

Birmingham and West Midlands Localised Guide for Open Data

European Community Share-PSI Best Practices

Version 1.0 published in July 2016



WMODF

Table of Contents

1.0 Introduction and Background	3
2.0 Principles and Benefits.....	3
3.0 Releasing Data	5
3.1 Targeting Data to release	5
3.2 Easing release process	6
3.2.1 <i>Building trust</i>	6
3.2.2 <i>Automation of data release</i>	7
3.2.3 <i>Training</i>	7
3.2.4 <i>Licensing</i>	7
3.2.5 <i>Standards</i>	8
3.3 Using Open Data	9
3.3.1 Cross Sector	9
3.3.2 Businesses	9
3.3.3 Community.....	10
Appendix A - Example open data proposal paper	11
Appendix B – SharePSI	13
Appendix C – Data Release Form	14

Prepared by

Digital Birmingham
West Midlands Open Data Forum

1st June 2016

1.0 Introduction and Background

The Open data agenda has been growing for some time across the EU community member states, with some cities and countries spearheading the agenda whilst others are playing catch up or waiting to learn from the experience of their counterparts. This has resulted in national and local institutions publishing vast amounts of information and documentation on the opportunities that open data potential offers. In amongst these publications, there are valuable case stories and practical examples on how these institutions have progressed their agendas. However, much of this is underutilised and often the opportunity to learn from it is lost.

In recognition of this and the hyperbolic growth of this agenda, the EU SharePSI project established a collaborative platform to develop and publish an agreed number of key Best Practices (BPs) to support cities and European countries to develop their open data strategies, policies and action plans. Birmingham City Council through its digital team, Digital Birmingham has worked with the SharePSI community in developing the BPs and utilising them in order to create this localised guide.

This localised guide has been developed in partnership with key stakeholders and activists in this region and references how the BPs have supported its development.

In particular this guide has been prepared by [Birmingham City Council](#) (BCC) in cooperation with the [West Midlands Open Data Forum](#) (WMODF) in order to add an English Midlands perspective to the guides established internationally (e.g. Open Knowledge Foundation [Guide](#)) and nationally (e.g. Open Data Institute [Guides](#)). A new local Government organisation, the West Midlands Combined Authority has recently been formed and it is hoped that this guide can support the establishment of open data as a tenet of this new public sector organisation. A local social enterprise [RnR Organisation](#) is undertaking an Open Data Commission with the aim of highlighting local expertise and experiences.

2.0 Principles and Benefits

What is Open Data?

For purposes of clarification open data is defined by the Open Knowledge Foundation as “Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike.”

As stated on the Open Data Institute [website](#), open data has proven benefits for our economies and citizens. The [Shakespeare Review](#) identified £6.8bn of total value in UK public sector data, a report in 2011 estimated that the EU market for public sector data would grow [to €40bn per year](#), while McKinsey estimated a global market powered by open data from across seven sectors would create between [\\$3tn and \\$5tn](#) a year.

A key principle is that all public data should be accessible openly, within the limitations of personal and commercial confidentiality. This should also apply to data generated and held within public sector contracts.

For example Birmingham City Council’s policy on open data (v3 2014) includes the following statement:

“The provision of open and linked data features significantly in the Leader’s Policy Statement (LPS) issued in June 2014 and the subsequent action plans. The availability of timely and good quality open data is required to support a number of key initiatives, including the Smart/Digital City and the devolution of services to communities.

The LPS states that the Council will establish organisational arrangements and funding streams for the Smart City Commission Board, to begin delivery of the Smart City Roadmap in the next three years. This includes a commitment to procure and establish an open data portal and identify and start the release of open data sets.¹⁴

Open data offers opportunities for local councils to work in more efficient ways and to stimulate economic activity. As the council is looking to new models of service delivery, services should position themselves to take advantage of opportunities as they present themselves e.g. when undergoing major service redesign and arising from major systems development.

The sharing and open access to datasets using common technical and data standards amongst public sector organisations will turn data into helpful information and business intelligence. It also has the potential to accelerate business growth and increase entrepreneurial opportunities. And it has the potential to encourage social innovation by citizens and communities.

