Release of Data
Final Report: Birmingham Data & Skills Hub

Final Report to Cabinet Office
Prepared April 2015

Consortium: Aston University, Boilerhouse, Digital Birmingham, Podnosh, RAWM
## Contents

1. Executive Summary ............................................................................................................. 3  
   1.1 Review of Goals .................................................................................................................. 3  
   1.2 Lessons Learned and Future Plans ...................................................................................... 3  
2. Introduction ............................................................................................................................... 4  
3. Challenges of Open Data in Practice ....................................................................................... 5  
   3.1 Data Extraction and Standardisation .................................................................................. 5  
   3.2 Process Issues .................................................................................................................... 7  
   3.3 Successes .......................................................................................................................... 10  
   3.4 Lessons Learned ............................................................................................................... 10  
4. Needs of Community Groups ............................................................................................... 11  
   4.1 Goals .................................................................................................................................. 11  
   4.2 Skills Gaps and Training: RAWM ..................................................................................... 12  
   4.3 Skills Gaps and Training: Podnosh ................................................................................... 13  
   4.4 User Needs: Data and Analytical Tools ............................................................................ 13  
   4.5 Data Tools: Boilerhouse .................................................................................................... 15  
5. Evaluation and Case Studies ................................................................................................. 19  
   5.1 Case Studies from Open Data Surgeries ........................................................................... 19  
   5.2 Evaluation .......................................................................................................................... 23  
6. Sustainability ............................................................................................................................. 23  
   6.1 Networking and Communication ....................................................................................... 23  
   6.2 Open Data Skills ............................................................................................................... 23  
   6.3 User Communication ......................................................................................................... 24  
7. Conclusions and Recommendations ...................................................................................... 24  
8. Appendices .............................................................................................................................. 26  
   8.1 Briefing ............................................................................................................................... 26  
   8.2 Content of Initial Training Sessions: Community Open Source 101 ............................... 27  
   8.3 Communication Activities ................................................................................................ 28  
   8.4 Data Release Form (completed for release of BCC housing stock data) ............................ 31  
   8.5 Feedback Questionnaire .................................................................................................... 33  
   8.6 Report on data available, visualisation options and innovations seen elsewhere .......... 33
1 Executive Summary

1.1 Review of Goals

The goals of this project, as stated in the original proposal and the interim report, were to:

1. Release new datasets and develop tools that enable regular future data releases to Birmingham’s Data Factory by reusing Leeds Data Mill (LDM) automated publishing tools and establishing scalable data extraction processes in Birmingham City Council (BCC).

2. Support citizens, communities, third sector and support organisations to understand the value of open data. It will provide a sustainable programme of community engagement enhancing the process developed through Social Media Surgeries (Surgeries) in developing citizens’ interest, focusing on developing skills to utilise open data through visualisation tools. Activities will focus on publishing, using and maintaining open data including crowd sourcing and enhancing data literacy and purposeful analytics.

3. Improve accessibility to open data by developing visualisation tools that enable third-sector and community users to understand data without requiring database skills.

4. Act as a catalyst to encourage grass roots engagement, enhance the activity of networks, support social innovation in communities, improve transparency and civic engagement and build trust to co-create programmes.

The technical aspects of dataset release were all successfully accomplished: a database schema for housing data was developed by HACT, and automated publishing tools and scalable data extraction processes were established by BCC. Some process issues within BCC around data release specifically related to data ownership and authorisation processes were difficult to manage. (For more details, see Sections 3.3 and 3.4.)

The community engagement programme was very successful. A large number of surgeries of different types were run, and local activists were involved both as learners and leaders. These activities were strongly focused on the objectives of each organisation: a key question that they wanted to resolve. Those involved in the sessions not only learned about accessing data, they were able to put their knowledge to practical use to achieve their organisation’s goals. (For more details, see Sections 4.2 and 4.3.)

The visualisation tools have been developed and trialled with the user groups. All the groups emphasised the importance of visual representation of results and the need to be able to specify simply and accurately regions of interest. The tool successfully met these requirements (see Sections 4.4 and 4.5 for more details).

Many of the users created blog pages that not only used data to make a point but also showed how they had used open data and analytical tools to create the blog itself. There is a strong motivation to share good practice and to support other groups in improving their skills. (See also Section 6.)

1.2 Lessons Learned and Future Plans

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. It has to be recognized that such interactions must be part of a continuing process: we recommend a local Forum that brings together community leaders, activists and local government to agree on what data should be published and how it should be presented. Digital Birmingham has set an example by establishing the West Midlands Open Data Forum, which could be built on (or replicated) to reach communities. The Birmingham Speaks platform (http://birmingham.dialogue-app.com/open-data-for-birmingham) is inviting citizens to get more involved in the type of data that they are interested in seeing made available, how to make it available and how it could be used.

Skills sessions for community groups should focus on what the user’s goals are and that will drive their interest in data and data skills. Data skills are a journey - if a user don't know how to use a spreadsheet or post something on the internet it's unlikely they will be able to make effective use of open data.

Skills sessions are most successful when individuals are given relaxed unthreatening help from trusted people - the surgeries offer a chance for someone who's already helped them with one problem (social media) to offer them help with another. The social media surgeries are a network run primarily by volunteers, and as such they have resilience and sustainability built into them. The process will become more sustainable through sharing and peer learning, using social media to connect individuals and point people to useful sources of data and analytical tools.

There is a clear requirement from users to be able to specify spatial regions of interest – ‘their patch’ – and this is only possible in a small number of Open Data tools (such as the visualisation tool developed in this project). However, this user need has to be considered in an appropriate ethical framework: there is the potential for individuals to be identified from data if the region defined is too small (contains too few data points).

It is clear that to make greater progress with the Open Data agenda in the future, there is a need to consider both internal and external reuse of data at the design stage of Information Systems to a much greater degree than at present. Also, where IT services are outsourced, as is the case for Birmingham, the requirements for Open Data and transparency need to be built into the initial contracts.

To help encourage local authorities to make more active progress in making their data openly available, there is a need for a clearer rationale of the benefits of transparency to local government itself.

### 2 Introduction

The goal of the Birmingham Data & Skills Hub project, as stated in the Interim Report, was to release new datasets and develop tools that enable regular future data releases for Birmingham’s new open data hub by reusing Leeds Data Mill (LDM) automated publishing tools and establishing scalable data extraction processes in Birmingham City Council.

It will support citizens, communities, third sector and support organisations to understand the value of open data. It will provide a sustainable programme of community engagement enhancing the process developed through Social Media Surgeries (Surgeries) and specific community development training programmes in developing citizens’ interest, focusing on developing skills to utilise open data through visualisation tools. Activities will focus on publishing, using and maintaining open data including crowd sourcing and enhancing data
literacy and purposeful analytics. It aims to act as a catalyst to encourage grass roots engagement, enhance the activity of networks, support social innovation in communities, improve transparency and civic engagement and build trust to co-create programmes.

3 Challenges of Open Data in Practice

The Birmingham Data Factory https://data.birmingham.gov.uk was launched in September 2014 and now contains 20 datasets. Four of the datasets were released under an FOI request. Several of the datasets are not marked with a date (about one quarter) or updated regularly (almost all): the monthly purchase card data (last updated in March 2015) and payments to suppliers over £500 (last updated in February 2015) are exceptions. It is also noticeable that the amount of metadata provided for datasets varies significantly. This illustrates well the benefits of automated data release that this project is working towards: more reliable and richer information can be made available in a regular fashion which improves the interpretation of the data by making it possible to track changes over time and ensuring that the most recent data is available to the public.

The technical issues around data release have largely been solved. The Leeds Data Mill CKAN Sharp tool has been embedded successfully in the GIS system to support publication to the Birmingham Data Factory.

The recommendations of the interim report still hold good. There is a need to provide good context well beyond what is typically provided e.g. in responses to FOI requests. For instance this could cover why we collect the data, how we collect it, and for whom, on top of providing geographic coverage, date range and frequency of update. It is also important for valid comparison of different data points.

3.1 Data Extraction and Standardisation

HACT (Housing Association’s Charitable Trust) were commissioned to develop a draft schema for data on homes. HACT has been working on a number of projects related to housing data for the last several years, including its work on Housing Big Data. Through this work it has developed an appreciation of the range of data that housing providers hold and the formats in use. In the absence of a shared format for the interchange of data, any attempts to process data from more than one organisation inevitably include a substantial overhead to process the data into a common format before actual analysis can commence.

In developing the schema, HACT had regard to the LGA’s guidance on schemas, since Birmingham City Council intends to seek publication of the schema on the LGA website at a later date. This schema is also intended to contribute to HACT’s broader programme of activity on the agenda of data transmission for housing providers. The schema (or a later version of it) should form a component of a more general housing data schema that HACT intends to develop, which will cover data on a broader range of issues related to housing provider activity. HACT will publish the schema under Creative Commons: Non Commercial – No Derivs (NC-ND).

Both due to good practice and practical considerations, in developing the draft schema for collecting data on housing stock, HACT’s approach has been to draw upon existing standards where possible. Where gaps exist in these, HACT has drawn from previous standards. The rationale for this approach is that:
Existing standards have been well researched, clearly defined, widely adopted and data compiled adopting such standards should be comparable between organisations.

Mappings from data held by various organisations to existing standards will be readily available and is a practical approach for ease of obtaining data.

Defining an entirely new data standard without regard to previously adopted standards would be counterproductive, as considerable effort may be required for organisations to produce essentially the same data in a new format.

The Statistical Data Return (SDR) 2013/2014\(^1\) was initially consulted. Unfortunately, the SDR only compiles aggregate information on housing stock, whereas for this exercise information at the level of individual properties is required. Therefore, the National Register of Social Housing (NROSH) data schema (Field Definitions Release 4 V9.7)\(^2\) was consulted.

