

Voyage of the Data Treader

Put your library on the map

Library Data Camp

Monday 11th March 2019

Voyage of the Data Treader is for those interested in library data, in using data to improve library services, and people who just like playing with data!

Manchester Central Library is the setting for our second library data camp, a free event for all.

Library Wikidata

Have you heard about the Libraries in Wales data project? For CILIP Cymru Wales, Kathryn Parry worked with the National Library of Wales Wikimedian Jason Evans, and Libraries Hacked, on the creation of visual datasets of libraries in Wales. All of the datasets are openly available and editable via Wikidata.

Inspired by this we invited Jason to show you (and us) how to add and edit your libraries' details on Wikidata as open data, and how to use the Wikidata

visualisation tools to "put your library on the map"!

A library data unconference

We are enthusiasts who love data, and most importantly, seeing how we can use library data to improve library services.

We'd love to see more library workers feel empowered and confident in using data within their service, learning from what others are doing (or not doing), and sharing their own data practices.

It would then be great to see more library services involved in **#opendata**, releasing public data about their library services, and allowing others to use this in various ways. How people will use open data will vary too much to try to second-guess it, but that's the exciting part!

It's important that everyone feels able to get involved and take part, and that this is an opportunity to learn. We're not looking for a particular level of data knowledge. Before our previous

event in 2017 we ran a survey to ask people how they defined their level of expertise in using data. Half of the respondents chose 'beginner', and so we need to make sure that everyone can learn from this experience.

We'd also like to hear from everyone about their experiences, and what they would like to discuss.

- What are your frustrations in using data?
- Do you find too much emphasis placed on quantitative data, rather than qualitative?
- Do you get enough time to look at data?
- Do you feel like you make data-informed decisions?
- What do you love most about data?
- Would you like to publish open data, but not have enough support to do it?

Unconferences are designed to allow people in similar situations to talk about the things they most want to, and to hear what others are interested in.

#DataTreaders

We're all Data Treaders; the ones organising this event are Claire Back (Plymouth Libraries), Aude Charillon (Newcastle Libraries), and Dave Rowe.

