

Making Everything Available: British Library Research Services and Research Data Strategy

Rachael Kotarski
The British Library

Torsten Reimer
The British Library

Abstract

The way that researchers generate, analyse and share information keeps evolving at a rapid pace. To ensure that it is well equipped to serve its global user base for years to come, the British Library is transforming the way it works too, from the physical buildings to its digital service portfolio. One key programme, Everything Available, will ensure the Library's continued support for research with services to enable access to information in an open and timely manner. This paper will describe the activities planned within Everything Available, with a particular focus on the aims of the Library's recently refreshed Research Data Strategy. It will give an insight into the challenges and opportunities faced by a National Library in providing relevant services in an 'open' world.

Received 21 January 2018 ~ *Accepted* 20 February 2018

Correspondence should be addressed to Rachael Kotarski, The British Library, 96 Euston Road, London. NW1 2DB.
Email: rachael.kotarski@bl.uk

An earlier version of this paper was presented at the 13th International Digital Curation Conference.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: <http://www.ijdc.net/>

Copyright rests with the authors. This work is released under a Creative Commons Attribution Licence, version 4.0. For details please see <https://creativecommons.org/licenses/by/4.0/>



Introduction

As the National Library for the United Kingdom, the British Library has had, since 1973, a statutory role in supporting research¹. During this – relative to an organisation with hundreds of years of history – short period, the research environment has changed significantly. In response, the Library had to change itself. Take for example the 24th November 1997, when the Library opened its building in Central London to readers². Original 1970's plans for the building to include a large card-catalogue space were all but obsolete by the time construction commenced. Instead readers were faced with computers, and the space designed for the card catalogue was subsequently occupied by the now iconic King's Tower³. The 20 years since the building's completion have again seen an upheaval in the way that research is carried out and supported. This change necessitates a rethink of the Library's physical and digital infrastructure, from the very nature of our collections to discovery solutions, digital preservation systems, repository services, machine learning and remote access.

The British Library's Living Knowledge⁴ vision for 2023 – the 50th anniversary of the British Library Act – is for the Library to be the most open, creative and innovative institution of its kind in the world. To achieve this vision, the Library is undergoing a period of change that will touch almost every process and aspect of the organisation. Living Knowledge will be delivered through five strategic change programmes:

1. 'St. Pancras Transformed' aims to adapt and expand the British Library's London site, making it one of the UK's most open, creative and innovative public spaces for the 21st century
2. British Library North will consider future plans for the Library's presence in Yorkshire and major developments in physical collection management
3. 'Heritage Made Digital' will set a strategy for our future priorities in the digitisation of our collections, considering researchers' needs both here in the UK and also internationally
4. 'Everyone Engaged' aims to grow the Library's reach to new audiences and communities across the UK, and deepen our engagement with our core users and partners
5. 'Everything Available' encompasses the continued development of the Library's services so that they remain relevant and supportive for individual researchers and research organisations in a rapidly changing digital environment.

1 British Library Act 1972: <http://www.legislation.gov.uk/ukpga/1972/54>

2 Tweet by @britishlibrary: <https://twitter.com/britishlibrary/status/933983761062072321>

3 History of the British Library: <http://www.bl.uk/aboutus/quickinfo/facts/history/>. The building was subsequently officially opened in June 1998.

4 Living Knowledge: The British Library 2015 – 2023: <https://www.bl.uk/projects/living-knowledge-the-british-library-2015-2023>

Everything Available

Along with society more broadly, the world of research is changing. New policy environments encourage or force changes in behaviour and generate new requirements from researchers. Governments and research funders actively support, and increasingly require, researchers to follow the principles of Open Science, including open access to research outputs and research data management. These are extending the kinds of information, support and training that our users are looking for. A focus on evaluation and the drive to demonstrate impact make research information management more important and require systematic tracking of research outputs and their use globally.

At the same time, key elements of the current scholarly communications systems such as pre-publication peer review and reliance on impact factor as surrogate of research quality are being questioned. Mega-journals aim to transform article publishing and data publishing is sometimes predicted to become as or even more important than the established journals; the humanities are debating the future of the scholarly monograph; and new, interactive and constantly updated forms of publishing challenge the established model of final ‘versions of record’. New publishing business models emerge, for example, consortial publishing where libraries cover the cost of open access publishing jointly so that authors can publish for free.