The NROSH aimed to set a standard for collecting information on social housing at the level of individual properties. The project was developed to the point where housing associations were able to produce information in the prescribed NROSH format, before being discontinued several years ago. The NROSH format was used to inform many of the fields and definitions within the schema; this also represents a practical choice, as many housing management systems still contain mappings in the back-end to allow data output into the specified NROSH fields.

NROSH contains a comprehensive list of fields relating to the individual properties. Upon evaluation, many of these were determined to be unnecessary for the scope of this project, so a filtered subset has been used as the basis for the draft schema produced by HACT.

Additionally, several of the fields asked for the format of the data to be completed using somewhat dated standards, which HACT has taken the opportunity to update.

This schema specifies fields that are not in their own right personal data, but it should be noted that through combinations of information it might be possible to identify living individuals and, *in extremis*, to identify sensitive personal data about them.

For example, if one of the housing providers that published data to this schema was focused on providing housing to a particular ethnic group, or to people with a particular sexuality, then publishing even a list of postcodes for that provider might identify some of their tenants (e.g. if there is only a single house with a single occupier in that postcode). Therefore this would indirectly give information on those tenants’ ethnicity or sexuality.

The schema contains 35 fields relating to address, local authority, council tax reference number, classification (distinct fields for social and non-social rented), type (house etc.), number of floors, block, age, floor space, number of rooms of different types (living, bedroom, bathroom, kitchen), energy rating, and heating system.

---


\(^2\) NROSH Field Definitions Release 4 v9.7, Communities and Local Government, created by T. Thornton, approved 12/12/06.
3.2 Process Issues

Digital Birmingham (DB) as a team has been the informal lead in making the council aware of the open data agenda and its benefits – they have acted as an internal lobbyist and catalysts. Although there is a section that deals with Freedom of Information (FOI) requests, transparency code and liaises with the Information Commissioner’s Office for statutory datasets etc., there is no individual or group officially responsible for Open Data within BCC. Digital Birmingham have progressed the Open Data agenda by the following means:

- Community engagement: Since 2010 DB has held regular workshops and events to engage SMEs and the wider data user community, including cooperation with, Birmingham Science City working groups

  Digital Birmingham co-founded the West Midlands Open Data Forum (WMODF) (with Birmingham City University) in 2013 to build wider awareness and to increase demand for BCC data; 3 events in 2014 were attached to wider regional events. The various events in 2014 attracted approximately 100 people. DB doesn’t have quantitative information on the success of such outreach events in terms of longer impact. However WMODF meetings were a clear contributor to Centro’s successful Oxygen Hackathon that created 3 useful applications for the organisation in 2014 and data and challenge driven innovation in the Heart of England NHS Foundation Trust (Innovation Engine project).

- A recent presentation by DB at the Digital Media Business cluster breakfast meeting about ‘Using Open Data’ in a business context reached over 10 businesses, 3 of which made follow-up contact and stated that they ‘hadn’t thought about using data in this way’ and thought this was really useful input.

- DB created Birmingham Data Factory (led project internally, undertook market evaluation, procurement and implementation).

- Developed processes for the internal Transparency and FOI team in order to train them to upload data onto the Birmingham Data Factory. It was agreed to initially encourage users to use the FOI route to make more data public because FOI request follow a well established process and are internally resourced.

- Digital Birmingham has developed a sign-off process for publishing open data to go beyond FOI requests, as part of this project.

- DB are members of a number of national and European Open Data expert groups, in particular the Cabinet Office Open Data Champions group; EUROCITIES Open Data working group, the Share PSI 2.0 thematic network.

- DB wrote the relevant parts of the Smart City Roadmap and Action Plan for Birmingham’s Smart City Commission. Open Data is a clear priority for this group and DB had made Open Data a cross-cutting priority in its business plan to ensure all funding bids help push this agenda.

- DB supported “Making a Difference with Data”, a conference for 100 public sector staff from health, housing, the voluntary sector and others organised by TheInformationDaily.com, held in June 2014 (http://madwdata.org.uk/events/Birmingham_Launch_Event).

- DB wrote BCC’s Open Data policy and strategy – which is published and agreed. This was part of the process of building a mandate, but it does not yet mean that senior managers have really thought about resource implications, or taken positive actions to achieve the policy objectives. Open Data needs an advocate and resources – there is no
implementation team (all work is currently done voluntarily), although, latterly, the Chief Executive of BCC has advocated the City to take an *Open by Default* approach.

**Effort vs Sustainability**

There is a stark contrast between the timely completion of FOI requests (which must be completed within 20 days, or 40 days if it requires additional time to consider the public interest test) and the work on this project. Production of the final version of the Housing Asset Register took the better part of 6 months lapsed time. Technical feasibility to extract and publish the data has been demonstrated, though a resilient internal sign-off process delayed publication slightly beyond the anticipated timeframe.

The difference between these two scenarios is due to pursuing a long term, sustainable approach with the project. Under the ‘Right to Data’ (Local Government Transparency Code) requests can now be made for *regularly updated* open data. The resources required to ensure that open data, once published, remains current are significantly higher than the resources required to gather and publish a one off data set in the first instance. FOI requests traditionally respond to a set of questions and create a single data set in response, sometimes by pulling together data from different sources. Little explanation has to be given as the data is published in response to a set of specific questions. To ensure that data can be published automatically at regular intervals, without having to ask officers to manually run reports and provide the data, more preparation has to be done. This requires effort and sometimes direct costs for development work. In the case of the Housing Asset register the preparatory work included:

- Requesting access to and review data in the source database
- Verification of the data fields by housing experts to exclude invalid fields
- Creation of a standard Business Intelligence report to generate the required data fields
- Evaluation of the data in context, renaming of field headers into human readable names, removal of obsolete data, data quality checks
- Consideration of privacy issues leading to aggregation of some data to avoid individuals being targeted with nuisance or sales calls
- Formal check by data protection / legal experts
- Amendment of the report to reflect the above
- Production of final dataset
- Production of detailed meta-data

Technically the work also included:

- Pulling a new data layer from the source database (Northgate) into the GIS system. This has been done for security reasons, to avoid direct access of Northgate, which holds personal data of tenants.
- Embedding the automated publication tool (CKAN Sharp) into GIS system
- Publishing final dataset onto Birmingham’s Data Factory
Only after all the above steps have been undertaken can publication be automated. In the long term, this will avoid further work and is much more effective than publishing a snapshot of the data on request against FOI requests.

Because the resources required to automate and maintain accurate open datasets are high BCC have found it difficult to make progress without grant-funded projects. Smart Spaces; Opticities and OTN are EU funded projects, which have and will produce several open data streams for Birmingham. As part of the Smart Spaces project an open API for energy data in public buildings has been built; its data can be downloaded from http://www.energysmartbirmingham.com/

Another organisational issue when dealing with open data is staff turnover. When certain experts leave (e.g. a number of business intelligence officers was lost due to recent service changes) the knowledge about individual datasets, the meaning of fields, how to access the source database, meta-data etc is lost. This increases the resources needed the next time somebody needs to access the data. Hence data management and open data needs embedding into everyday council business.

A move to regular updating public information from databases is the long-term solution, but requires more up-front investment. In the meantime, data ownership linked to permission to publish remains a key issue. Information that the public wants may involve access to multiple databases with multiple data asset owners; each of whom has to authorise the request. Because IT services are outsourced to Service Birmingham there is a requirement for individual owners’ agreement because of contractual arrangements. Therefore DB developed an in-house approval process to achieve release of open data under this project.

A release form has been developed (Appendix 8.4) and process has been tested with the Housing Information Governance Team and Legal Services. This required several iterations of the dataset and the form and will allow the Housing Data to be published.

**Other issues explored**

The wider ICT and commissioning/procurement strategy should require database system providers to provide better interfaces to enable a more human-friendly way of accessing and interacting with data in the first place. Commissioning contracts should take account of reuse needs. Human-Computer Interaction and system accessibility should be given a higher priority when developing or procuring new systems. Contracts need to ensure that data ownership will remain with BCC.

The LGA has also recently published the practitioners’ guide: ‘Local Transparency Guidance – Publishing Data’. BCC could work with staff to use this template and improve the Council’s responsiveness in making data open.

It can be difficult for BCC to respond to data requests because it currently lacks a proper data catalogue. BCC maintains an Information Asset Register (IAR) (c. 350 assets) containing the data asset owner, system and licence in use, title for each asset. The IAR does not contain information about locally held small data bases e.g. Excel or Access files. There is no global catalogue of data stored within each asset making it difficult to identify the data source in response to e.g. one FOI request easily. The matter of database consolidation is currently being considered as part of a contract review with Service Birmingham. Hence when third parties ask for ‘all data’ they don’t realise how much this entails and the work involved. This
lack of coordination makes response to FOI requests difficult to manage, let alone creating Open Data repositories. DB would like to do more in this area because ultimately it would reduce the effort and wasted time to access data internally. However, it is proving difficult to build a convincing enough business case given the short-term financial constraints on the Council.

3.3 Successes
The BCC team have been successful in the following aspects of the project:

- Created a process to publish data sustainably by direct access to the databases underlying information systems;
- Built knowledge in the Service Birmingham GIS team to support the data publication process;
- Demonstrated successful automated extraction and publication of data from a big system (Northgate) and are now looking to extend this to data in the Adults & Communities Directorate (health related data);
- Developed an authorisation process and documentation that considers ethics issues. This document allows the Information Governance Officers to brief the Head of Service and the Strategic Director as required, to gain formal sign off.

3.4 Lessons Learned
Stronger and clearer internal processes are needed to make the policy and strategy a reality. At the moment, too much relies on individual good will and departmental culture. There should be a change of approach so that Open Data becomes the norm, unless there is a genuine public interest in non-disclosure, rather than the opposite case. This will be a similar culture change as was required for FOI. In addition, there needs to be a simpler authorisation process that involves a smaller number of people and stages, particularly for data drawn from multiple systems.

The Open Data initiative needs to be embedded with external, outsourced, service providers as part of their contractual requirements.