Sponsors

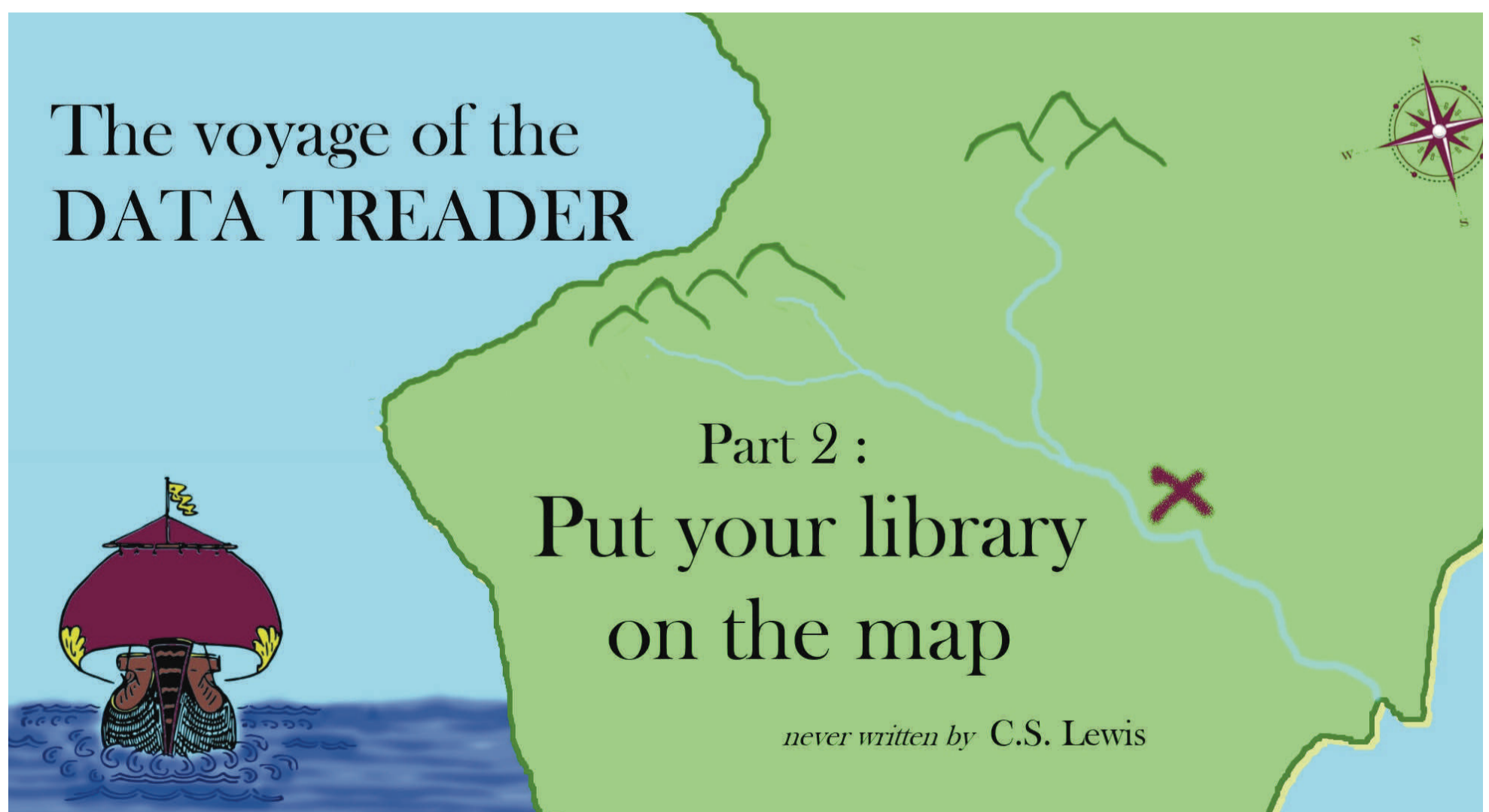


UKGovCamp sponsors

We've been kindly sponsored by UKGovCamp, and the Libraries Taskforce. We're also delighted to have been hosted by the beautiful Manchester Central Library. Many thanks!

**LIBRARIES
TASKFORCE**

The Libraries Taskforce sponsors



Libraries and Wikidata

Background

In our first Voyage of the Data Treader, Library Data Camp in November 2017, we had discussions on a wide range of topics, such as:

- Open Data 101;
- How to use an API;
- Publish your first dataset;
- Extracting data and automation;
- Social media surgery
- Open Data and social value.

(For more evidence of this do visit <https://voyage.datatreaders.co.uk>).

We also ran an introduction to data exercise where we pretended we were pieces of data, and indulged in sorting and filtering ourselves. It was a lot of fun!

Providing library staff with learning opportunities has always been a key goal, so we invited Owen Stephens to provide training on using the Open Refine software tool for cleaning and manipulating messy data (and tidy data for that matter).

(Note: Open Refine also works very well with Wikidata so check it out!).

For 2019, we decided to focus on doing something practical, as well as ensuring we learnt lots and put our open data goals into practice. So we're learning about Wikidata as a tool for library staff to publish, query, and browse open data.



