
From stored knowledge to smart knowledge

The British Library's content strategy
2013 - 2015

Lucie Burgess
*Head of Content Strategy,
Research and Operations
The British Library
lucie.burgess@bl.uk*

The British Library is the national library of the UK and is one of the greatest libraries in the world. Our purpose is to make our intellectual heritage accessible to everyone, for research, inspiration and enjoyment. In order to achieve this, we build, curate and preserve the UK's national collection of published, written and digital content; we make our collections and expertise accessible to researchers from all disciplines; and we partner with others to connect people with information wherever it is held.¹

The British Library's collection consists of some 150 million items and at 693km of shelf space would stretch from London to Paris and back. It encompasses 14.5 million printed books; over 800,000 journal titles; 59,000 newspaper titles from the UK and overseas; rare and precious manuscripts from the Lindisfarne Gospels to the Beatles lyrics; music scores, plays, stamps, patents, prints and drawings, maps, sound and moving image; and exists in formats from ox-bone and parchment to web archives and born-digital audio-visual files.

The British Library has a responsibility to maintain the UK published archive, fulfilled through UK legal deposit legislation² which empowers us with five partner libraries to collect a copy of every print and digital publication published in the UK, including websites. We commit a significant annual budget to purchased acquisitions, £14.9m in 2012-13.³ We also collect through significant voluntary arrangements (for example, UK pop music) and donations (for example, some literary archives). We fundraise to purchase significant heritage items, such as the remarkable St Cuthbert Gospel.

WHY DO WE NEED A CONTENT STRATEGY?

In 2012 we set out to re-examine our collection development strategy, which we have articulated as a 'content' strategy – *From stored knowledge to smart knowledge*. We use the word 'content' rather than 'collection' to explain that we both collect, and, increasingly, connect to content in an ever more digital and open world. Ultimately, strategy is about making choices in a world where resources are limited, and where demonstrating value from public money is essential. We wanted our content strategy to answer questions such as:

- What should be the balance between collecting from the UK and from overseas, in different formats, and in disciplines / subjects in order to best support research?
- What should be the balance between collecting for future generations of users, and contemporary collecting / connecting to support services for today's users?
- How much digital collecting should we do? What impact might this have on digital infrastructure, operations and resources?
- What new opportunities and innovations could we exploit?
- What should we prioritise, do differently or stop doing altogether?

While we did not develop definitive answers to each of these questions, we developed a set of guiding principles which each of the library's content experts now uses to inform their everyday collection development activities. Through the strategy we aim to demonstrate that our collections are at the heart of the British Library and the vital role they play in underpinning UK research, economic growth, culture and learning.

The UK is particularly rich in research-level publishing, producing some 6% of the world's research articles.⁴ The British Library's content continues to underpin the excellence of the UK research and higher education, facilitates significant cost-savings in research libraries through acquisition of unique content and initiatives like the UK Research Reserve⁵ and so helps to maintain the UK's research competitiveness internationally.

The information world is undergoing profound change. It will be obvious to librarians that the information world continues to undergo unprecedented change:

Users want their libraries to be somewhat like Google or Amazon. Rapid technology innovation (such as internet shopping, mobile devices like tablets and smart phones, social media) has transformed user expectations. Libraries are expected to provide an 'anytime, anywhere' experience, independent of device and, unfortunately, they often fall short of what their users expect.

Researchers have a strong preference for material that can be accessed online. The main UK publishers deliver content on the web; most subjects have abstracting and indexing services with click-through access; e-book publishing is growing rapidly. There are lots of different models: free, rental, subscription and streaming through many players including Google Scholar, OverDrive and Scribd. Digital content brings huge users benefits such as text mining and seamless search, but adds complexity for libraries.

Content is becoming even more complex and dynamic. User research conducted by the British Library suggests that many formats of content are important to researchers;⁶ such as web sites, audio-visual content, grey literature, geospatial information, statistics and research datasets, user-generated content and social media. By 2020, the British Library estimates that 70% of all content globally will be published in a digital form.⁷

Open Access is changing research communication. Following the Finch Review,⁸ which recommended an open access (OA) mandate for UK publicly-funded research, a variety of OA models are being implemented, and there has been much debate about the costs and impact on libraries, publishers and researchers. The British Library already provides access to some OA journals, books and datasets, and we played a key role in the development of Europe PubMed Central⁹, a subject-based repository of OA journals.