Pressures on funding coupled with increasing costs in some areas are forcing libraries to make efficiencies and seek new business models. The emergence of new (and breakthrough of more established) technologies change user expectations and the environment we work in: digitisation of content, mobile devices, instant remote access and data analytics are just a few examples. Technologies like artificial intelligence have the potential to change the concept of a library and the way librarians work even further. Physical space requirements are evolving too as reading rooms move from places for the solitary study of content to collaborative ‘office’ spaces for researchers.

This is accompanied by a shift from the established library model that has the library’s own collections at its heart to a service model that enables users to find, access, use and share content regardless of where it is held or what format it is in. Our current information services are based on most material being within our own collections. Our services (including reading room and document supply), our operations, the design and use of our buildings, staff roles and skillsets are largely still driven by this analogue model. As more and more people get their information from open access sources using mobile devices, we need to change our services. Also, the amount of data (in the widest sense) that is made available globally increases at a much faster rate than our collections. This means that, relatively speaking, the chances that we can meet the information needs of our users solely from our ‘own’ material are reducing. We therefore need to think much more on how we can help to make material outside our collections findable and usable, and how we can maintain usability of that material by contributing to its preservation.

These issues are challenging for any research library, but are arguably even more difficult considering that the British Library serves a global audience that extends well beyond higher education. The British Library provides services to any researcher who may walk through its doors, and widening this to any that may access its online services; is a further challenge. As well as those that you might call ‘traditional’ academics such as post- and undergraduates, lecturers and professors, the British Library supports researchers who may be family historians, authors, small business

owners, local government employees, marketing and management professionals, consultants or a mixture of any of these.

In order to develop new services, the Library needs to combine its own strengths, such as the provision of national services and metadata expertise, with the expertise of innovators and other organisations from the private or public sector.

Considering one of the main strengths of a national library – its longevity – the British Library is keen to work with others to help advance and sustain the global infrastructure for open science. This includes providing support for persistent identifiers, considering our role in digital preservation beyond our own collections and working with other organisations on distributed models for global infrastructure.

The Everything Available programme has been designed to enable the Library to meet these challenges through a rethink and redesign of its service portfolio, supporting individual researchers and research organisations. Specifically, its objectives are to:

1. Engage with researchers and research organisations to capture and understand their needs
2. Deliver improved discovery, access and analysis services for researchers of all kinds for both British Library collections and external content, thereby increasing usage and supporting creation of new knowledge
3. Improve the user experience of interacting with the Library's physical and digital collections and services, as well as those of third parties
4. Build the Library's capability to manage and share research outputs
5. Deliver new services for managing and connecting the digital collections of third parties (cultural, higher education, third sector and private sector)
6. Secure the long-term future of digital cultural artefacts and research outputs by ensuring content is stored in accordance with relevant preservation standards
7. Ensure the Library can serve the most appropriate version of content to both commercial and non-commercial audiences whilst meeting our legal obligations.

For more detail on the strategy behind Everything Available see (Reimer, 2018).

British Library Research Data Strategy

One suite of services within the remit of Everything Available will be those delivered as part of the Library's Research Data Strategy. Following on from an initial data strategy that ran from 2010 to 2015, the vision of the refreshed strategy is to see research data as embedded into the Library's collections and services as text is today.

While developing the data strategy, we examined the strategies of peer organisations for statements on research data. This included the strategies of 14 other national libraries, where we found that at the time, half of those examined did not explicitly mention data at all. Given the continuing importance of data as a resource, and the increasing recognition of data as a valuable output in policies and activities ranging from the Research Councils UK policy on open access to research data⁵ to the

⁵ RCUK, (now UKRI's) Common Principles on Data Policy: <https://www.ukri.org/funding/information->

G8 open data charter⁶ and more recently the FAIR data principles (Wilkinson et al., 2016), we felt it important to expand on the Library's commitment to research data expressed within Living Knowledge by refreshing and updating the research data strategy to ensure it supports and aligns with these initiatives where possible.

Themes were used to structure the activities planned for the British Library's data strategy around four core elements: data management; data creation; data archiving and preservation; and data discovery, access and reuse. For each of these four themes we developed a series of user stories. These stories demonstrate – at a high level – the kinds of activity a successful research data strategy would enable.

Data Management Theme

As an independent research organisation, the Library is developing its policy and practices around the data created as part of its research activities. The policy reflects on the kinds of data produced, the context in which it is created and the existing data policies and plans of collaborators. Training for British Library staff will also be key.