Open Data brings greater scrutiny to the data that government holds, as has been noted in the USA. This has the potential of exacerbating three data integration problems that arise before data are ready to be released to the public. First, the problem of inclusion concerns what data are included or excluded from the datasets that agencies index on data.gov. For example, distorted, incomplete, or misleading data could be intentionally released to damage or promote the cause of certain groups.

Second, the problem of confusion refers to the fact that it is almost impossible to rectify bad data once they are published on the web because they trickle downstream to numerous databases that are not under the control of a single entity. In contrast, it is far easier to resolve bad data upstream. For example, almost 2 decades ago, the U.S Social Security Administration

---

3 This paragraph and the next two are based on Peled, Alon. "When transparency and collaboration collide: The USA open data program." Journal of the American society for information science and technology 62.11 (2011): 2085-2094.
successfully created its Master Death File that directly feeds numerous other databases to ensure that the deceased no longer receive correspondence.

Last, the problem of diffusion is created by the fact that Open Data did not offer a mechanism to time-version datasets, to create historical integration points among composite datasets, or to allow data to retire or expire. Knowing how the citizen’s composite virtual self (i.e., data woven across federal databases) appeared to government officials at a given point in time is critical to understanding why, for example, citizens were arrested or deemed as worthy of receiving welfare benefits. Without such time-versioned mechanisms, data.gov is a useless index where one cannot connect datasets in a meaningful way. Open Data architects celebrated the idea that “data, once posted, cannot later be sequestered or taken down”\textsuperscript{4}. This “post it forever and delete nothing” argument ignores the problem of deriving meaning from historical data. The National Archives and Records Administration and the GAO admitted that this problem is risking the data quality of all federal databases\textsuperscript{5}.

From this study, and our experience in this project, it is clear that to make stronger progress with the Open Data agenda in the future, there is a need to consider both internal and external reuse of data at the design stage of Information Systems to a much greater degree than at present.

4 Needs of Community Groups

This section describes how the project team interacted with community groups in Birmingham to meet their needs for release of data, presentation of data and interactive tools to make easy viewing/understanding of data. The results of the work carried out by RAWM, Podnosh and Boilerhouse are then discussed in Section 5.

4.1 Goals

RAWM initially targeted 3 community organisations: Balsall Heath Forum (BHF), Castle Vale Community Housing Association (CVCHA), and Sparkbrook Neighbourhood Forum (SNF)/ACP group (chosen since they were the lead organisations for the DCLG Our Place programme in Sparkbrook). These organisations were involved in two other initiatives – Our Place and BCHCT’s Healthy Villages (Birmingham Community Health Care Trust). It was believed that involvement in this programme would benefit those organisations by augmenting their participation in these programmes.

All organisations had access to a wider group of community activists and were themselves running community volunteer programmes; CVCHA had already run a number of social media surgeries.

Podnosh ran Open Data Surgeries, which are informal gatherings where people who know what they’re doing share skills. ‘Patients’ sit alongside ‘surgeons’, someone who understands good ways to use the internet and who can demonstrate how to use free tools. The aim is to help individuals and communities understand how they can use information and data to address issues that matter most to them. Podnosh initially spoke to 5 community organisations:


Birmingham Open Spaces Forum (who want to demonstrate how much income their many voluntary groups generate in the city); Birmingham Conservation Trust; Birmingham Settlement (understanding the needs of vulnerable older people); Gateway Family Services (provenance of staff); Handsworth Helping Hands (quantifying the financial savings achieved by their volunteer-run street cleaning and recycling work).

4.2 Skills Gaps and Training: RAWM

Initial Discussion with all organisations started in September 2014: BHF were the first organisation to agree involvement, followed by the other two. Because of other commitments there was a delay so the initial meeting with all participants could not be held until the 19th November, though a number of meetings with individual organisations took place prior to this date.

RAWM met with or talked to staff from all organisations and provided a briefing (see Appendix 8.1) for all organisations, enabling staff to report through the appropriate internal procedures Initial discussions outlined organisations’ stated need for statistical information for the specific localities within which the community groups operated, often more focused than Ward/SOA definitions and often required for specific subjects or service areas that may not be available in general data.

What became apparent through these consultations was:

- Organisations’ reliance on publicly available statutory data.
- Organisations had difficulties in establishing formal relationships with, or identifying, public sector (BCC) officers to obtain data relevant to their work.
- The organisations’ lack of appreciation of the value of their own data.
- Organisations lack of awareness of their ability to influence/manage other open source tools that could affect perceptions of their geographic area(s).

Early sessions provided an outline of available open data, such as Land Registry as related to house prices and accommodation type related to community group area. This form of data was not previously known to the groups. These session identified the need for a wider training programme describing open source data tools.

RAWM is a voluntary sector Strategic Partner of the Department of Health, NHS England and Public Health England programme and is involved in the evaluation of a variety of programmes. Its research shows that there is very little training offered to VCS organisations on the uses of data. The majority of support is offered within an expert/recipient relationship, where the VCS is the recipient. This support tends to emphasise the use of data only within funding and commissioning applications rather than as a source of evidence for positive social action.

Additional research for this project identified that there is no training offered to community groups concerning the utilisation of open source tools including open data as a method of adding value and enhancing the organisation’s digital footprint.

---

6 Voluntary and Community Sector
RAWM staff have also worked with Open Mercia, a regional collective of open data advocates, to develop a bank of data scientists who are interested in working with community groups to develop their datasets and provide analysis. This is a long-term investment in developing community organisations’ appreciation of data.

The Release of Data Programme enabled RAWM staff to work with the identified community groups to promote understanding and use of open source tools including open data as a way of promoting their organisations and services.

The 8-step Open Source tools and resources programme, Open Source 101, designed by RnR Organisation as part of their involvement in open source activity, was seen as an appropriate programme to widen VCS awareness and use of open source tools and resources. It helps to develop skills in data identification, collection and presentation. It also starts to address some of the issues identified in discussion with participant organisations. An overview of the initial training sessions can be found in Section 8.2.

### 4.3 Skills Gaps and Training: Podnosh

The Open Data Surgery approach was based on Social Media Surgeries, which are a simple and powerful way to show community groups, active citizens and voluntary organisations how to use the internet to connect, collaborate and communicate. The surgeries build both confidence and skills, and help develop vibrant civic conversations and create social capital in neighbourhoods.

In total, 18 surgeries were run in 3 locations (Central Birmingham, King’s Heath, and King’s Norton and a fourth floating surgery made it easier for Podnosh to help those interested but not able to get to a surgery). There were 23 ‘surgeons’ and 37 ‘patients’. The sessions contained a mixture of social media and open data activities, as driven by the needs of the patients. It is noticeable that was strong support from technically more experienced people to act as surgeons (indeed, they outnumbered the patients at several events and started learning from each other.). This suggests that this approach should be sustainable beyond the lifetime of the project. The work done by the surgeons and patients was shared as blog posts (www.bevocal.org) so that other people running social media surgeries can find them and use them to spread the learning: this also supports sustainability.

### 4.4 User Needs: Data and Analytical Tools

Infographics are a powerful way of communicating stories based on data: one participant stated that “They stay in the mind.” Discussions with community representatives to explore the potential of visualisation tools were kick-started at an event hosting by Boilerhouse in January, where they presented examples of existing visualisations: Where can I afford to live? (http://www.bbc.co.uk/news/business-23234033); public health longer lives map (http://healthierlives.phe.org.uk/topic/mortality); Health and Social Care in the West Midlands (http://www.hscwm.org.uk/); Neighbourhood Planner (http://neighbourhoodplanner.org.uk/).

It was mentioned that there is mistrust of institutional data amongst some community groups so being able to use open datasets and combine them with organisation’s own data is a welcome development. It might be questioned whether the original source of the data in an ‘open dataset’ is always clear (i.e. is it originally derived from a ‘mistrusted’ institution) or whether the organisation presenting the data through a tool or portal is the one that is trusted.
Several groups were interested in layering their own data on top of other maps or combining datasets in various ways. For example, groups at the Open Data Surgeries wanted to be able to link the location of the people they helped with deprivation indices. However, currently available datasets are very unwieldy, difficult for these users to query, and beyond the reach of the computing resources available.

This interest was very encouraging as it shows that the individuals were thinking beyond just using existing data and visualisations and were considering the goals of their analysis. However, currently the tools required to layer data on maps require more advanced programming skills than the majority of users have. Combining data from different sources generally requires the data to be downloaded and then analysed using spreadsheets or similar programmes: again requiring relatively advanced programming or scripting skills.

Another common issue that was raised was the need to be able to specify a geographic region of interest quite precisely. However, much government data is aggregated in areas (e.g. whole local authorities or wards) that are much larger than the neighbourhoods that most organisations are interested. Even when more detailed information is available, it is still aggregated in regions governed by postcodes (e.g. Lower Super Output Areas). Sites such as www.police.uk are ideal since they allow the user to define a polygon and then present time-filtered data on incidents within that region. For example, Figure 1 shows a user-defined region for Cannon Hill Park (with the total number of crimes reported for February 2015). This complex region clearly requires a user-defined boundary.

Figure 1: User-defined region for crime statistics (www.police.uk).
To summarise, what local groups and individuals want is:

- Very local information: the council-wide or ward level data they can access is not useful to them as they are working at ‘hyperlocal’ level.
- Data that is easy to understand and share: existing map-based visualisations can be hard to understand.
- Community groups want credibility for the data they have collected and published.

### 4.5 Data Tools: Boilerhouse

The discussion with community representatives described in the previous sections led Boilerhouse to define a key innovation, which enables users of map-based data visualisations to click on a Lower Super Output Area (LSOA) or a selection of LSOAs and see relevant statistics for just that area. This enables people to look at data for a small geographical area that contains, say, a housing estate or neighbourhood they live or work in or represent. N.B. the average number of households in an LSOA in England and Wales is 672 and the average population 1614 (figures from 2011).

