributed by anyone. Open Data is data made available, via the internet, in an electronic format which supports its ready re-use, and with open licensing which allows its reuse.

The development of this strategy has been overseen by the Data Management Board supported by a short term cross-sector working group, with further contributions from the wider public sector. Its purpose is to help achieve the Data Vision for Scotland and the associated Action Plan. The Vision sets out ambitions for a Scotland which, by 2020, recognises the value of data and responsibly makes use of data to improve public services and deliver wider societal and economic benefits for all. Improved use of data will support achievement of Scotland's National Outcomes.

> This strategy seeks to create a Scotland where non-personal and non-commercially sensitive data from public services is recognised as a resource for wider societal use and as such is made open in an intelligent manner and available for re-use by others. Making data open will support:
>
> 1) Delivery of improved public services through public bodies making use of the data
>
> 2) Wider social and economic benefits through innovative use of the data
>
> 3) Accountability and transparency of delivery of our public services
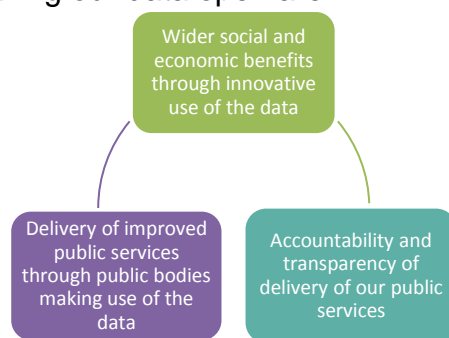
Over recent years there has been a move towards making more data available to others, such as the introduction of publication schemes and Freedom of Information requests. Open Data is a natural evolution and, if built in to an organisation's information management structure, can become part of the business process.

---

[1] Where organisations hold datasets which contain personal and/or commercially sensitive data it may be possible for them to either anonymise the data or to release parts of the data which are not personal and/or not commercially sensitive (this should be carried out in line with the organisations information governance procedures and in line with good practice such as the Data Linkage Principles).

Section 1 – Why make Data Open?

The key drivers for making our data open are:

Wider social and economic benefits through innovative use of the data

Delivery of improved public services through public bodies making use of the data

Accountability and transparency of delivery of our public services

Improved public services and transparency
Making data that the public sector holds open enables communities and individuals to understand more about public services, gain insight into their own community and contribute to future design and delivery. Studies by the EU have identified that people are most interested in gaining access to information about their local area.

Making data more accessible can also help those delivering public services make better use of data themselves.

Innovation
Publishing our data in a reusable form empowers others to use the data for new and exciting purposes. Open data can be used in isolation or with other data to provide new insights or to develop apps which allow people to use the data in useful ways. Examples of this include the publishing of transport data which has allowed the development of apps which allow users to better get around.

Audiences
In making data open we need to take account of the different audiences who will be accessing the data and the different needs and skills they have.

- Seeker – People who need the data presented in an easily accessible and digestible ready to use format, such as visualisations or tables
- Enthusiast – People who lack formal skills but will try using available tools
- Specialist – People who will use their own skills to analyse and interpret the data

There are a number of pieces of legislation which require public bodies to publish data. Existing legislation includes:
- EU INSPIRE Directive,
- EU Directive on the re-use of Public Sector Information (PSI Directive),
- Freedom of Information Scotland Act (FOISA)
- Environmental Information Regulation

But we do not need to be limited by these requirements. Taking a proactive approach to the publication of data offers benefits to both the organisations publishing data and those who will make use of it. In doing so we can take advantage of the increasing volumes of data being created. We are using digital technologies to deliver services e.g. sensors which make lighting sensitive to the presence of pedestrians or to measure traffic flow. Such data, some real time, can provide insight for wider city management.

Section 2 –Progress towards Open Data

Making public data available is not a new concept, however, making the data available in such a way as to allow it to be easily re-used and shared, is for many a new way of approaching the publication of data. There is a growing International Open Data movement which recognises the value of data and seeks to encourage both the publishing of it and the innovative use of data.