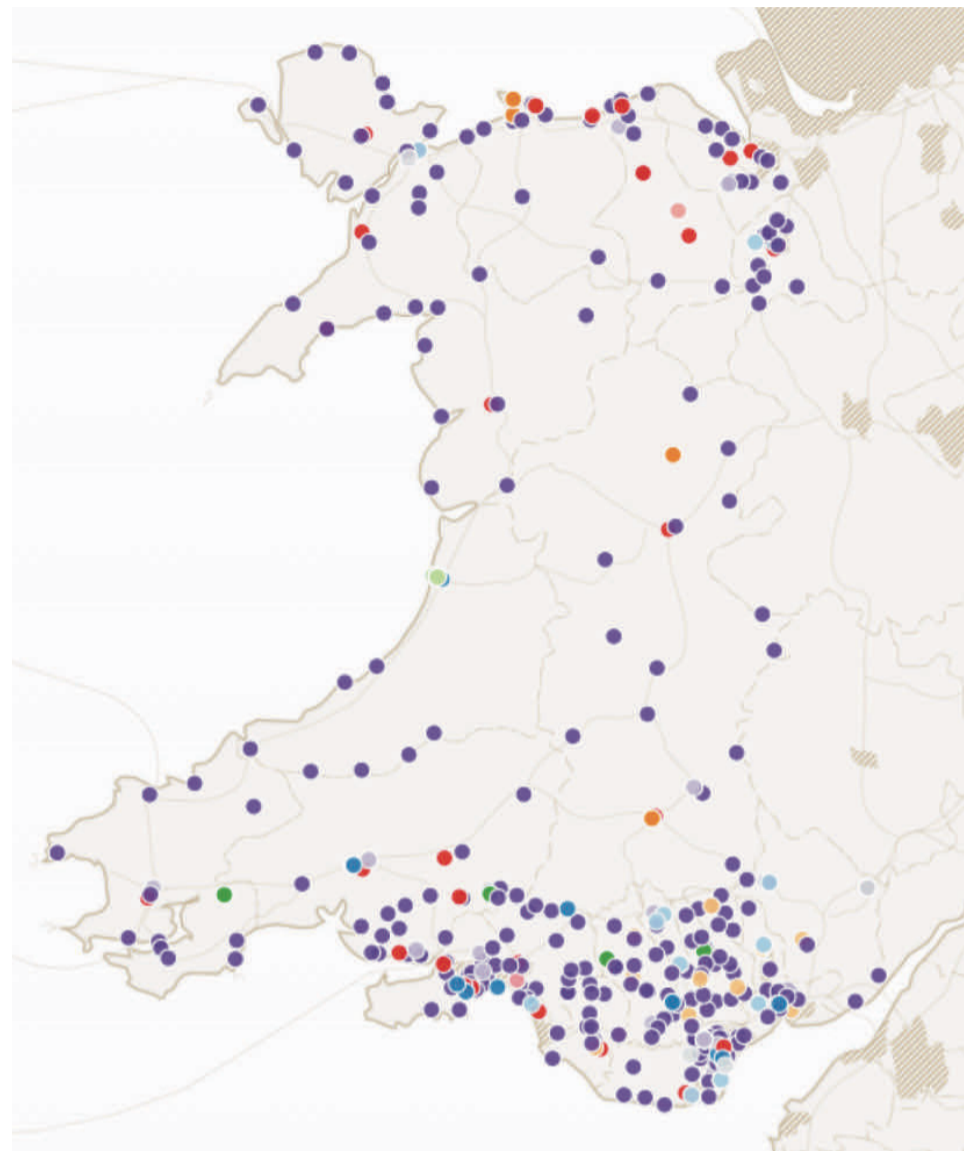
The Wikidata logo

Wikidata

Wikidata is a free database of structured data.

What does that mean?

Free means both free to get involved and publish data to, and free to use. The data on Wikidata is licensed under the **Creative Commons Public Domain Dedication 1.0**. Data publishers and data consumers can be confident



Libraries in Wales by library type, extracted from Wikidata and put on the map

in using Wikidata, as they have a clear licence that protects both parties.

Structured means that the data held in Wikidata has a clearly defined set of fields, allowing concepts (e.g. libraries) to be linked to one another.

A library, for example, will have a name, description, and coordinates; the coordinates in particular allowing it to be plotted on a map. It will be defined as a library, but can also be classified as a number of different types of library (academic, public, etc).

Libraries can be linked to their parent organisation (the *operator*), and the country in which they are located, as well as all sorts of other linked information. This allows for data on all those other things to be retrieved at the same time.

It also means that when retrieving data there is no issue in consuming data that has been edited by multiple people, or from many different sources, because we can be confident the same fields exist.