#### 4.5.1 Background

Boilerhouse researched data visualisations carried out by other Breakthrough / Release of Data projects (full report in Appendix 8.6) and also looked at housing-specific visualisations on a wider front. With local authority projects, they found continuing use of traditional GIS tools (which many LAs have invested in). These (e.g. data atlases) and other visualisation tools can be hard for non-experts to use. As an example, consider the following instructions from a recent project:

*This dataset contains multidimensional data (a data cube) which can be displayed as a grid to compare two dimensions at a time. Use the drop-down menus below the grid to choose which dimensions to show as rows and columns (and, optionally, to filter the other dimensions by value).*

Frequently there are no concessions to non-experts in use of labelling (median, percentile etc.) and narratives. Publishers like the Guardian and the BBC, and organisations seeking crowd-sourced information e.g. mySociety ([https://www.mysociety.org/](https://www.mysociety.org/)) tend to create the most user-friendly tools.

They have reviewed the data visualisations associated with previous breakthrough/release of data funding and other publicly-funded sources like LGInform. They have also looked at private sector-funded visualisations from property businesses (e.g. Zoopla), consultants and publishers (e.g. Guardian datablog) and reviewed data from ONS, VOA, BCC, and private sector providers that offer free APIs (see separate attached report for information). On the whole, visualisations that are accessible to all users, including non-data expert community and citizen groups that are the subject of this project, tend to be produced by media companies and private sector providers in the property sector.

In the UK, sites such as [http://www.rightmove.co.uk/](http://www.rightmove.co.uk/) allow people to carry out sophisticated searches on housing and accommodation for sale or rent, based on proximity to a location defined by postcode. This site combines data from multiple sources, including the Land Registry (for prices), and provides such information as a map of local schools, broadband speed etc.
In Germany, there are portals to provide information on apartment rental in specific geographic areas. Figure 2 is taken from Immowelt, a property portal using their own data by district presented in good graphs: http://www.immowelt.de/immobilienpreise/hamburg-winterhude/mietspiegel

Figure 2: Rent comparison and time series for Hamburg.

When creating visualisation tools, 50-66% of the time may be spent accessing, checking and manipulating available datasets. Many national datasets are not available at sufficiently fine granularity for local decision-makers.

4.5.2 System Design

An effective means of making data accessible to people that are not specialists is to think of a story to tell. Boilerhouse idea was to tell a story about the cost of renting property and its affordability – something as interesting to many individuals and families as it is to policy and decision-makers.

They set out to present rental values from private, council and registered social landlords (housing association) properties in different parts of the city along with income data so that people would be able to see the proportion of disposable income used up by renting in different sectors and neighbourhoods.

To tell these stories you often need just two or three key pieces of data (in this case, rental values and postcodes) which may need to be extracted from the very large datasets typically released by the council. ‘Council only’ data on a specific topic may not be that interesting, but can become more valuable when mashed up with data from other sources. For example, our tool enables comparison of council rents with private sector rents and (if we could have accessed the data during the project) social housing rents. Council housing data on rents and
stock is much more open than social housing data: ironically, private sector data may be the most open of all.

For example, our tool enables comparison of council rents with private sector rents and social housing rents. It then mashes up this data with open data about incomes from government sources from DWP available at local level via Nomis. The tool also presents data about housing benefit which is obviously relevant in this context.

As part of designing the visualization tool, Boilerhouse have participated in a meeting with members of Birmingham community groups from Castle Vale, Balsall Heath and Sparkbook (these are areas involved in the DCLG Our Place scheme). They demonstrated some data visualisations that they had previously done, and led a conversation about what sort of access to what sort of data would be useful to groups working with neighbourhoods and communities. Three key issues emerged:

1. Data published by ‘the authorities’ is very often at too large an area (e.g. wards) to be relevant to communities that tend to operate in a smaller geographical footprint.
2. These groups are often unaware of data that is openly available (e.g. LGInform profiles).
3. They collect a lot of data themselves but struggle to present it and get it taken seriously by officials. Therefore they would welcome access to ‘crowd-sourcing tools where they could upload and display data. Bearing these conversations in mind, and our information-gathering on existing data tools, they designed, built and tested a tool informed by these issues.

4.5.3 Data Tool

The tool developed in the project can be accessed at http://birminghamhousingdata.org
Some screenshots of key features are shown in Figure 2, Figure 4 and Figure 5.

Figure 3: Visualisation tool home page.
The design of the tool follows Shneiderman’s information visualisation mantra\(^7\): start with an overview, then allow the user to zoom and filter, with details available on demand. The figures show how the user can successively obtain more detailed information and the control they have over the type of information (Figure 4) and then the spatial region considered (Figure 5).

5 Evaluation and Case Studies

This section contains several case studies of local groups who have been able to answer important questions using skills that they have developed at training events run by this project. Several of them are written up in more detail at http://bevocal.org.uk, a site to help community groups and charities to make the most of data to do something useful.

5.1 Case Studies from Open Data Surgeries

5.1.1 Marcus Belben

Following an accident involving his son, he became more interested in road accidents locally, and then more interested in how communities can pull together to support each other and address local issues. Swan Corner Community group was formed and ‘road safety’ has always been a major concern for the group. The group is campaigning at the moment to include the road in the new BCC 20mph initiative.

Marcus has used CrashMap before, but wanted to know what other data is available. CrashMap is based on data taken directly from the Department for Transport: road-collisions.dft.gov.uk. Another useful site is doorda, and Marcus also looked at data.police.uk and www.police.uk He saw that a critical roundabout is on the border of policing neighbourhoods, so this leads to issues with response. The group would like to modify police behaviour by using data to support their arguments. On www.police.uk, it is possible to define area by hand and get crime figures: this was particularly valuable when dealing with a tightly localised issue. The ‘surgeon’ encouraged Marcus to download the data (in CSV format) so that he could analyse it further in a spreadsheet and then generate graphics. Marcus found it surprising how easy it was to use the tools and to present data in form that would be valuable to his group. The ease of searching for data in a graphical format and generating graphs was also surprising: he had expected it to be necessary to trawl through large lists of numbers and data, but tools made the information he needed very clear.

He writes as follows:

“We have been documenting problems ourselves … but data freely available online has really made the difference to prove the need for change. Thanks to Podnosh and their Kings Heath and Moseley social media surgery, we’ve been able to access all the information we need for this, and pretty much anything else.”

“Police.uk has a particularly useful feature - you can draw around the area you are interested in and it will give you a breakdown of useful stats for any time period, so for my small stretch of Brook Lane:”

Marcus has created an excellent blog post (http://birminghamlives.blogspot.co.uk/2015/03/research-your-street.html) about his use of Open Data.
Figure 6: Graphs showing reported crime in user-defined region.

5.1.2 Suzanna Carter - BCT The Coffin Works
- 09/01/2015: Initial Statement: We’re a new museum: we want to make sure we are collecting the correct data to inform evaluation marketing and understanding of who our visitors are and where they come from.
- 09/01/2015: An understanding that there are already free tools that can analyse post code data and use existing social media to understand our supporters and followers demographic.
And that we’re guinea pigs for Podnosh and happy to be so. It will be useful for BCT to have an understanding on how this can help in all projects.

- 28/01/2015: Today I have learnt about free Google forms. I usually use typeform which I pay for on subscription. I will be able to save money using this tool. Nick tested the usability with two users and we refined. It has brought up questions about how we need to train our volunteers at FOH and also utilise more the iPads we have. We will test both paper and iPad forms and see how we get on. The index of multiple deprivation is a cumbersome dataset that is hard to work with.

5.1.3 Emma Woolf – Birmingham Open Space Forum (BOSF)

- 19/01/2015 Initial Statement: We need to look at money generated by groups - how to collect and collate information from parks with friends groups about income generated. In particular, we want to show how much funding we have brought in to areas of deprivation.
- 19/01/2015: We set up a Google form to start collecting the data we need and had a discussion about whether information should be put is as text or as numbers.
- 10/03/2015: Sourced data from police crime maps and open data communities deprivation maps (ODCDM). We are also making our data more open – working with 10 other groups.

Blog posts:

![Money Raised by 10 of Birmingham's Volunteer Friends of Parks & Open Spaces Groups in 2014](Figure 7: Graph generated by BOSF based on data collected through Google forms.)
5.1.4 Simon Baddeley – Handsworth Helping Hands

- 26/01/2015 Initial Statement: Clear streets and refuse in Handsworth, have a better understanding of who owns which properties and see if there are better ways of organising the data they do have.
- 26/01/2015: a chance to be asked questions about Handsworth Helping Hands to clarify my thinking on why we’re doing what we’re doing and what we get out of it.

5.1.5 Becky Pickin – Dudley CVS

- 23/02/2015 find sources of data: dooda, LG inform. We had a hands on go at forming a natural neighbourhood (within the parameters of super output areas!) and hopefully that will provide useful data for that neighbourhood and community groups which work with people in that neighbourhood.
- The more you drill down, the more you get asked about what kinds of data you’ve got, so it’s important to stay focused and keep it simple.

Video: [https://www.youtube.com/watch?v=RTlnRtlkgJKU](https://www.youtube.com/watch?v=RTlnRtlkgJKU)

5.1.6 Maria Harvey – Birmingham LGBT

Maria has been a Social Media surgeon in the past.
• 24/02/2015 Initial Statement: to find out more about how data could be used to help
Birmingham LGBT centre & projects
24/02/2015 you showed me some examples of what data is available and how it can be used to find
things out. We looked at the idea of perceptions of safety and that the centre isn’t a safe place to be at
night, we looked at how to find the data to see how realistic that
• was. We used http://www.police.uk/west-midlands/BWAA/crime/+3rxDLN/
• 24/02/2015 Since the last Tuesday I’ve thought more about data and we looked at data to
find out about the older population of Birmingham that is relevant to my project. I hadn’t
thought about using specific data as a resource to target our resources.
• Blog posts:
  • http://bevocal.org.uk/2015/03/30/police-crime-data-for-improving-the-perceptions-of-
safety-in-communities/
  • http://bevocal.org.uk/2015/04/01/consistency-and-collaboration-for-data-collection-
around-niche-areas-of-commuunity/

5.2 Evaluation

The richness of data from the community engagement work has been generated through
individual blog posts and informal discussions. An online survey aimed to gather more
quantitative feedback from user participants but response to this has been limited. The task-
specific questions are contained in Table 1 in Section 8.4.