In 2011 the European Commission published its report on Open Data[2] which identified open data as an "engine for innovation, growth and transparent governance".   In addition the report recognised that Open Data presents an opportunity for citizens to better engage with the public sector and the development of policies and delivery of public services. In 2013 the G8 members signed an Open Data Charter, which set out 5 Strategic Principles to support the release of data. The 2013 Open Data Barometer report[3] sampled 77 countries and found Open Government Data policies and programmes in over 55%.

Data.gov.uk provides access to a huge range of data sets ranging from statistics to spend information.   Across Europe a number of regions and cities have been developing data portals which provide people with a convenient way to access data sets in one place.

Within Scotland we have made significant progress in publishing spatial data driven by the EU INSPIRE Directive, which sets out requirements to publish certain spatial data in a prescribed format. Public sector organisations across Scotland are publishing their key spatial data sets and making the metadata available through the Scottish Spatial Discovery Infrastructure. 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 any spatial data created by local government in a consistent form.

Scotland's Environmental Web (SEWeb) provides a gateway to environmental information published by a range of bodies involved in protecting and improving Scotland's environment.

In local government, local authorities including Glasgow, Edinburgh and Aberdeen City have been working to make much wider sets of data from their cities open and available for re-use. In South West Scotland the Crichton Institute is leading on a regional approach. The publishing of data has been supported by civic engagement activities where participants take part in competitions to come up with innovative uses of the data and/or to develop apps using the data.

There is also a growing interest in Scotland from individuals keen to understand more about the areas they live in and/or to make use of new data being made available by organisations. Groups such as the Open Knowledge Foundation are helping to build up a growing community of enthusiasts.
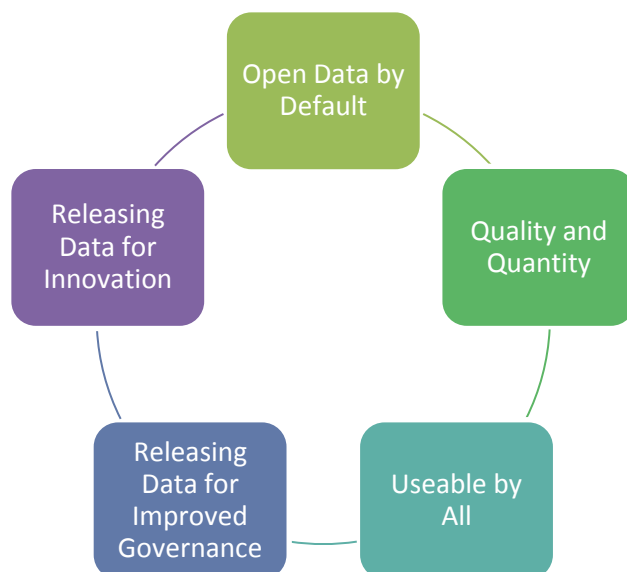
---

[2] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF
[3] http://www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf

Section 3 – How will we make our Data Open?

This strategy adopts the G8 **Open Data Principles**, set out in the diagram below. We will follow an Intelligent Approach to doing so.

Open Data by Default – Those holding public data should make it open and available for others to re-use. Those collecting new data should ensure that the opening up of the data for re-use is built in to the process so that Open Data becomes part of the business process.

Quality and Quantity – The amount of public data we own is huge but the quality of that data will vary. Published data will be supported by metadata so that consumers of the data understand it and are aware of any limitations within it. We will seek to release data in a timely and frequent manner.

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

Releasing Data for Improved Governance – We will release data which supports delivery of better public services. We will use our data to improve the services and policies we deliver. We will seek through the release of data to better inform and engage with citizens.

Releasing Data for Innovation – We will encourage and empower others to make use of the data we release to develop new products and services, for non-commercial and commercial use, which will create wider economic and societal benefits. We will encourage use of the data in education to increase awareness and participation and inspire a new generation of data users.