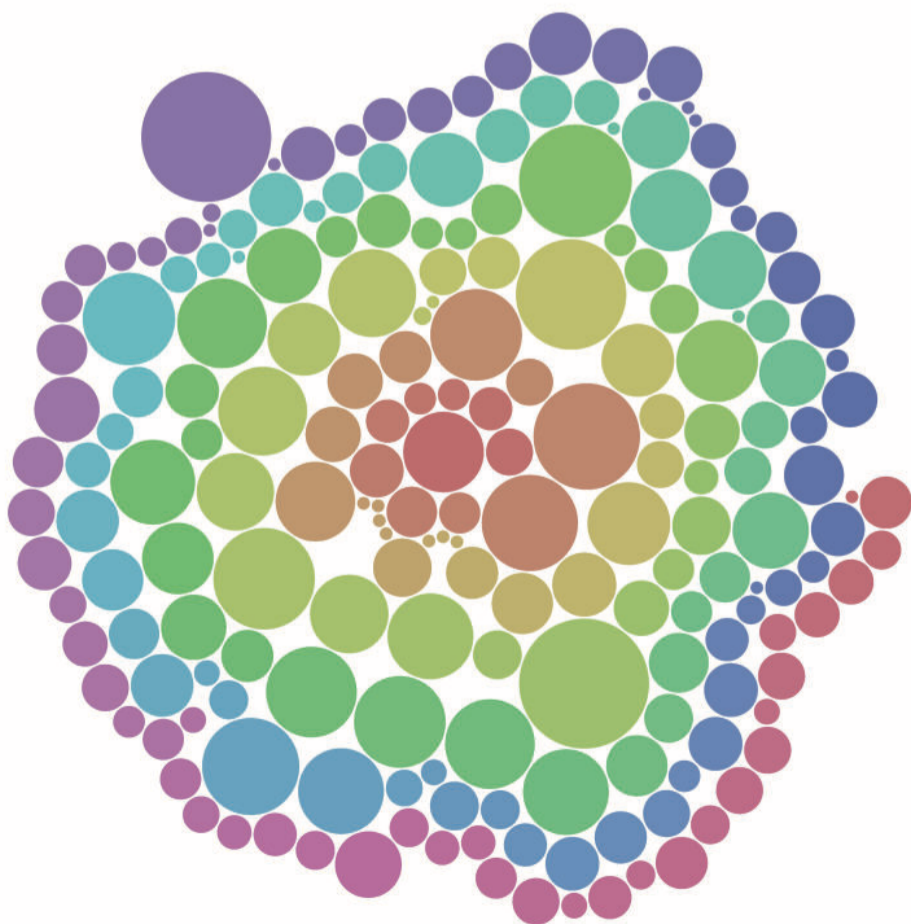
We could query Wikidata to see all the libraries, or be more specific and look for academic libraries in particular. We can select what fields we want to return, and what fields we want from linked data. It's a powerful tool to hold our data in.

Other reasons for Wikidata

Wikidata is part of the Wikimedia family (such as Wikipedia), and many of the concepts and principles that people will know from those projects will also apply to Wikidata.

It has a strong parent organisation, good branding, and most importantly: a huge community of enthusiastic editors, administrators, and people using the data.

These things all make Wikidata something worth investing time in, and learning about. And it doesn't look like it's going to go away any time soon so the question we should be asking is:



Counts of UK public libraries by operator. Can you guess which is yours? (No)

what can we do to make full use of it?

Data on libraries will always exist in Wikidata, and be available to the public. Library staff need the skills and confidence to edit this data, correct it, and take advantage of it by adding to it. Whether you have photos of your library that you wish to share, or just want to update an address or description, this is all possible using the Wikidata interface.

Wikidata and Wikipedia can be linked together. If your library has a great Wikipedia page then this can be linked to the library data that is held in Wikidata. This method of linking data, and making it discoverable, leads to plenty of opportunities to use the data from Wikidata in other applications.

Perhaps this could even lead to the library information displayed on a local authority website being sourced from Wikidata. Holding the data in one place

means editing it once and seeing it change everywhere.

Libraries already on Wikidata

In 2018 Kathryn Parry led an initiative to create a map of all libraries in Wales.

There would be ways of doing this without storing the data in any public

and open repository, but the work completed would be lost to any future developments. So after much effort on Kathryn's part in compiling a list together of various library types (public, health, school, national, and academic), and engaging with the relevant library organisations, this data was ready to be loaded into Wikidata.