Our hypothetical user is a British Library staff member who is able to use a template data management plan to write her proposal (Figure 1).

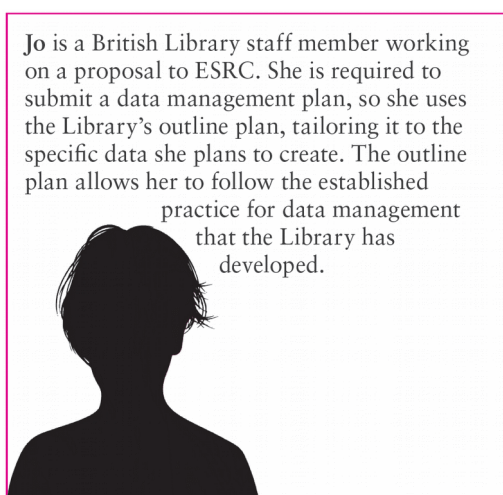


Figure 1. Data management example from the British Library's Research Data Strategy 2017.

Data Creation Theme

The Library's collections are a rich source of historical and social data, and can provide an important historical and social context for other data. Examples of the Library's data are available through data.bl.uk, and include datasets formed of collection metadata (EThOS, 2016); images of digitised content (British Library Labs, 2016a) and their text (British Library Labs, 2016b); extracted data from the web archive (The UK Web Archive, 2013) or datasets derived for specific projects (Do et al., 2017).

[for-award-holders/data-policy/common-principles-on-data-policy/](https://www.gov.uk/government/publications/open-data-charter)

⁶ Open Data Charter: <https://www.gov.uk/government/publications/open-data-charter>



Figure 2. Data creation use case, from the British Library Research Data Strategy 2017.

The Library will continue to promote and explore the value of its collections as data, providing opportunities to open up collections in ways that make their content more widely available but also presenting challenges in ensuring the onsite services and training currently provided can support these new activities.

Our hypothetical user demonstrating the vision for data creation is an epidemiologist, looking at long-term prescription trends of particular drugs, and the impact that the media has on those trends (Figure 2). Our example also expresses the desire to be able to make user-derived datasets available to other researchers.

Data Archiving and Preservation Theme

As the Library begins to create more datasets, and as data is ingested, the Library needs to ensure it is able to archive and preserve data content as well as its traditional text and image-based collections. The Library will explore how its systems and expertise can be leveraged to provide data archiving and preservation services to third parties, in conjunction with its existing services and those developed as part of wider Everything Available activities.

There are already services at the British Library where preservation could be better integrated, including the e-theses online service (EThOS)⁷ and data.bl.uk. We are actively looking at bringing these separate services together within a single multi-tenancy platform. This will support the preservation of their content, as well as better methods of user discovery and access, including text and data mining.

The user story for data archiving and preservation reflects the need for better preservation of the variety of research and its outputs that are published in an ad-hoc way. A preservation service would also allow the creators of data to guarantee any persistent identifiers they may assign.

⁷ EThOS e-theses online service: <http://ethos.bl.uk>

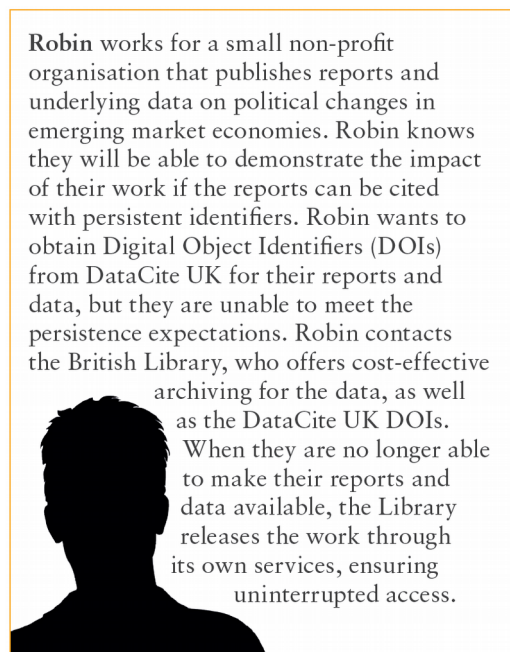


Figure 3. Data archiving and preservation use case, from the British Library Research Data Strategy 2017.