To support realisation of these principles we need to consider the following:

Open data by default	➡	Organisations' open data plans
Quantity and quality:	➡	Improve the quality of the data
Useable by all:	➡	Discoverability of Open Data
	➡	Accessibility and Usability of Open Data
Releasing data:	➡	Identification and Prioritisation of Open Data

We also need to address:

Exceptions where it is appropriate to charge for Open Data

Licensing data for re-use

Organisational open data publication plans

To realise the vision of a Scotland which makes its public data open, we need all organisations to develop and implement their own Open Data publication plans. An open data publication plan will set out the organisation's commitment to making its data open and identify how and when they will achieve this, including the data that will be made open. Organisations should identify core data which they plan to make open (this includes one off data sets and time series data sets which will be published on a recurring basis). Where appropriate organisations may link to existing Freedom of Information publication schemes.

National Action - We will develop a resource pack which includes a suggested template to support production of a plan.

Quality

The range of data held by organisations is very different and the quality of the data will also vary. We will seek to ensure that the data which we publish is of a high quality and is comprehensible and accurate. Where there are gaps in data we will clearly articulate these and encourage feedback mechanisms to help further improve the quality of the data. Where we are providing time series data we will seek to ensure that we provide this on a continuing basis. This will be a large task for many organisations and will potentially require new processes to be built in to achieve this.

National Action - We will ensure organisations can access training to assist them in recognising the value of data and improving the quality of it.

Discoverability

In publishing our data we need to ensure that we do so in such a way as to make it easy for others to find the data and to understand what is contained within the data. Providing accurate **metadata** with our data is essential. Metadata provides an opportunity to enhance the value of the data and provide clarity on what the data is and what it is not, any limitations to the data and/or its use can also be included within the metadata. Providing good metadata can support those using the data in not only finding it but also bringing together other data sets which relate to it, for instance data consumers can search for data by theme.

Organisations in Scotland are already publishing data on their websites or within dedicated data stores and some are already working together on regional and

thematic publishing platforms. Organisations, in particular those with smaller volumes of data, should consider making use of existing publishing platforms or collaborating with others.

To support greater discoverability of data we will seek to establish a Scottish Data Discovery Site. This will provide those searching for data with a common entry point to all of Scotland's Open Data. The site will link to existing websites and data stores and to those which are developed in the future. Through the provision of good metadata when organisations publish their data, the site will allow people to search for data in a number of ways (e.g. by organisation, by theme).

> National Action - We will develop a resource pack which includes guidance on creation and use of metadata.

> National Action – We will establish a Scottish Data Discovery Site

Accessibility and Usability

To support reuse of data it is important to recognise that data can be made available in different formats. There is a universally recognised 5* schema proposed by Sir Tim Berners-Lee for the deployment of Open Data[4]:

★   available on the web (whatever format) but with an open licence, to be Open Data

★★   available as machine-readable structured data (e.g. excel instead of image scan of a table)

★★★   as (2) plus non-proprietary format (e.g. csv rather than excel)

★★★★   all of the above plus, use open standards from WC3 (e.g. RDF (Resource Description Framework) and SPARQL (SPARQL Protocol and RDF Query Language) to identify things, so that people can point at your stuff

★★★★★ all of the above, plus link your data to other people's data to provide content

We will seek to publish our data as 3* and work towards where appropriate, and there is demand, offering it in higher formats.

> National Action - We will develop a resource pack which includes guidance on format and available technologies.

Identification and Prioritisation

In making data open we seek to do so in an intelligent and managed manner. To do this we need to identify the data which we already hold. Then we need to prioritise for release.

Engagement with potential users of the data (including citizens, community groups, industry, academia and other public bodies) will help to identify data which is of most of interest. Such engagement will also inform the importance to users of the format in which data is released.

> National Action - We will establish channels of communication to identify key data sets and the types of data which users are interested in and thus in which organisations should work towards making open.

---

[4] http://www.w3.org/DesignIssues/LinkedData.html

Charging for Data

Our starting point is organisations should not charge for data. Open data should be made available at no cost to the user. There are a small number of exceptions to this, such as payment for additional services, licensing restrictions and trading funds (e.g Registers of Scotland). Where the payment is for the provision of an additional service which adds value to the data, such as analysis or computation of the data, any charges for service are out with the scope of this strategy and should be in line with an organisation's own policies.