Fortunately, the National Library of Wales has a Wikimedian in residence, Jason Evans, who was able to assist in the technical process of checking and loading that data.

Having done that, the data is now there and secure, and ready for anyone to maintain. This allows us to build on that work, and add more associated data, such as photos. And it serves as an inspiration to the rest of the UK, and other library sectors.

Since that project there have been further projects, such as the #Lluniau-Llyfrgell initiative (#PictureMyLibrary), instructing people on how to add their own photos to Wikimedia Commons, and link to these photos from the library data held in Wikidata. Clear, well-structured, openly licensed, linked datasets. What's not to love?

Later in 2018 an open data list from the Libraries Taskforce, listing public libraries in England, was also loaded into Wikidata, giving a further boost to the number of libraries being held on the platform.

Get involved

We'd love it if everyone was able to get involved in making an edit to Wikidata.

If learning comes through doing, then this is a perfect time for us to get going on Wikidata. We know there are loads of people who are enthusiastic about libraries, and who would love to update a central data repository that held data about their local libraries. The key is to get started!

Although Wikidata has necessary community standards and administration, even when people are aware of these it can understandably seem a little daunting to just start editing.

Everyone should be encouraged to use Wikidata, and to contribute their data to it. Wikimedia people are friendly and welcoming, and during the event we'll have Jason Evans on hand to help with queries if anyone is nervous! And even beyond that time feel free to reach out to us, or to Wikimedia to get involved.

Easy Guide To

Using Wikidata: Viewing and editing library data

Editing Wikidata

Wikidata holds large amounts of data about things you may know. You can contribute names, values, descriptions, coordinates, photos, and more.

Viewing Wikidata

Wikidata provides powerful query tools, enabling discovery of information about other library services, including the ability to search, browse and visualise this data.

These are the steps



#1

Go to [wikidata.org](https://www.wikidata.org)



#2

Discover your library



#3

Edit data



#4

Add data



Congratulations !

Once you've got started you can carry on maintaining your library data, adding new information, and uploading photos.

Source and Inspiration: Kathryn Parry and Jason Evans, Libraries in Wales <https://tinyurl.com/llyfrgelldata>

Library Data Stories

Most popular Harry Potter?

Plymouth Libraries have released open data on the books borrowed from their libraries in 2017.

2017 was a big year for Harry Potter. It was the 20th Anniversary of the release of Harry Potter and the Philosopher's Stone. In October, the British Library held their record-breakingly popular History of Magic exhibition, which included releasing a set of accompanying open data, taken from their collections and the British National Bibliography.

Harry Potter enthusiasts also calculated that 2017 was the year in which J.K. Rowling set the epilogue of the final Harry Potter book, the Deathly Hallows.

So, what were the most popular Harry Potter books in 2017, from analysis of the Plymouth borrowing data?

The results show that it is almost the same as the chronological order of the books: the most popular being the first book, and the least popular, the last.