To quote Birmingham City Council’s [open data policy](#): “A Smart City is not measured by the amount or type of technology it deploys – it is a measure of how citizens, business and communities engage and utilise it to achieve their goals and objectives”

The City Council’s review of Information, Communication Technology and Digital Services (ICT&D) as part of its Business and Budget Plan 2016+ includes the following extracts:

- Change in ICT&D technologies are rapid and growth in the uptake of the latest, emerging technologies such as mobile devices, social media, and high-speed broadband together with the use of open data present huge opportunities for the Council, our communities, citizens and suppliers.
- Potential benefits from ICT&D include access to and exploitation of a wealth of available ‘Open Data’.
- As part of the Future Council Programme we will deliver an information management strategy (aligned to corresponding strategies and created utilising the governance framework) that recognises a one council- single version of the truth for identity and address. The Information Management Strategy will describe how ‘Open Data’ will be used to support our digital and creative businesses and how in the future we will work with community based social media to open up decision making and policy debate.
- The initiatives of ‘Open Data’, intelligence and data transparency are in line with Central Government policy; more public data is being made available online, which includes publishing information about service costs, contracts and plans.

There are similar commitments by other local authorities in the region e.g. the open data strategies and projects featured in Wolverhampton’s [digital strategy](#).

Working in partnership is the key approach to releasing and utilising public sector open data, determining the data to release requires working closely with citizens and the business

communities. Fostering this relationship will then increase the uptake and usage of the data allowing the benefit of transparency, decision making and economic growth to be achieved.

This was the key finding of the [final evaluation](#) of the [Birmingham Data and Skills Hub](#) - *“The most important lesson from this project is that the greatest assets to the open data movement are the individuals and local organisations who can use data to achieve their aims. The interactions with the user community were vital for the success of the project and should become the norm for future open data work”.*

A [node](#) of the Open Data Institute (ODI) has been established in Birmingham with the support of [Innovation Birmingham](#). This demonstrates the local commitment to exploit some of the opportunities of open data.

It is recommended that the city council explores working with the ODI node to further develop its processes and procedures within a framework for the release of data and links to start ups.

3.0 Releasing Data

3.1 Targeting Data to release

The city should work with interested parties to develop the ‘open by default’ policy and practice and while that is being developed, the following aims can be adopted:

A body looking to release data can look to a number of sources to determine data to release:

- Public bodies can use Freedom of Information requests as a simple method to determine demand, respond by publishing the data for others to use.
- Setting up an online forum to capture demand for example:
<https://birmingham.dialogue-app.com/open-data-for-birmingham>
- Auditing existing publications of data that form part of statutory returns to central Government
- Auditing data published on the website used by the public sector body where the data is currently only available in PDF and PowerPoint reports
- Work with internal stakeholders who are champions of open data and therefore are keen to see their data published
- Work with local lobbying groups who exist to support releasing data and actively engage with their local communities, in the case of the West Midlands Open Mercia and the West Midlands Open Data Forum
- Building in open data release to projects as part of the funding agreement e.g. Birmingham is beginning to release Transport data as part of the [OpenTransportNetwork](#) EU funded project.

The body can also identify relevant and appropriate training to address the different needs officers and members have around open data.

SharePSI best practice: [Develop an Open Data Publication Plan](#); [Identify what you already publish](#)

3.2 Easing release process

The task of releasing data is not always smooth or uniform but there are tools and techniques that can be followed to ease the process, build trust and speed up release.

These processes need to support the organisation's policy and strategy as in the [Birmingham example](#). Birmingham's [open data strategy](#) includes a section "How We Will Do It" that outlines the use of a data platform, extraction processes and the use of standards.

As part of Open Data Day 2016 - James Cattell started a [crowd sourced guide](#) to releasing dataⁱⁱ. It refers to:

- How to find data (audit your organisation, ask data users, catalogue what you find)
- How to prepare data (remove personal information, remove confidential information e.g. commercial information, data affecting national/local security or enforcement)
- How to check for other sensitive fields e.g. comments, free text, hidden fields
- How to ensure the right license is in place (3rd party data, licensing)
- How to best describe data (choosing the right file name, linking data, providing metadata)

3.2.1 Building trust

To ensure the organisation is willing to release open data it is important to identify key stakeholders and bring them on board. This can be done by:

1. Identifying data controllers, processors and subject matter experts
2. Working with such data holders to agree a sign off process incorporating a form determining the metadata, licensing, field description, and a report highlighting the appropriate aggregation level to release data ensuring anonymity for any personal or sensitive data.
3. An example of such a form can be found in Appendix C

To understand how to ensure anonymity, Birmingham has developed a process for location based data:

1. Take raw data in an internal legacy system, where there is a location address
2. Determine with the data owner if the release of the most granular data is sensitive
3. Run the data through a series of aggregation processes, placing the addresses in ONS statistical areas, Wards, Districts or the whole city area.
4. A report is created that determines the number of areas where there are 5 or less locations (i.e. individual instances). Therefore the locations could be triangulated which could make individuals / specific assets identifiable (e.g. homes with vulnerable inhabitants).
5. With agreement of the data owner, the data can be released at the appropriate aggregated level; this mitigates the fears data owners may have over useful data that could be seen as sensitive

6. The automation process then uses this aggregated process to regularly extract and load data onto an open data platform or area

An example of the output of the above process is here:

<https://data.birmingham.gov.uk/dataset/social-housing-aggregated>

SharePSI best practices: [Dataset Criteria](#), [Provide metadata](#), [Provide descriptive metadata](#) and [Enable quality assessment of open data](#)

3.2.2 Automation of data release

Automated publication is the preference to keep data up to date and minimise the impact on business resource, but crucially to ensure that data is maintained and kept relevant.