6  Sustainability

6.1 Networking and Communication

A key part of making Open Data accessible and useful is collaboration with networks outside of
Birmingham as well as continuing the successful collaborations with VCS organisations inside
the city. Contact has been made with York City Council, Derby City Council, the London Data
Store, the Trafford Innovation Lab, and Surrey County Council. The value in these connections
is to enable best practice to be shared, resources (e.g. for tool development) to be pooled, and
for greater standardisation (e.g. of data schemas) so that access can be made more consistent
(and therefore easier and cheaper) across the country.

It is important to work with a variety of organizations to ensure that we reach the full range of
people who could make use of Open Data. For example, the Open Data Camp had a better
age range – more in 30-45 band and a few under 30. Hackathons tend to appeal particularly to
younger individuals.

Members of the project team have been active in disseminating our work in other venues: see
Section 8.3. Nick Booth from Podnosh spoke at an Open Data Event in Manchester to explain
the surgeries and has also picked up on existing relationships with those who run the surgeries
in Manchester. The June Surgery (run in part by the Voluntary Services Council there) there will
be an Open Data one.

6.2 Open Data Skills

Open Data surgeries in Birmingham will continue to run offering open data and access and
analysis skills. Podnosh will continue to talk to surgeons about what’s available and what has
been learnt. The strong contribution from a large number of ‘surgeons’ gives us confidence that there is sufficient engagement from the user community to support these surgeries in the future as has been the case with the previous social media surgeries. The funding from the projects has created the interest and support from mentors to widen their existing social media skills to support wider engagement and transfer of skills around open data.

In Dudley we worked with the VCS because they too have started running social media surgeries in their own patch. Currently the additional skills picked up by their staff can be fed into the surgeries. In general, the surgeries grow where people wish to run them. A more universal provision of them would need investment in order to nurture the network. We also need to communicate this work more widely to those running surgeries to encourage them to use the blog posts and stories to offer a starting point for sharing open data skills with the people they are helping.

The surgeries are not based on a teaching model, but a shared learning model focused on problems and available tools. This should make it more sustainable because material doesn’t need formal updating. However, some teaching material, such as that used by RAWM in their training sessions, can be generically useful, and a hosting site or institution for this would help to provide a more permanent home. User feedback on the process needs to be improved, so in future surgeries, the intention will be to ensure that users complete the survey during the session itself.

We should seek to identify opportunities to develop the skills of VCS organisations in data identification, collection and presentation with a regional and national perspective.

6.3 User Communication

Communication between users is another important aspect of sustainability: peer-based learning is a powerful tool in this environment, so that users can see how other people in their position have managed to use data to achieve their goals. Here the blog site http://bevocal.org.uk will have an important part to play.

We should use social media to put the local activists in touch with each other. In the longer term, recommender system\(^8\) approaches could be used to guide users to data sources and case studies that they would find helpful. (Recommender systems are software tools that provide the user with suggestions for new items that they might like based on their (and others) previously expressed interests. They are ubiquitous in modern online interaction, e.g. these systems underpin the ‘Other items you might like’ lists provided by Amazon.)

7 Conclusions and Recommendations

Thanks to the efforts of the project team and Birmingham City Council throughout this project, council data on rents and stock for council housing is now much more open than social housing data. The database schema developed for the project by HACT will be valuable to other local authorities who publish open data in this domain.

Significant time was spent during the project trying to access data about local rents in Birmingham for housing association properties. Clearly this data is essential for the visualisation to deliver its full value. However, this proved impossible to get: we applied to the Department of Communities for access to data from CORE (COntinuous REcording) which collects data from housing associations, but this was not forthcoming. The only openly available data on housing association rents was a single figure: an average rental for all sizes of property across the whole of Birmingham. Clearly this figure has little value for our project. Ironically, private sector data may be the most open of all. The visualisation tool uses an API supplied free of charge by the online property search company Zoopla.

It is clear that for all the users, it is vital to be able to define spatial regions within which data should be reported. There are two technical issues to resolve: one is the spatial resolution of the data that is stored (i.e. is it stored at a sufficiently high resolution to support region definition); the second is the need to develop tools that allow the user to define regions (such as with our visualisation tool or www.police.uk). However, this user need has to be considered in an appropriate ethical framework. As noted in Section 3.1, there is the potential for individuals to be identified from data if the region defined is too small (contains too few data points). There are known ways to address this, either by aggregating to larger regions or by adding noise to the location data so that it is not possible to identify exactly which property is referred to.

Several groups were interested in layering their own data on top of other maps or combining datasets in various ways. Currently the tools required to layer data on maps require more advanced programming skills than the majority of users have. Combining data from different sources generally requires the data to be downloaded and then analysed using spreadsheets or similar programmes: again requiring relatively advanced programming or scripting skills. Future work should make this easier for most users.

At the moment, most users are looking at ‘snapshot’ data at a particular moment in time: there is not much work on temporal patterns (trends, change points) yet. It seems likely that this is due to the fact that most of the users we were working with are new to the use of open data. The need for temporal analysis may increase in the future as people’s facility with open data tools increases and data accessibility becomes more sustained.

As an example of these issues, the area represented by King’s Heath Residents Forum does not match political boundaries. The nature of the streets in the neighbourhood changes a lot based on day of week and time of day. However, the data that the group receives is often very crude and uninformative: for example, when considering accident statistic, one large accident distorted summary statistics. Thus there is a need for users to be able to dig into data.

We identified different types of user: some don’t have (in their words) a ‘data head’. These users still want to be able to search and tag information. Users need both to be able to access existing data but also need to have the tools to analyse this data to determine new insights: in particular, visual impact is very important. More work is needed to make the tools available (such as Tableau Public) more accessible to a wider range of users.

Despite some of these reservations, there are many case studies of success within this project. It is clear that local action groups are on a pathway – the sector is moving in the direction of making more use of data to provide evidence for the arguments that they put forward and to communicate with their members.
The difficulties that the project team had with publishing the housing datasets show that stronger and clearer internal processes are needed to make BCC’s Open Data policy and strategy a reality. At the moment, too much relies on individual good will and departmental culture. There should be a reversal in the rules of access so that Open Data becomes the norm, unless there is a genuine public interest in non-disclosure, rather than the opposite case as currently. (It is a similar culture change as was required for Freedom of Information). It is a positive sign that the Chief Executive of BCC has advocated an approach of Open by Default. In addition, there needs to be a simpler authorisation process that involves a smaller number of people and stages, particularly for data drawn from multiple systems.

To help encourage local authorities to make more active progress in making their data openly available, there is a need for a clearer rationale of the benefits of transparency to local government itself. This rationale should consider philosophical goals (e.g. there should be a debate of equals between the council and its citizens), financial goals (e.g. Open Data should reduce the frequency and cost of FOI requests, and improved services. This rationale should be illustrated by case studies demonstrating the benefits of releasing data to the public.

Finally, it is worth noting that the project team are actively working to extend the collaborations formed in this project. Aston University has submitted a proposal (in collaboration with Warwick University, the University of Birmingham, and Birmingham City University) supported by the Chief Executive of BCC to the Leverhulme Trust for a £10M ten-year research centre: the Leverhulme Research Centre for Co-Created Cities: Citizens, Assemblages, and Diversity. The goal of this Centre is to understand better the role multi-layered social interconnections and urban diversity play in strengthening the fabric of society. The Centre, which will be led by Aston University, will co-create a longitudinal research programme in partnership with citizens to understand civic engagement, diversity, the process of becoming adult and how this changes with new technologies, investigating the ‘constructive disruption’ of city life: enabling communities to take ownership of their city. By exploring living, working and growing up in a diverse city the research findings from the Centre will be central to solving the issues and problems at the core of future decision-making in cities across the world.

8 Appendices

There are six appendices which contain the briefing document given to participants at training events, the content from one of the training sessions, activities undertaken by the project team to disseminate the results of this project, data release form on the report, the user feedback questionnaire and the report following the research undertaken by Boilerhouse on data available, visualisation options and innovations seen elsewhere.

8.1 Briefing

This briefing was provided to the participant organisations to outline the purpose of the project.

The Breakthrough funding enables Local Authorities to make certain datasets available and encourage individuals and organisations to understand the value of data and open data analysis to help address community issues that matter most to them.

The Birmingham programme will utilise data, released by BCC Housing Department into a new data hub, with the aim of encouraging the continued placement of data into this hub in the future.
The programme is a collaboration between BCC (Digital Birmingham), Regional Action West Midlands (RAWM), Podnosh and Boilerhouse.

Digital Birmingham is working with BCC Housing to identify the datasets and facilitate the extraction and placement of the data.

RAWM has identified 3-4 communities to work with, communities with active community groups linked to other national and local initiatives i.e. Our Place and Healthy Villages.

Boilerhouse will be undertaking a review of existing tools to visualise datasets, working with communities and RAWM, to enhance understanding of the use of data by communities, and identify where and how such activities can inform better decision making and community activities.

Podnosh who undertake Social Media Surgeries will develop a toolkit, training material and engagement concept for communities as well as training existing social media surgeons and facilitators in the use of data for community benefit.

The programme is aimed at supporting individuals, communities, third sector and support organisations to understand the value of data and open data.

It is envisaged that there will be a sustainable programme of community engagement, enhancing the process developed through ‘Surgeries’ to develop interest, focusing on developing skills to utilise data through visualisation tools. Activities will focus on publishing, use and maintenance of open data including crowd sourcing of data, and enhancing the data literacy of community groups.

The programme aims to act as a catalyst to encourage grass roots engagement, enhance the activity of networks, support social innovation in communities, improve transparency and civic engagement and build trust to co-create programmes.