FROM STORED KNOWLEDGE TO SMART KNOWLEDGE

The library will develop the collection primarily in terms of disciplines and subjects

The framework of disciplines (Arts and Humanities, Social Sciences, and Science, Technology and Medicine) and subjects will lead our collection development decisions because this enables us to be responsive to the user communities we serve.

We aim for coherence within the disciplines, collecting in all key subjects as far as possible, prioritising some subjects according to UK research priorities, usage and national provision; and

recognising that a comprehensive approach is neither financially sustainable nor appropriate. We prioritise collection of unique content, whether in physical or digital form.

We will continue to develop format-based collections, which support research in all disciplines. These include web archives, manuscripts, maps, theses, official publications, news media, sound and moving image content. For example, we have collected 14,000 instances of websites through the UK Web Archive¹⁰ that reflect the diversity of lives and interests throughout the UK. In order to promote inter-disciplinary use of our collections, we are encouraging matrix working between subject curators and format experts – for example, between curatorial specialists in politics and web archives.

The intellectual and cultural heritage of the UK is central to the British Library's content strategy, but we will continue to collect overseas content of value to research

As the UK's national library, the intellectual and cultural heritage of the UK is at the heart of our collection development through legal deposit. The British Library has collected a copy of every significant UK book, newspaper, journal and other printed publications for over 300 years. On 6 April 2013, new legislation was passed which empowers the British Library and five other libraries to harvest websites and to collect digital publications that are published in the UK and preserve them for ever. We have already collected 80,000 e-journal articles and are scaling up our e-book collecting. We have collected a snapshot of the entire UK web space – some 3.9 million websites in the '.uk' domain. We are also conducting web crawls on themes such as the reform of the National Health Service, Scottish independence and the death of Nelson Mandela. We will continue these activities. It is important that this ephemeral digital material is not lost to a digital black hole of the 21st century and can be used by researchers and the public.

In any given subject we will supplement legal deposit content, voluntary deposit and donations through contemporary acquisitions, both from the UK and overseas. Our collecting is not defined by geography, but rather by subject. Our coverage of subjects such as European studies, Asian and African studies and Australian and American studies illustrates our global outlook and the UK's relationships with the rest of the world.

We collect actively material of major significance to UK heritage and the UK's relationship with the rest of the world. A recent example is the draft score of Benjamin Britten's much-loved *The young person's guide to the orchestra*. Increasingly we collect digital heritage material as well as ancient manuscripts or books; for example, we recently acquired the digital images, text and data in the archive of the evolutionary biologist William Donald Hamilton.

We will collect more digital content to reflect publishing trends

We will prioritise collection of and connection to digital content over print as a strategic response to the accelerating pace of change in the external environment. We will move to digital quickly for contemporary collections, particularly for journals; although we will continue to collect print (and digital) for UK heritage collections such as literary fiction.

We will continue to license and subscribe to digital content in order to provide a better user experience both within the library's physical spaces and on the web. We will collect and connect to freely available but dynamic or unpublished online content. For example, we provide the growing EThOS¹¹ service in partnership with UK universities, which gives free online access to over 300,000 UK theses, both to born-digital versions and through a digitisation-on-demand service. We are collecting digital audio-visual content excluded from legal deposit; in January 2013, we launched a new multimedia 'Broadcast news' service, which enables reading-room users to access TV and radio news bulletins produced by 15 broadcasters since May 2010.

In order to collect and connect to digital content at scale we are making substantial improvements to our digital infrastructure, from acquisition through to preservation and access. For example, we are developing new highly automated workflows to acquire digital content direct from publishers, to ingest it into our digital library system and make it searchable through 'Explore the British Library' (explore.bl.uk/).

We will increasingly connect to digital content and embrace openness

It is essential that we make our content available through online channels if we want to remain relevant. At the British Library, we are achieving this goal through a mixture of through-linking to open resources; subscription and licences; through new licence models that allow delivery direct to

users from publisher websites, for content which the library does not wish to hold; through mass digitisation of physical holdings; and through digitisation-on-demand services.