However this is dependent upon internal ICT approaches and skill levels. If this is not possible then the relevant information owner can upload the data themselves.

The use of Transport data is best served by automation in order to meet the demands of the quantity of the data, the fact that much of the data can be real-time and also it is easily anonymised with existing business models in place that can exploit it. Nationally the Transport Systems Catapult is collating the cataloguing of available datasets: <http://imdata.co.uk/>

Birmingham is using a combination of data extraction tools and geographical information tools to automate the publishing of data and transforming it into something a little more useful. As an example Birmingham is publishing the data from the car park availability signs to the open data platform. Clicking on the map button visualises the data:

<https://data.birmingham.gov.uk/dataset/birmingham-parking/resource/bea04cd0-ea86-4d7e-ab3b-2da3368d1e01>

This is an example of good practice and should be part of the Open by default process.

SharePSI best practice: [Open Up Public Transport Data](#)

3.2.3 Training

There are already a number of free courses and tools available on the web to learn about open data. In addition, many local councils offer access to internal eLearning portals that often include courses on information governance, business analysis and other complimentary skills to make the most of data. These could be used for self-learning or in a more structured approach e.g. through a training needs analysis for the whole team. Example resources:

- Free: [European Data Portal's eLearning programme](#)
- Paid for: [ODI open data courses](#)

3.2.4 Licensing

The open data release strategy should specify the aim of releasing data under as open a licence as possible to encourage uptake and reuse. The Open Data Institute provide excellent advice on this: <http://theodi.org/guides/publishers-guide-open-data-licensing>

Working with spatial data in the UK will inevitably mean dealing with Ordnance Survey derived data. Fortunately the OS have moved to a more permissive model and open licensing in the last few years. There are many examples of data being released by different bodies setting the precedent of what can be released. In addition a public body can apply for OS exemption to confirm the data they are releasing does not breach any Public Sector Mapping Agreement Rules.

Comprehensive information can be found here: <https://www.ordnancesurvey.co.uk/business-and-government/help-and-support/public-sector/guidance/index.html>

SharePSI best practice: W3C Best Practice 5 – [Provide data licence information](#)

3.2.5 Standards

Ideally any open data release will adhere to data standards appropriate to that sector - this encourages reuse and uptake, in particular supporting citizens and businesses to develop tools and techniques that cross geographical borders.

However each sector is at different stages of development, in addition to this being a fluctuating area of data development, but this should not be a barrier to releasing data. Where no standard exists, release the data and then attempt to work with industry bodies and standardisation projects (e.g. funded by Innovate UK or EU funds) to define standards.

It is worth investigating a number of areas for UK and international data standards:

- Local Government Association, <http://opendata.esd.org.uk/>
- City level data is moving forward under the smart cities banner: <http://www.bsigroup.com/en-GB/smart-cities/Smart-Cities-Standards-and-Publication/PAS-182-smart-cities-data-concept-model/>
- City comparison data - <http://www.dataforcities.org/>

SharePSI best practice [Standards for Geospatial Data](#)

3.3 Using Open Data

3.3.1 Cross Sector

The data required by public sector organisations originates in many different organisations; therefore in order to release as much data as possible a cross sector approach collaborating between many different actors is required.

Many organisations across the EU are choosing to use CKAN as the platform to release their data, or to record their metadata. The advantage of this is that the information can be downloaded at the granular level of the local city region or aggregated with neighbouring datasets across the country or even the EU.

The principle of this was established as a result of the EU INSPIRE directive where combining and sharing datasets across borders and across sectors would allow the analysis of data to tackle the environmental challenges facing the EU area. Within the UK a variety of tools were developed to help ease the publishing of data. BCC uses “MISO portal” an online tool from Dotted Eyes that takes the data (spatial) and also captures the metadata (also in a standard format, UK GEMINI v2) automating the upload to data.go.uk which then federates to the wider area of the EU data portal.