RAWM has had one meeting with the identified community partners and is arranging a meeting with Boilerhouse and community groups in the New Year to outline available open datasets and look at some visualisation.

RAWM is also facilitating correspondence between community partners and Boilerhouse to address queries concerning available data.

RAWM is also working with BCHT concerning the release of data related to the Healthy Villages initiative.

RAWM would expect community partners to attend sessions and/or engage community activists to attend sessions.

8.2 Content of Initial Training Sessions: Community Open Source 101

Fair Data Access and Use – How to use the Freedom of Information Act and write effective requests.
Wikipedia – introduction to Wikipedia and sister projects e.g. Wikiversity, Wikimedia, Wikidata and demonstration of editing an entry to encourage community groups one way to control their digital footprint by shaping what was written about their area.

Open Street Map – use of Open Street Map as a tool to enhance community activity by finding and presenting local information.

Introduction to research – An introduction to the use of search engines and advanced search tools, illustrating how simple searches can be enhanced using filters to find specific information.

Turning Data into Pictures: Google Fusion – use of visualisation to enhance the narrative connected with data.

Turning Data into Pictures: Tableau Public – introduction to several free data visualisation tools, using a dataset from the Birmingham Data Factory, to show community groups how visualisation can enhance presentations of information.

Excel Spreadsheets – collection and collation of data; basic sorting and analysis skills.

Google Analytics – developing organisations’ appreciation of their digital footprint: how it is measured and how it can be used.

Where is the data? Introduction to data at various levels, including output areas, and identification of freely available datasets under various themes including housing, transport and health which community groups can use to add weight and significance to reports, funding applications etc.; included an opportunity to see the Boilerhouse data visualisation tool.

8.3 Communication Activities

Partners used the following blog posts, articles and Twitter to reach VCS groups.

- [http://www.rnrorganisation.co.uk/strategic-management/smart-cities](http://www.rnrorganisation.co.uk/strategic-management/smart-cities)
- [http://www.rnrorganisation.co.uk/blogs/opendataandsocialenterprise-worldscollide](http://www.rnrorganisation.co.uk/blogs/opendataandsocialenterprise-worldscollide)

RAWM contributed to additional external events to broaden awareness of the work on this project.

- **VCSSCamp Barnsley, Nov. 2014 – 3rd RnR Organisation co-organised national unconference for VCS infrastructure organisations sponsored by NCVO Big Assist – session for 15 groups about VCS organisations using data and open data, including some open source tools, and references to Birmingham Skills and Data Hub project including development of an ethical framework to encourage health bodies to consider release of data**

- **New ideas on social social value, Dec. 2014 – regional conference on Social Value Act with 50+ voluntary sector, social enterprise and public sector delegates from West Midlands - referred to community groups involvement in Birmingham Skills and Data Hub project with specific reference to use of data and good commissioning practice**

- **Big and open data for social good, Feb. 2015 – Nesta national conference – featuring panel discussions on use of big and open data for good by Citizens Advice Bureau, DatakindUK, NCVO, Nesta, Indigo Trust, Third Sector Research Centre – referred to community groups involvement in Birmingham Skills and Data Hub project and possibility of larger VCS organisations supporting/mentoring smaller VCS organisations around data skills and knowledge in Q & A sessions**

- **Open Data Camp, Feb. 2015 – 1st national 2-day unconference about open data – ran session for 15 delegates on community groups use of open data with specific reference to Birmingham Skills and Data Hub project; co-ran session on Data and Charities with colleague from New Philanthropy Capital for c.30 delegates on barriers to VCS organisations using data including lack of digital literacy, making specific reference to Birmingham Skills and Data Hub project and use of open source tools; participated in**
sessions on digital literacy and writing better request for information using the Freedom of Information Act

- **Better Data, Mar. 2015** – One East Midlands/Regional Voices regional event – took part in several discussions about data with Department of Health & NICE representatives and informed them about our work on the Birmingham Skills and Data Hub project and use of open source tools
### 8.4 Data Release Form (completed for release of BCC housing stock data)

<table>
<thead>
<tr>
<th>Open Data Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Sets for Publication on Birmingham Data Factory</td>
</tr>
<tr>
<td><a href="https://data.birmingham.gov.uk/">https://data.birmingham.gov.uk/</a></td>
</tr>
</tbody>
</table>

**Directorate:** Place  
**Section:** Housing  
**Data set name:** BCC Housing Stock  
**Business application/Manual application (e.g. Excel/Access):** Northgate/(SX3)  
**Description of data and metadata:**  
The data describes the locations in an aggregated area (Lower Upper 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

<table>
<thead>
<tr>
<th>Data Format</th>
<th>Size (in MB)</th>
</tr>
</thead>
</table>

**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:
<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is all or some of the data owned by someone other than the Council?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, have we got permission in writing to release the data?</td>
<td></td>
</tr>
<tr>
<td>Give details:</td>
<td></td>
</tr>
<tr>
<td>Have any FOI requests for the data been refused, or would we reasonably</td>
<td></td>
</tr>
<tr>
<td>refuse to release the data if it were requested by via FOI?</td>
<td></td>
</tr>
<tr>
<td>Does the data fall under any exclusion (not necessarily cost)</td>
<td>No - Property addresses have been released previously through FOI4275 and FOI10211</td>
</tr>
<tr>
<td>EINA / Ethical Considerations and any decisions made to adjust the data</td>
<td></td>
</tr>
<tr>
<td>before release</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>The housing stock data that describes over 64,000 social housing</td>
<td></td>
</tr>
<tr>
<td>properties owned by BCC was extracted from Northgate. Matthew Cloke,</td>
<td></td>
</tr>
<tr>
<td>data expert in BCC advised on the suitability of the data fields/values</td>
<td></td>
</tr>
<tr>
<td>extracted. As a result: Obsolete fields have been deleted. Information</td>
<td></td>
</tr>
<tr>
<td>has been totalled up in an area to prevent any individual property being</td>
<td></td>
</tr>
<tr>
<td>identifiable for more sensitive information e.g. which house has</td>
<td></td>
</tr>
<tr>
<td>insulation To avoid the risk of identifying individuals the data set is</td>
<td></td>
</tr>
<tr>
<td>being aggregated in the GIS system to Lower Super Output Area leave as</td>
<td></td>
</tr>
<tr>
<td>this has 86% of the areas with more than 5 properties in and will be</td>
<td></td>
</tr>
<tr>
<td>published as described in the 'Data Example' tab.</td>
<td></td>
</tr>
<tr>
<td>If all questions are answered ‘No’, consideration should be given to</td>
<td></td>
</tr>
<tr>
<td>placing this data on the Open Data platform.</td>
<td></td>
</tr>
<tr>
<td>Is the data to be placed on Open Data platform?</td>
<td>YES</td>
</tr>
<tr>
<td>Form completed by:</td>
<td>Stuart Lester / Heike Schuster-James</td>
</tr>
<tr>
<td>Date:</td>
<td>26/02/2014</td>
</tr>
<tr>
<td>Signed off by:</td>
<td></td>
</tr>
<tr>
<td>Date:</td>
<td></td>
</tr>
</tbody>
</table>
8.5 Feedback Questionnaire

This is the online questionnaire used to get feedback from the users at Open Data Surgeries.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>How aware are you of what open data is?</td>
</tr>
<tr>
<td>2</td>
<td>What are the objectives your group aims to achieve by using open data?</td>
</tr>
<tr>
<td>3</td>
<td>Which categories of data are you most interested in having access to? (Choice of 9 categories).</td>
</tr>
<tr>
<td>4</td>
<td>Which tasks do you think could be better facilitated with the availability of open data for Birmingham?</td>
</tr>
<tr>
<td>5</td>
<td>What is your general view on open data in promoting the social, political and economic health of Birmingham?</td>
</tr>
<tr>
<td>6</td>
<td>If you have used open datasets, were they made easily available by the Government/Council?</td>
</tr>
<tr>
<td>7</td>
<td>If you have used open datasets, please give a short description of the scenario where you have used them, describing your experiences and highlighting new things that you may have learned.</td>
</tr>
<tr>
<td>8</td>
<td>Were the openly available datasets complete for their purpose?</td>
</tr>
<tr>
<td>9</td>
<td>What were the benefits of using open data?</td>
</tr>
<tr>
<td>10</td>
<td>If you needed access to a specific open dataset, are you aware of whom to contact within the Government/Council?</td>
</tr>
<tr>
<td>11</td>
<td>What kind of support do you think should be provided to stakeholders interested in using open data?</td>
</tr>
<tr>
<td>12</td>
<td>How did working with the project team shape your understanding about the usage and benefits of open data?</td>
</tr>
<tr>
<td>13</td>
<td>What are your suggestions for improving the availability of open datasets for Birmingham?</td>
</tr>
</tbody>
</table>

Table 1: User feedback questionnaire.

Report on data available, visualisation options and innovations seen elsewhere

8.6 Report on data available, visualisation options and innovations seen elsewhere

As part of the deliverables under this programme, Boilerhouse undertook to research innovations in visualisation of local open data with particular reference to housing data.

What data is available?

There is a huge amount of data on housing potentially available to this project

- open data from the census, indices of deprivation, land registry and other national sources:
- data from council housing systems: properties and their attributes; rent levels; rent collection data; repairs and improvements etc; houses in multiple occupation
- data from council reports and surveys; state of housing stock
- data from private sector sources, eg online residential property search companies like Rightmove.co.uk, Zoopla.co.uk, home.co.uk

We cannot show it all, we need to decide what is the most useful data to show through this project.

The **geographies** data is available for varies enormously, some data is available at LSOA level (the smallest geography used for collecting census data, representing around 100 households). LSOAs fit into local authority boundaries, but not political ward boundaries (ie LSOAs may span two or more wards). Some data is available only at ward or even local authority level.

Data for the geographies themselves (local authorities, wards, LSOAs etc) is available from Ordnance Survey. Tools for things like mapping postcodes to LSOA also exist.