Data Discovery, Access and Reuse Theme

As new methods of analysis and visualisation are constantly developed, the Library will need to find ways of providing relevant data tools and training not just for users but also for staff. Whether enabling researchers to bring their data to the Library or allowing them to bring Library-held data to their analysis, new forms of access need to be developed. The sheer file size of some datasets will make them difficult to provide to offsite users without a rethink of the digital delivery service for that item. And while providing access to the physical items, a large map for instance, in a reading room may require an appropriately-sized desk and guidance on handling techniques, analysing geographic data on a reading room terminal may require training in the structure of the data, or even basic coding knowledge.

Our discovery, access and reuse case study hinges on the support we need to provide to users on site for research data, but also the availability of analysis tools that will allow appropriate use of data collections (Figure 4).

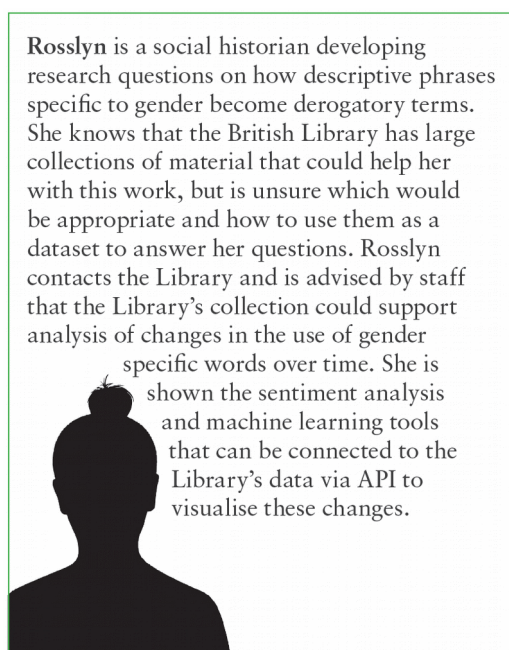


Figure 4. Data discovery, access and reuse use case, from the British Library Research Data Strategy 2017.

Maintaining Momentum

The Library is already active in research data through activities such as DataCite UK⁸, projects such as THOR⁹ and FREYA¹⁰ and by providing its collections as data through data.bl.uk. There will be work to do to ensure that these activities not only remain sustainable but become business-as-usual for a national library. The approach to realising some of the user stories presented here, and ensuring they remain sustainable, may well change over the coming years. The strategy is purposely high-level, enabling us to remain agile and reactive to new technologies and opportunities as they arise.

Final Thoughts

Many aspects of Everything Available and the Research Data Strategy will pose new challenges to the Library's technology, physical infrastructure and frontline staff. Staff training is therefore a key requirement. In turning our aspirations into business-as-usual activities, we will certainly learn a lot about the skills and knowledge required by frontline library staff to support these services. We will also learn more about the technologies, skills and knowledge required by support and curatorial staff. We aim to share this knowledge where we can.

Partnership and cooperation will be key to delivering both Everything Available and the Research Data Strategy, and these challenges are not all particular to the British

⁸ DataCite UK: <https://www.bl.uk/datacite>

⁹ THOR: Technical and Human Infrastructure for Open Research: <https://doi.org/10.5438/6423>

¹⁰ FREYA: <http://www.project-freya.eu/>

Library. So we are keen to work with cross-cutting organisations such as CENL to explore new models of support from national libraries, as well as continuing the dialogue and exploring the opportunities with the IDCC community as a whole.

References

- British Library Labs. (2016a). Theatrical playbills from Britain and Ireland. British Library. doi:10.21250/pb1
- British Library Labs. (2016b). Theatrical playbills from Britain and Ireland. (OCR text only). British Library. doi:10.21250/pb2
- Do, B., Zhao, L., & Wang, K. (2017). SherlockNet. British Library. doi:10.21250/sherlocknet1
- EThOS. (2016). UK Doctoral Theses (EThOS) Abstracts and Metadata - 01/03/2015. XLS. British Library. doi:10.21250/ethos1
- Reimer, T. (2018). The once and future library: the role of the (national) library in supporting research. *Insights*, 31, 19. doi:10.1629/uksg.409
- The UK Web Archive (2013). Host Link Graph – JISC UK Web Domain Dataset (1996-2013). The UK Web Archive. doi:10.5259/ukwa.ds.2/host.linkage/1
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. doi:10.1038/sdata.2016.18