To help further standardise the publication of data the ODI is acting as the “national” guide for the UK. They publish guides and maturity models here: <http://theodi.org/guides>, which guide various organisations in the principles of releasing data, covering a variety of topics such as:

- Licensing
- Metadata technical standards
- Business cases to facilitate releasing data

Information and research on Data maturity models for the UK social sector is being done by Data Evolution, a project being undertaken by Data Orchard CIC and DataKind UK, and their latest report (2016) can be found on the DataEvolution.org.uk website.

The city can work with the local ODI node and other appropriate organisations to develop and implement these models.

SharePSI best practices: [Develop and Implement a Cross Agency Strategy](#), [Develop a federation tool for open data portals](#) and [Standards for Geospatial Data](#)

3.3.2 Businesses

Supporting businesses to utilise the data released is a key area of ensuring take up, understanding and maximising the value of the data. The well-established practice of running hack days and data dives helps to tease out problems in the data, capture the needs of the business community and build relationships that can ensure the right data is being released.

Within the UK, the establishment of the ODI Nodes is used to share best practice, foster commercial relationships and create a focal point for the local region. Birmingham has an ODI Node based out of the [iCentrum](#) building, supported by Innovation Birmingham in order to encourage the incubator companies to build business models around open data.

On behalf of the businesses the lobbying of public sector to encourage release of certain datasets can be taken on to a certain extent by the ODI Node. To date this has been done primarily by WMODF and by Open Mercia where a range of people beyond business interests come together to work with data. A key part of the open data ecosystem is having groups and organisations that represent the different users and Birmingham has established many groups.

At a national level the wider ODI and the Future Cities Catapult are actively lobbying and promoting the use of open data from an economic perspective e.g. <http://theodi.org/guides/engaging-reusers>. Central Government is looking at action plans for each department around the release of open data and DEFRA has committed to release over 8000 datasets by mid-2016. It needs to be noted that many local councils are still more driven by transparency aims than economic benefit, when opening up data.

To make open data work for a business, a large market is required. This means that comparable data needs to be available for large geographies so it can be combined for national or regional services. Therefore cities need to establish data release standards that allow developers to match data and create new services and products utilising the same data with the potential of a large customer base.

More work is still to be done on growing the skills base when working with data; this will help foster greater business growth. The city needs to work with the ODI node to develop a relevant curriculum.

SharePSI best practice: [Support Open Data Start Ups](#) and [Establish an Open Data Ecosystem](#)

3.3.3 Community

At the local level it is the community interest in openness and transparency that takes precedent around open data.

The West Midlands Open Data Forum brings together local activists with advocates within public sector bodies to help build the demand and the business cases, and to convey the pros and cons of releasing data. The forum encourages honest, frank and open discussions that recognise the views of both those wanting data to be released, and those organisations that have licensing or real world examples of where releasing data can create dilemmas - these are discussions worth having.

The two well established measures of community demands of data release are the Freedom of Information process and the setting up of online forums where requests can be made. The latter is often implemented as part of an approach to Smart Cities to help gauge the public desire to change how services are delivered and identify new ways that communities can be involved. Birmingham City Council uses the [Birmingham Speaks](#) website for such discussions.

SharePSI best practices: [Enable feedback channels for improving the quality of existing government data](#) and [Establish an Open Data Ecosystem](#)

Appendix A - Example open data proposal paper

The following form has been adapted from the proposal kindly developed by the West Midlands Fire Service

Business case - release of Open Data Set

Proposal

To take information currently held by **[Insert organisation name]** and make it publicly available as a public data set. This would enable third party organisations, and individuals to take our data and analyse or use it as the basis to create new products. Any new products that are created could be as result of presenting our data in a new form or through linking with data provided by other organisations.

What is Open Data?

For purposes of clarification Open Data is defined by the Open Knowledge Foundation as “Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike.”[1]

Define the Data Set

[Insert definition of proposed data set]

Next Steps

We propose to take the identified data set and ensure that it does not contain information that can identify individuals or any commercially sensitive data. The data set will be produced in an accessible format and will be made available on our website.

The dataset will be made available under **[Insert name of license]**

Further information can be found and accessible formats [2], and different licenses [3] can be found via the [Open Data Handbook](#) and [Open Data Commons](#)

Potential benefits

1. The organisation would benefit reputationally from opening up data – fitting with the general trend towards open data and transparency, and this would be publicised widely on social media and other channels.
2. There is an organisational benefit from working with people from other agencies, making links and fostering cooperation.
3. There is a potential benefit to analysis and business processes from further linked analysis of this data set.
4. There is a potential benefit from new ways of doing things, as ideas and themes developed when the dataset is released and shared may reveal opportunities for innovation.