The **dates** data is available for also varies. The latest census data is from the 2011 collection. Data from local authority, government department or other public authority data collections or line of business systems may be more current (eg figures of JSA claimants from DWP or house prices from the Land Registry). However, availability depends on publishing practice by the authorities concerned.

Data from the **private sector** may be available where there is some benefit to them in publishing it., eg many property businesses publish house price or rental indices.

**What data might we want to show and why?**

The project is being driven by the community. We are consulting with community groups through meetings organised by RAWM and social media surgeries organised by Podnosh, to find out what data they want to get hold of and how they might like to see it presented so that it supports/informs the work they do.

The meeting held on 6 January at the Boilerhouse offices included representatives from:

- Abdullah Rehman, Basall Heath Forum
- Ruth Miller, Castle Vale Community
- Pauline Roche, rawm.co.uk
- Ted Ryan, rawm.co.uk
- Mohammed Ashfaq, KIKIT, Sparkbrook
- Mohammed Shafique, Chair, Ashiana Community Project, Sparkbrook Neighbourhood Forum
- Naeem Qureshi, Manager, Ashiana Community Project, Sparkbrook Neighbourhood Forum
- Jeremy Monk-Hawksworth, Black Country Make It

The discussion revealed that:

- Data collected by public authorities is often at too large a geographical area to be of use to them. eg Castle Vale neighbourhood is only part of Typurn Ward, with wards being the 'lowest' level at which official data is published
- Data that is published is not easily accessible. For example, no one at the meeting has seen the pdf ward profiles that show census and other data at ward level, and are freely
available from LG Inform (as pdf documents only)

- There was agreement that simple data was needed to engage people – the majority tenure map was a case in point

- Community groups collect much data themselves, but struggle to get this into the public domain because the data is not ‘official’ and they do not have easy means to publish and promote it.

- The ability to publish by adding data to ‘crowd-sourcing’ tools collecting data defined by them would be welcomed

- Where they are involved in working with ‘the authorities’ in projects where they are a conduit to people in the local community, they often get no feedback on outcomes of projects. Eg they may have collaborated in initiatives to reduce GP or hospital visits, but they do not get data about whether the projects were successful or whether their contribution was useful. This means they have no evidence of effectiveness to refer to when bidding for future projects.

- In terms of data they would like to see, there was interest in ‘consumer’ data like rents and how council and national policies affect these. For example, one representative said that it would be interesting to be able to get data showing a link between the bedroom tax and prices of one bedroom flat rentals. The suggestion was that housing benefit recipients were being moved from council properties (where there is a shortage of one bed flats) into one bed flats in the private sector, with a resulting impact on rentals for those renting privately.

- It would also be useful to be able to see the impact of council polices – eg not putting families in flats above the fourth floor – on prices.

The housing ‘wish list’ from this discussion comprised:

- No of houses with more than 5 occupants
- Concentration of applicants on housing list for 7 years
- Void Properties
- Areas of overcrowding for more than five years
- No of people homeless
- Housing migration - living less than one year
- Development Housing programmes by investment
- Fuel poverty more than 20%
- Leaving Care - on location and migration
- Parents kicking out their children over 16

Not all of this data is ‘housing’ data – eg children leaving care – although it may have a housing implication. Not all of the data will be open data or released by BCC within this project timescale, but the list has been forwarded to the project team to investigate.
Commentary of data tools and visualisations researched for this project

Our research started by looking at some of the visualisations created through the Release of Data and Breakthrough Funds and other publicly-funded data schemes. We then looked at visualisations created by data journalists, data consultancies and experts.

The purpose of this activity was to find good practice to bring into this project and to avoid re-inventing the wheel.

Release of Data and Breakthrough Fund projects 2014/15 are listed at http://www.local.gov.uk/local-transparency/-/journal_content/56/10180/3926733/ARTICLE. Many of these are work in progress and little has been published yet.

Local Open Breakthrough Programme 2013/14 are listed at http://www.local.gov.uk/web/guest/local-transparency/-/journal_content/56/10180/4049888/ARTICLE

The following projects from these lists have elements of visualisation or other aspects relevant to our project:

**Trafford Council** has created InfoTrafford as the council’s ‘data observatory’ which contains open data about Trafford and its communities in one datastore (www.infotrafford.org.uk ). There is a small amount of data on housing, some of it only at borough level, enabling comparison of Trafford with the North West and England, but not Trafford wards or LSOAs with each other.

The site has interactive features, and includes for example a map of council tax bands. However this is presented in GeoWise InstantAtlas software and is not easy for the novice to understand or navigate:

http://www.infotrafford.org.uk/dataviews/report/multiple/fullpage?viewId=240&reportId=44&geoTypeId=48,15,33,17

**Cambridgeshire County Council: Cambridgeshire Insight Open Data**: Stimulating Economic Growth and Encouraging Innovation http://opendata.cambridgeshireinsight.org.uk

There is a housing data visualisation here - http://www.cambridgeshireinsight.org.uk/interactive-mapping/housing/cambridgeshire-atlas-sub-region-housing-profiles

This is done in Instant Atlas and not sufficiently intuitive, in our view, for the citizen/community (or senior decision-maker) user

**Cheshire East Council: URIs for natural neighbourhoods**: A public interface for defining geographical areas at neighbourhood level, other than formally supported administrative or census output areas. http://neighbourhoods.esd.org.uk/

http://www.cheshireeast.gov.uk/council_and_democracy/council_information/research_and_consultation/open_data.aspx

The concept of defining neighbourhoods in this way is relevant to community driven projects like ours.

**Devon County Council**: Local Government Community Data – working with communities, businesses and IT experts find what data are most required and how best to make them useful to communities and business.

https://new.devon.gov.uk/factsandfigures/open-data/opening-up-our-data/
We found data that we would want to use in our project (e.g. housing tenure data) and saw good use of data directly from DCLG’s Open Data Communities resource.

**Hampshire County Council:** Hampshire Hub [http://www.hampshirehub.net/](http://www.hampshirehub.net/)

One of the site features is Area Profiles - collections of charts and key statistics about different places. They are created automatically from the data, and are currently available at County, Unitary, District, Ward and Parish level within Hampshire and the Isle of Wight. Profile data is laid out nicely, e.g. as tables with bar chart. Housing data shown (search on housing):

**Housing data shown (search on housing)**

Portsmouth: Homelessness and Housing

Households on the housing register (excluding households waiting for transfers)

House building

Indices of Deprivation 2010: Indoors Living Environment subdomain

Housing Benefit recipients

Housing Benefit and Council Tax Benefit claimants

Households (by tenure)

Homelessness (family rate)

Homelessness (annual)

Households (by deprivation dimensions)

Housing (sharing / central heating / occupancy rating by age)

CO2 emissions reduction (per capita)

Indices of Deprivation 2007

Fuel poverty

Fuel poverty (people receiving income-based benefits in homes with a low energy efficiency rating)

Some data is displayed in a ‘spreadsheet view’ – this is very hard to use for the layperson and it is hard to see the benefits of displaying like this? Is it to show lots of data in small space?

Data sources are not always obvious. However this is a beta project and presented as work in progress.

**Leeds City Council:** Leeds Data Mill - [http://www.leedsdatamill.org](http://www.leedsdatamill.org)
Leeds Data Mill holds 19 housing-related data sets (although some on inspection turn out not to be housing data). Data is held as CSVs mainly. It appears that the website can visualize the data as graphs and maps. However, these tools are not intuitive – and it may be that they are not working yet. (NB Barnet uses the same tools, and they do not seem an easier to use on http://open.barnet.gov.uk/)

Leeds housing data available is as follows:

- Permanent Dwellings by Tenure Type in Leeds: the number of new permanent dwellings that has been built down by the tenure type.
- Council Housing Stock, including sold and demolished dwellings.
- Tenanted housing stock: all council housing which are currently tenanted, i.e. not empty properties. It provides details on property type, area, ward, and by post code sector
- Council house bids: details on how many bids are being made on specific vacant council properties. Information includes full postal address, type of property etc
- Council properties under-occupation: these have all had a reduction in their housing benefit.
- Council tax charges: by property band
- Private Sector Long Term Empty Properties by Ward: empty properties (6 months+) in the private sector.
- Households in Temporary Accommodation: broken down by local authority and temporary accommodation type from 2011 Q2 to 2014 Q2
- City wide empty properties trends (regardless of tenure or how long empty)
- Long term empty properties owned by Ltd companies
- Domestic Energy Performance Certificates By Energy Efficiency Rating

Peterborough and Bath and North Somerset are using Redbridge’s DataShare platform. North Somerset’s has no housing data on it. Peterborough has discretionary housing payments in a table but the visualisation tools do not appear to work.

