Empowering citizens to use their data: a public library perspective.

Luke Burton

Digital Development Manager, Community Hubs, Libraries and Parks
Newcastle City Council

luke.burton@newcastle.gov.uk

Open data is a term which has been receiving a lot of attention in many fields recently, and that includes the library and information world. The term is not always well understood, so the first step is to define what we mean by open data. Then open data can be placed in the context of libraries by giving examples of how it is being used and discussed in public libraries in particular.

Open data has been defined in a number of ways but the simplest and most succinct is that given by the Open Data Institute (ODI): 'Open data is data that anyone can access, use or share. Simple as that.' (Open Data Institute, 2017).

To give a little more detail the Open Definition describes open data as:

'Data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike'.
(Open Knowledge Institute, 2016).

Open data fits perfectly with the ethos and skills base of library and information professionals who are, by definition: 'those professions which are engaged in the creation, organization, diffusion, and preservation of information and knowledge.' (Greer, R. C, Grover, R. J. and Fowler, S. G. 2007). What could be more natural then for information professionals, and librarians in particular, to play a key role in the creation, storage, organisation and dissemination of open data? This premise was taken as the basis for the work that Newcastle Libraries took as the starting point for their work on open data.
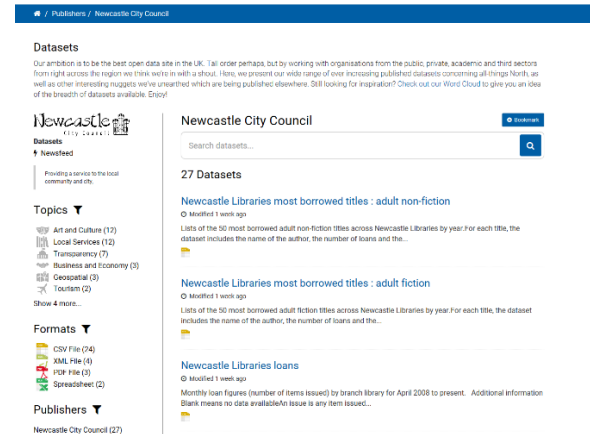
Newcastle Libraries Service has a large amount of data relating to a range of topics, as many library services will do, including, visitor figures, issues, computer use, opening hours, buildings, energy use and catalogue data. This data could be used in a number of ways both internally and externally and by releasing it as open data it allows others to use it in any way they see fit, some of which will hopefully be of benefit to the library service. Internally the data is intended to be used to drive service delivery and development as the library service faces significant budgetary challenges over the next three years.

The first step in the process of releasing the data was to speak to colleagues within Newcastle City Council who were working on open data allowing the library services to fit its work into the Council's policy which states that: 'as a council, we're interested in open government, and how this can help us be a cooperative council' (Newcastle City Council, 2017). The idea of transparency and accountability is an integral to the decision to release this data; as a library service Newcastle Libraries are

only the custodians of the data, it belongs to the citizens of the city. By releasing the data, one version of the truth is provided which can be used for internal reports, by elected members, for Freedom of Information (FOI) requests and to allow anyone to create services and products that may be of interest or use to the library service.
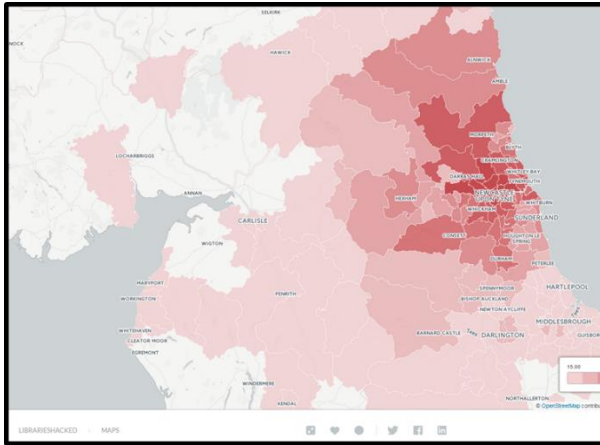
The next step was to ensure that the data was released in a format that would be of use to as many people as possible and without them having to do too much, if any, work to make it useable. Local developers were very helpful in providing information on formats and how to described the data, the metadata. The files were released as Comma Separated Values (CSV) files making them easy to use programmatically. The files were originally released under the Open Government Licence v3.0 (OGL v.3) (The National Archives, 2017), this was later changed to Creative Commons 0 (CC0) licence which releases the data into the public domain (Creative Commons, 2017). It was important to describe the data effectively so that it could be used without having to contact the Council for further information.