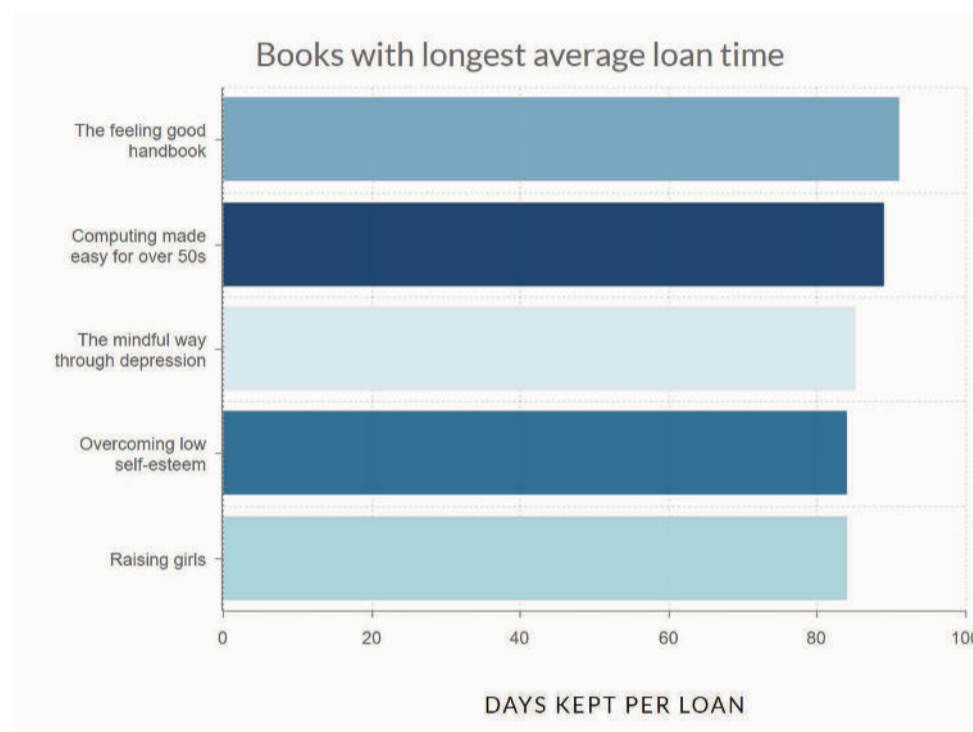
Does it represent people giving up on the books having got so far in? Or is it

tern was completely different.

Books people hold on to

Which books do people keep for a long time, before eventually returning to the library? Is it those that have the most pages and take the longest to read?

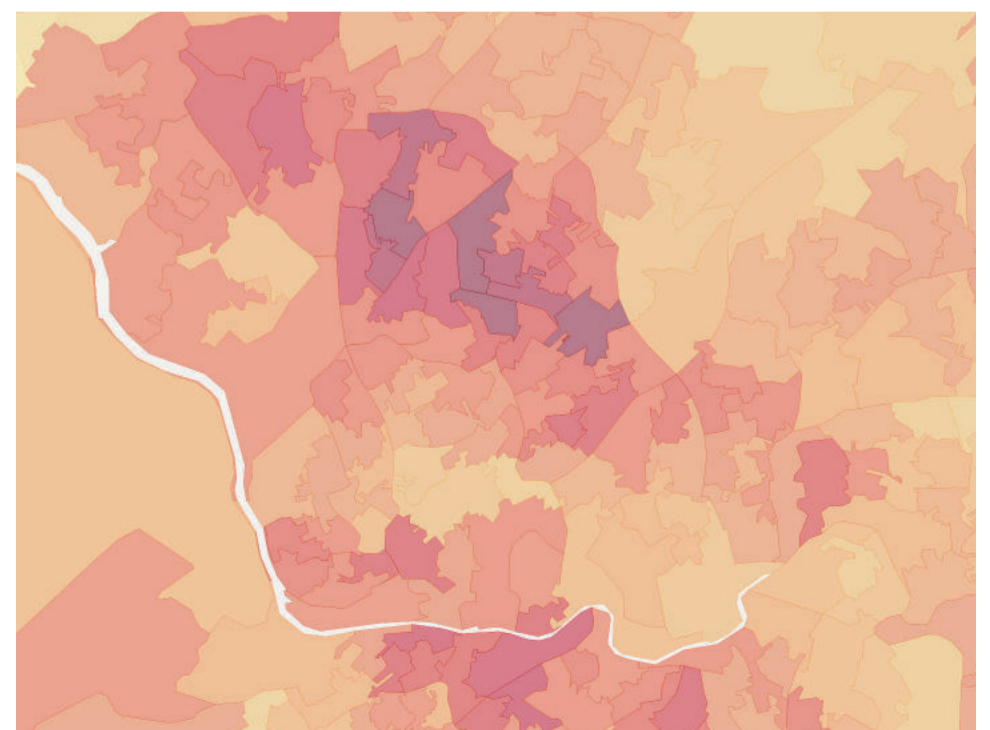
The top results were overwhelmingly books on managing well-being, or else helping with computing. Perhaps it is not so surprising that people spend a considerable time with these books, but it doesn't look to be related to the length of the book. The average loan length for these top books was in the region of 90 days, meaning people would likely have had to renew them at least 4 or 5 times. This kind of evidence may be worth considering when set-



Books held on loan for the longest time

ting loan periods and maximum renewal policies.