Potential risks

1. There is a possibility some of the data will not exactly match data already released publicly and inconsistencies will be noticed. The likelihood is low and the impact is not significant – it is mitigated by adding a comment stating the data may contain inaccuracies.
2. There is a financial cost risk in releasing this dataset. The financial implications relate to costs of hosting the dataset, and maintaining it to ensure that data is current. This risk can be mitigated through a number of free hosting resources if it is deemed too costly to host ourselves. The risk of relevance can be mitigated through clear labelling of the data and setting a review timetable.

Governance

Accountability for this dataset will sit with **[Insert name of body within organisation]**. On a day to day basis this dataset will be managed by **[Insert job role of individual]**. Use of this data set will be reviewed **[Insert review period]**.

[1] <http://opendatahandbook.org/guide/en/what-is-open-data/>

[2] <http://opendatahandbook.org/guide/en/appendices/file-formats/>

[3] <http://opendatacommons.org/licenses/>

Appendix B – SharePSI

<https://www.w3.org/2013/share-psi/>

Share-PSI is a pan European network offering advice on implementation of the [European Directive on the Public Sector Information](#), better known as the *(Revised) PSI Directive*.

It [comprises](#) many of the government departments responsible for implementing the (Revised) PSI Directive across Europe along with standards bodies, academic institutions, commercial organisations, trade associations and interest groups. A series of workshops in 2014 and 2015 identified what does and doesn't work, what is and isn't practical, what can and can't be expected of different stakeholders.

Advice is offered as a set of [Best Practices](#). Each of these is a standalone document that is based on one or more case studies presented during the workshop series. Best Practices are supported by evidence of their implementation, details of which are provided in one or more guidance documents that are produced by EU Member States. These are referred to within the Share-PSI network as *Localised Guides*. Each guide is tailored to a specific EU Member State or region, bearing in mind the local legislative, administrative and infrastructural environment in which the (Revised) PSI Directive is implemented. As well as geographically localised guides, other guidance documents exist for specific sectors such as geospatial and business data.

Example of other localised guides here:

https://www.w3.org/2013/share-psi/wiki/Localised_Guides

Where linking to a best practice, use the URI

<https://www.w3.org/2013/share-psi/bp/>

Appendix C – Data Release Form

Open Data Policy Data Sets for Publication on Birmingham Data Factory https://data.birmingham.gov.uk/ Directorate: Place Section: Housing	
Data set name	BCC Housing Stock
Business application/Manual application (e.g. Excel/Access)	Northgate /(SX3)
Description of data and meta-data	The data describes the locations in an aggregated area (Lower Super Output Area): Average Rent Number of Properties Max Rent Min Rent Sum of Rent Total of no. of Bedroom properties Average number of bedrooms per area Max number of bedrooms Min number of bedrooms Total bedrooms Constructed By No Cavity Insulation recorded Number of Residences on each Floor Heating Type Number of properties with no loft insulation recorded Number of properties with no record of a smoke alarm
Data Format	Size (in MB)
Risk Analysis before Publication	
Does the data set include personal data?	
No	
Would release cause revenue loss for BCC?	
No	
If YES, give details:	
Would release cause competitive disadvantage to a supplier / contractor?	
No	
If YES, give details:	
Is the data currently being licensed to third parties for commercial use. If so, on what terms?	
No	

If YES, give details:	
Is all or some of the data owned by someone other than the Council?	
No	
If YES, have we got permission in writing to release the data?	
Give details:	
Have any FOI requests for the data been refused, or would we reasonably refuse to release the data if it were requested by via FOI?	
Does the data fall under any exclusion (not necessarily cost)	
No - Property addresses have been released previously through FOI4275 and FOI10211	
EINA / Ethical Considerations and any decisions made to adjust the data before release	
<p>The housing stock data that describes over 64,000 social housing properties owned by BCC was extracted from Northgate. The resident data expert in BCC advised on the suitability of the data fields/values extracted. As a result:- Obsolete fields have been deleted. - Information has been totalled up in an area to prevent any individual property being identifiable for more sensitive information e.g. which house has insulation To avoid the risk of identifying individuals the data set is being aggregated in the GIS system to Lower Super Output Area level as this has 86% of the areas with more than 5 properties in and will be published as described in the 'Data Example' tab.</p> <p>If all questions are answered 'No', consideration should be given to placing this data on the Open Data platform.</p>	
Is the data to be placed on Open Data platform?	
YES	
Form completed by:	
Date:	
Signed off by:	
Date:	

Note: This form also includes a sample of the data in a separate sheet.

ⁱ Service Development Opportunities Birmingham City Council Open Data Policy 3.0 page 6 of 7

ⁱⁱ Version accessed and available on 20/07/2016