The data is now published in two repositories the Data Mill North (Data Mill North, 2017) and on GitHub (GitHub, 2017). On the Data Mill you will find datasets that have been tidied up and for which notes have been created explaining what the data refers to. The data on GitHub still requires work to clean it up and make it useable but it was felt it was more important to make the data available and clean it up dependent on demand and allow others to do some of the cleansing of data.

Newcastle City Council data sets on Data Mill North (https://datamillnorth.org)

In order to empower citizens to use their data it was felt necessary to run some events that demonstrated what could be done with the data as well as giving people the skills to use the data that belongs to them. Two data days were held, one to see what ideas developers might come up with and one a workshop to teach data visualisation skills to those that don't code or programmme in order to lower the barriers to entry in this area. These events highlighted the data, and exposed the library service, to a new audience as well as highlighting what could be done with the data that might benefit the library service and residents of the city. Some examples of the way the data was used include a text to speech service that scanned a barcode on an item and using library catalogue data read the titles description back; something that could be used on audiobooks for visually impaired people. The day for those learning new skills had worksheets available that allowed those new to data visualisation to create something that not only looked attractive but was of practical use, such as plotting library membership over time or the location of certain buildings. These worksheets were, and still are, available on GitHub (GitHub, 2017).

Visualisation of Newcastle Libraries' membership using Carto (Carto, 2017) produced by Libraries Hacked (Libraries Hacked, 2017)

The next steps in this work will be to run other events that will use and highlight the data such as data as art. A key future role for the data is also that of data driven decision-making as new and different services are developed and in the hope that others may develop services that benefit the library service and by extension the city or that allow the Council and library service to gain better insights into what they are doing.

Open data can open up a range of possibilities for different types of organisations, including library and information services. They can act as a single version of the truth accessed by all, transparent and accessible. It can allow new insights to be gained and new products to be developed by those who have the skills that are missing from an organisation.

References

Carto (2017) *Carto*. [online] Available at: https://carto.com/ [Accessed 15/05/17].

Creative Commons (2017) *CC0*. [online] Available at: https://wiki.creativecommons.org/wiki/CC0 [Accessed 3/5/17].

Data Mill North (2017), *Newcastle city council*. [online] Available at: https://datamillnorth.org/publisher/newcastle-city-council [Accessed 3/5/17]

GitHub (2017) *ToonLibraries*. [online] Available at: https://github.com/ToonLibraries [Accessed 3/5/17].

Greer, R. C, Grover, R. J. and Fowler, S. G. (2007), *Introduction to the Library and Information Professions*. Westport Connecticut and London. Libraries Unlimited.

Libraries Hacked (2017) *Libraries hacked*. [online] Available at: https://www.librarieshacked.org/ [Accessed 15/07/17]

The National Archives (2017) *Open government licence* [online] Available at: http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/ [Accessed 4/5/17]

Newcastle City Council (2017), *Open data policy* [online] Available at: https://www.newcastle.gov.uk/your-council-and-democracy/open-data-and-access-information/open-data/open-data-policy [Accessed 4//5/17]

Open Data Institute (2017), *What is open data?* [online] Available at: https://theodi.org/what-is-open-data [Accessed 7/5/17]

Open Knowledge Institute (2016) *What is open data?* [online] Available at: (http://opendatahandbook.org/guide/en/what-is-open-data/ [Accessed 7/5/17]