The British Library already engages in and supports open access in its many forms. We provide integrated discovery via our online search to selected open access journals, books and datasets. Increasingly, we intend to link to open content on the web as far as possible – for example, providing links to journal articles published under Creative Commons licences in university repositories or on publisher websites. Through UK legal deposit, we will also ensure comprehensive long-term access to this content.

We will continue to digitise our physical holdings, funded through commercial partnerships and philanthropic donations. Where possible, we are making digitised content available via open licences. For example we have made 60,000 digitised 19th-century books freely available as public domain content and recently uploaded one million images to Flickr Commons for full re-use. In future, we would like to provide links to openly available digitised books where we hold a physical copy, through services like the Hathi Trust. We are also trying to make our data, as well as content, as open as possible; for example, the British National Bibliography, which is available as an open linked dataset.

We will add value to content through curation and encourage our users to add value through community

We want to engage with our users to supplement our curatorial expertise and in so doing, add value to content. Through the UK Soundmap project,¹² launched in July 2010, we asked people to record the sounds of their environment, whether at home, work or play. Since then, over 2,000 recordings have been uploaded by some 350 contributors. Through the British Library's Geo-referencer project¹³ 2,400 scanned maps of 17th-, 18th- and 19th-century England and Wales were tagged with modern geospatial co-ordinates through crowd-sourcing in less than a month.

The Library's 200-strong curators and content professionals come from all disciplines and subjects and are experts in their research fields. Our expert staff research and interpret the collections, understand their provenance and provide context; they are data scientists, historians, project managers, diplomats and a bridge between the ancient and modern. They negotiate a complex world

of copyright-controlled and open content, paid for and free, physical and digital, and their work spans the disciplines.

We want to find better ways to connect with our users, in a meaningful and sustainable way. We want to crowd-source more of the rich expertise of our user communities. The combination of curated resources with community will provide an increasingly important catalytic process in enriching our content.

CONCLUSIONS

From stored knowledge to smart knowledge represents evolution rather than revolution for collection development at the British Library. But the period to 2015 is only a part of the journey. We are planning for an even greater scale of delivery of digital content directly to researchers in the future. This requires continued ingenuity and innovation, and investment at a time when budgets are under considerable pressure. To 2020 and beyond, we expect to see more dynamic changes; greater collection of apps, of dynamic research datasets, more mobile content, social media, audio and video. But there will be continuities as well as change. Over time, when rights and law allow, we hope that users will be able to engage with much more of the library's extraordinarily rich content to support research, innovation, culture and inspiration.

The British Library would be interested to hear your feedback on our content strategy, and to hear perspectives from your own library. For more information, please see http://www.bl.uk/aboutus/stratpolprog/contstrat/british_library_content_strategy_2013.pdf or please contact:

email [lucie.burgess@bl.uk](mailto: lucie.burgess@bl.uk)
Twitter @LucieCBurgess

REFERENCES

- 1 Extract of British Library statements of purpose, developed in 2013, as yet unpublished
- 2 The Legal Deposit Libraries Act 2003 and the Legal Deposit Libraries (Non-Print Works) Regulations, 2013
- 3 British Library Annual Report and Accounts, 2012 13, p. 27 (<http://www.bl.uk/aboutus/annrep/2012to2013/annualreport201213.pdf>)
- 4 Publishers Association, 6 Bridging the valley: written evidence submitted by the Publishers

Association [to the House of Commons Select Committee on Science and Technology]', April 2012

- 5 See <http://www.ukrr.ac.uk>
- 6 From regular reading room and web surveys
- 7 Research conducted by Outsell on behalf of the British Library, for the BL's 2020 Vision, unpublished, 2010
- 8 *Accessibility, sustainability, excellence: how to expand access to research publications: Report of the Working Group on Expanding Access to Published Research Findings*, Chaired by Dame Janet Finch CBE, published by Research Information Network, 2012
- 9 See europepmc.org/About
- 10 The UK Web Archive is a partnership between the British Library, the National Library of Wales, the Wellcome Library and Jisc see <http://www.webarchive.org.uk>
- 11 See ethos.bl.uk
- 12 See sounds.bl.uk/sound-maps#
- 13 For more information see www.bl.uk/maps/georefabout.html