Other visualisations of housing/property or related data

Lumino City 3D at Centre for Advanced Spatial Analysis UCL: http://www.bartlett.ucl.ac.uk/casa


Peterborough DNA: http://www.peterboroughdna.com/living-data Peterborough received £3m from TSB through Smart City programme. Site is very hard to use and very slow. Contact: helena@peterboroughdna.com

The BBC is one of a number of publishers who have created visualizations showing where people on different incomes can afford to live. A calculator published in December 2014 shows users where in the country they can afford to live - and see whether it would it be cheaper to rent or buy. All they need to do is enter their requirements and how much they would like to pay on rent or mortgage repayments each month. Affordable areas can then be seen on a map. http://www.bbc.co.uk/news/business-23234033
www.thisismoney.co.uk published in March 2013 a static heat map of England Map showing the level of affordability of property in England, calculated by measuring average prices against average earnings of residents. Other maps showed levels of changes in affordability since 2007, and the most affordable and unaffordable areas outside London. ‘Least affordable’ properties are those with the highest ratio of house prices to earnings. http://www.thisismoney.co.uk/money/mortgageshome/article-2295295/Property-affordability-map-Most-reasonable-house-prices-England.html

The Daily Telegraph has created static maps combining house price data with commuter fares to come up advice on ‘where cheap property prices meet cheap train lines’. http://www.telegraph.co.uk/finance/personalfinance/10958987/Map-where-cheap-property-prices-meet-cheap-train-lines.html

The Greater London Authority website publishes The London Rents Map showing average private sector rents for different types of home across London. Users enter a location and a property type and the map shows data for that area with the area location marked on the map. The data on average rents is given at postcode district level (SW19 or E7, for example), and is based on a sample covering the last 12 months (up to the date shown in the last update box) so does not fully reflect the most recent short-term trends in the market. Rent data has come from the Valuation Office Agency (VOA). https://www.london.gov.uk/priorities/housing-land/renting-home/rents-map

Rightmove publishes house price data and shows example sales on a local map. For an area (eg Birmingham) it provides commentary of types of property that make up majority of sales and their average sale price as well as average prices for other popular property types (eg semi-detached houses and flats). It provides comparison of prices with nearby areas (in this case, Birmingham City Centre, Sparkhill and Small Heath). There is also information on rises in ‘overall sold prices’ in the area in the last year and in comparison to previous years. http://www.rightmove.co.uk/house-prices/detail.html?originalIncode=&outcode=&rcode=&country=england&locality=&buildType=1&propertyType=0&landRegTenureType=ANY&year=30&radius=0.0&sortOrder=descending&locationIdentifier=REGION^162&referrer=detailPage&columnToSort=PRICE_ASC

The Guardian published an item in its Datablog in March 2013 with the headline Housing on the minimum wage: how many hours would you have to work? The item provides and interactive heatmap of showing how many hours of minimum wage work would be required to pay the median rent on a one bedroom flat. It doesn’t take account of housing benefit or other allowances. Data is from the Valuations Office Agency. Users click on the map to see how the data varies - or can search for their address and the map will zoom into the area. http://www.theguardian.com/news/datablog/interactive/2012/mar/19/housing-minimum-wage-map

My Society has built a service called Mapumental Property, which uses open public transport data from across the country to show users areas where they can live that are within an acceptable commuting time of their office, school or other destination. Mapumental was built with funding from Channel 4’s 4iP project. The site helps people to work out where to live and by providing a map of everywhere within a certain number of hours/minutes’ commute from a particular point. https://mapumental.com/services Mapumental Property includes property data
from Zoopla to show people what properties are available that involve a commute within a time specified by the service user. The user enters their destination postcode and the time they need to arrive there. Then they specify the length of commute (in time) that they are prepared to tolerate, then they click on ‘rent’ or ‘buy’. Properties that meet the criteria appear on the map. http://property.mapumental.com/map/B55NR-a-0900/4a6baed5#t45

In the course of our research we identified a wide range of housing-related data that might be visualized through this or similar projects

Some of the data sources are not known to us.

We have added initial ideas about how data might be visualized and the value of doing so.

<table>
<thead>
<tr>
<th>Data subject</th>
<th>Details</th>
<th>Source</th>
<th>Visualized as</th>
<th>Value of showing this data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tenure of all households</td>
<td>Owned outright; rented private; rented council; rented from private registered providers* etc</td>
<td>Census 2011 - ONS</td>
<td>Heat map with majority per ward breakdown in pop up Could be done per LSOA?</td>
<td>Tells you something about residents’ wealth, mobility etc</td>
</tr>
<tr>
<td>Number of households by ward</td>
<td></td>
<td>Census 2011 - ONS</td>
<td>Heat map showing wards/LSOAs with most households</td>
<td>Shows residential areas – but maybe only useful at lower level</td>
</tr>
<tr>
<td>People per household by ward</td>
<td></td>
<td>Census 2011 - ONS</td>
<td>Heat map showing wards with different sizes of households predominating</td>
<td>Shows where families live (tho how deal with HMOs?)</td>
</tr>
<tr>
<td>Overcrowding by ward **</td>
<td></td>
<td>Council?</td>
<td>Heat map</td>
<td>Linked to deprivation</td>
</tr>
<tr>
<td>Property types by ward</td>
<td>Number of residential properties of different types eg Detached Semi-detached Terraced Flats 1/2/3/4/more beds</td>
<td>Census ? Land registry?</td>
<td>Graph/chart Rank volumes per ward or LSOA?</td>
<td>Useful to citizens seeking accommodation in a particular area</td>
</tr>
<tr>
<td>Location of houses of multiple occupation</td>
<td>Council data from registration</td>
<td>Could show individual properties – is there a privacy</td>
<td>Shows where HMOs concentrated – useful for community activists?</td>
<td></td>
</tr>
<tr>
<td>(HMOs)</td>
<td>Rent values</td>
<td>Rent by property type by ward - private - social (council and registered private)</td>
<td>Zoopla API Council Statistical data returns*</td>
<td>Heat map Graph/chart</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Affordability</td>
<td>Affordable Like BBC tool</td>
<td>Based on average prices per ward?</td>
<td>Interactive tool</td>
<td>Would show which parts of Birmingham you could afford to live in</td>
</tr>
<tr>
<td>House prices</td>
<td>Prices by property type by ward</td>
<td>Land registry</td>
<td>Heat map Graph over time</td>
<td>Indicator of affordability of areas Indicator of economic growth/decline</td>
</tr>
<tr>
<td>Council tax bands</td>
<td>Predominant council tax band by ward Average council tax paid by ward</td>
<td>Council data</td>
<td>Heat map</td>
<td>Shows value of residential property in the area based on 1991 valuations</td>
</tr>
<tr>
<td>Housing deprivation</td>
<td>Use indices of deprivation housing domain</td>
<td>Take from IMD</td>
<td>Heat map Aggregate into wards?</td>
<td>Shows for activists where housing is an issue</td>
</tr>
<tr>
<td>Condition of residential property by ward</td>
<td>Show by Private Council Registered private</td>
<td>Not sure – council research data?</td>
<td>Heat map of LSOAs/wards by condition</td>
<td>Makes case for regeneration funding</td>
</tr>
<tr>
<td>ASB</td>
<td>ASB reports ASBOs by ward or LSOA breakdown by tenure type?</td>
<td>Crime stats? Council data</td>
<td>Heat map with ASBs aggregated to lowest area that does not compromise confidentiality</td>
<td></td>
</tr>
<tr>
<td>Housing benefit recipients by ward</td>
<td>Households with one or more housing benefit recipient</td>
<td>Council data</td>
<td>Heat map Graph/chart</td>
<td>Shows areas where numbers of people on low incomes are living</td>
</tr>
</tbody>
</table>
**Fuel poverty**

- fuel costs above the national median level which if spent leaves people with a residual income below the official poverty line
- English housing survey
- Heat map of LSOAs with significant numbers of households in fuel poverty
- Indicator of deprivation

**Households on the housing register**

- Not sure what this is but saw it on another council data hub
- Council data
- Most use in comparing areas including with rest of W Midlands and England

**House building starts by ward**

- Show by Private Council Registered private
- Planning dept?
- Heat map Graph over time
- Additions to housing stock that might impact on homelessness or affect house prices or rentals

**Housing benefit**

- Show private sector market rents against housing benefit
- See ***

---


** overcrowding definitions are if a boy and a girl aged 10 years or older need to sleep in the same room. 1 sleeping room has more than 2 people, 2 more than 3, 3 more than 5 etc, NB sleeping room sizes are defined


On 29 January we proposed to BCC that we focus on the following data for our project (datasets in red subject to BCC being able to provide them):
<table>
<thead>
<tr>
<th>Item</th>
<th>Level</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly income (takehome)</td>
<td>Constituency</td>
<td>Nomis</td>
</tr>
<tr>
<td>Private rents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>By property size (bedrooms)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>By type (flat, maisonette etc)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcodes captured at LSOA via Mapit</td>
<td></td>
<td>Zoopla</td>
</tr>
<tr>
<td>Social housing rents (HAs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>By property size (bedrooms)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>By type (flat, maisonette etc)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcode? To be aggregated to LSOA</td>
<td></td>
<td>Housing associations via BBC WMBUS</td>
</tr>
<tr>
<td>Social housing rents (BCC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>By property size (bedrooms)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>By type (flat, maisonette etc)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregated to LSOA level</td>
<td></td>
<td>BCC housing</td>
</tr>
<tr>
<td>All housing stock in the area</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private / council/ other social</td>
<td></td>
<td>2011 census</td>
</tr>
<tr>
<td>By property size (bedrooms)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>By type (flat, maisonette etc)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSOA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenure (owned outright, social rented etc)</td>
<td>LSOA</td>
<td>2011 census</td>
</tr>
<tr>
<td>Households</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSOA</td>
<td></td>
<td>2011 census</td>
</tr>
<tr>
<td>Occupancy (bedrooms)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSOA</td>
<td></td>
<td>2011 census</td>
</tr>
<tr>
<td>Properties in multiple occupation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcode aggregated to LSOA</td>
<td></td>
<td>Council</td>
</tr>
<tr>
<td>Availability of housing for rent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcode aggregated to LSOA</td>
<td></td>
<td>Zoopla api</td>
</tr>
<tr>
<td>JSA claimants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postcode aggregated to LSOA</td>
<td></td>
<td>Nomis</td>
</tr>
<tr>
<td>Housing benefit claimants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- private tenants</td>
<td>LSOA level</td>
<td>BCC</td>
</tr>
<tr>
<td>- bcc tenants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- other social housing provider tenants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing deprivation rank in Bham</td>
<td>LSOA</td>
<td>ONS</td>
</tr>
</tbody>
</table>

We also asked BCC to investigate the possibility of finding additional datasets from this wish list put forward by Naeem Qureshi of the Ashiana Community Project:

- No of houses with more than 5 occupants
- Concentration of applicants on housing list for 7 years
- Void Properties
- Areas of overcrowding for more than five years
- No of people homeless
- Housing migration - living less than one year
- Development Housing programmes by investment
- Fuel poverty more than 20%
- Leaving Care - on location and migration
- Parents kicking out their children over 16