Organisations should make the raw unprocessed data available at no cost. Some existing data may have licensing restrictions which require payment for use of data. When creating, commissioning or buying new data organisations should work on the principle that the data will be made open and available for re-use at no cost. The EU Re-Use of Public Sector Information Directive was updated in 2013 and is due to be transposed into UK legislation in 2015. It introduces the principle that charges for re-use of data should be set at no more than marginal cost, with exceptions in certain circumstances (such as trading funds). It also sets out that there should be transparency around any charging mechanisms which exist.

> National Action - We will promote awareness of the Re-Use of Public Sector Information Regulations 2015, to ensure organisations are aware of them and what they need to do to comply with them.

Licensing

When publishing data organisations should ensure that they make the data available under the Open Government Licence (OGL)[5]. This will ensure that others can make use of the data. Where it is not possible or appropriate to use the OGL then organisations should seek to provide clear licensing information to potential users of the data, so they are clear as to how they can use the data.

> National Action - We will develop a resource pack which provides guidance on licensing.

By considering each of these issues, organisations will be well placed to make their data open in an intelligent manner. Making data open will invite feedback from those who use the data.

Receiving feedback on the data we make open is valuable. Feedback and engagement can:
- Improve the quality of the data
- Provide insight in to the data
- Bring wider groups in to the design of public services

Organisations can mitigate any perceived risk and improve the feedback they receive by ensuring that when publishing open data that they provide clear and accurate metadata, which states what the data is and any limitations it may contain.

---

[5] http://www.nationalarchives.gov.uk/doc/open-government-licence/

Section 4 – When will our Data Open?

A number of organisations are already making some of their data open. To realise our vision and to meet the growing interest from users we encourage all organisations to have an Open Data publication plan in place and published on their website by December 2015. Organisations currently publishing data in a format which does not readily support re-use, should within their plan identify when the data will be made available in a more re-usable format. The ambition is for all data by 2017 to be published in a format of 3* or above.

Measurements
The Digital Public Services Measurements and Benefits Framework will help us to assess the progress in making more of our data open. The Framework contains 16 measures which provide coverage of the main benefits which Scotland's Digital Future – Delivery of Public Services Strategy aims to achieve. The open data measurement is:

- More Open Data sets are available for use by our citizens, businesses and research organisations (Benefit 7)

Measurement of this benefit will be undertaken during 2015. We will capture from organisations the number of data sets made open, the format of the data (its star rating) and where possible the frequency of download. There is no current baseline so this exercise will establish one and allow us to track progress.

Through the proposed Scottish Data Discovery Site we will collect feedback from users to understand how they are using the data, to help us measure the impact of using the data.

National Actions
To support organisations implementing this strategy the following actions will be undertaken at a National Level, as highlighted in section 3:

1) **Development of a resource pack to support this strategy**. The resource pack will be a living document which will include guidance (range of areas including; publication plan template, creation and use of metadata, format and available technologies, licensing), case studies and examples of best practice. Promote use of the resource pack through workshops/webinars. Development will commence in January, with the resource pack being developed during the first half of 2015, followed by promotion of it.

2) **Provision of a procurement framework offering training**, tailored for different roles in an organisation (this would include recognising the value of data, improving the quality of data). Framework to be in place by early 2016.

3) **Scope out and establish a Scottish Data Discovery Site**. The site will provide one place for individuals to go to, to access open data from across Scotland. Scoping to be completed by July 2015. Site to be in place early 2016

4) **Undertake engagement with potential users of open data to understand the types and themes** of data which would be of most use and should be prioritised to make open. Identify opportunities to undertake regional activities/events to promote re-use of data within communities. Where possible this will be done through working in collaboration with existing civil society organisations. Initial engagement to be completed by July 2015.

5) **Promote awareness of the Re-Use of Public Sector Information Regulations 2015**, to ensure organisations are aware of them and what they need to do to comply with them. To be undertaken following publication of the UK regulations (anticipated March 2015)

This publication is available at **www.gov.scot**