Mapping library membership



Map of LibrariesWest membership, relative to population, in the Bristol area

LibrariesWest is a consortium of library services in the South West, made up of Bristol, South Gloucestershire, North Somerset, Bath and North East Somerset, Dorset, Poole, and Somerset.

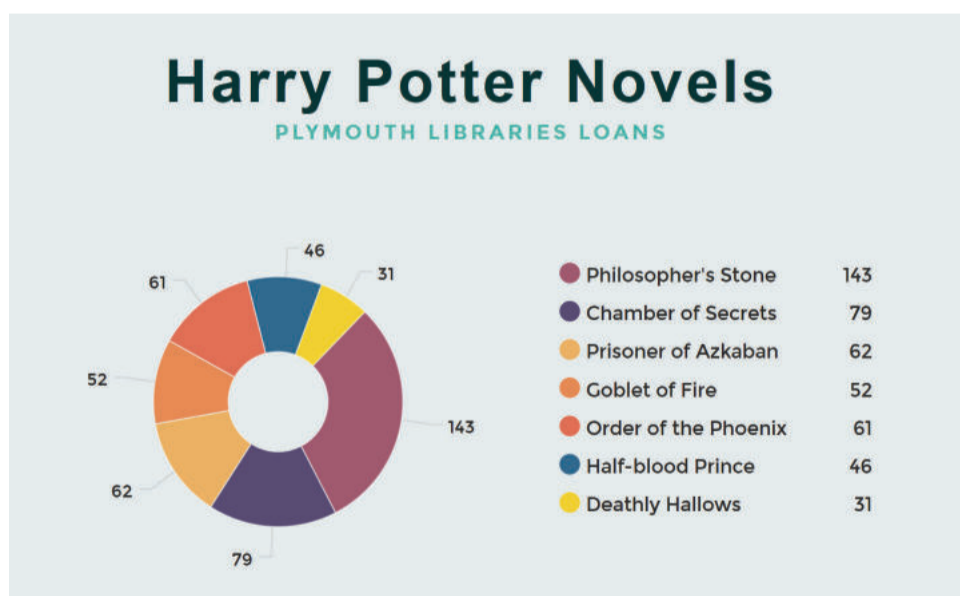
The services work together in many ways, but one particular aspect of the consortium is to share a Library Management System, including catalogue and membership database. Citizens in each authority can use libraries, and borrow books, from across the consortium, gaining access to a catalogue of 2.5 million items. This is either directly in libraries or through reservations.

LibrariesWest release open data on their library membership. The data includes active membership counts per Lower Level Super Output Area (LSOA). LSOAs are areas that are uniform in terms of population, so are useful for statistics. Typically an LSOA will represent about 1500 people.

So what does mapping this membership data show? Well, the 'hotspots' of library membership are those urban areas where libraries are located. This may seem obvious, but regardless of library locations, library services will be trying to attract membership from the whole population, not just those who live close to libraries.

This is also not just showing that areas of lower population have fewer library members. People from these areas are far less *likely* to use their library service. The differences can be as pronounced as 40% library membership in some inner city areas, and as low as 5% membership in rural areas.

Keeping an eye on membership mapping can aid in outreach. Being able to view data that is updated regularly could allow for results of initiatives to be immediately visible.



Harry Potter books borrowed from Plymouth Libraries in 2017

that there are still so many people starting out on the Harry Potter journey that the early ones will always remain the most popular? It might be interesting to compare with DVD stats to see if there are similar patterns of use between the different films.

Newcastle Libraries also release data on most borrowed fiction, and the pat-

An analysis of loan data, from the South West of England, listed the books people held on to for the longest amount of time. To be eligible, the book had to have been borrowed at least 100 times, and a number of days was calculated to show how long people kept hold of the book before returning it, on average.