British Library strategic developments

Based on a paper given at the 26th UKSG Conference, Edinburgh, April 2003

This paper provides an overview of recent strategic developments at the British Library. It highlights the external factors which have influenced the shaping of the strategy and outlines the key strategic themes for the Library. The BL's role in relation to serials collecting, long-term preservation and archiving, rights management and document supply in an electronic world are also discussed.

LYNNE BRINDLEY
Chief Executive,
The British Library

Introduction

I have worked in the information and library world for a long time now, and it seems to me that we are going through the biggest period of change that we have ever seen. Certainly, it is a time in which all our existing norms are being seriously challenged. Here are a few headlines that have appeared in publications over recent months:

> *'Second thoughts on bundled e-journals'* – is that what you really want?
> *'Commercial copying as clear as mud'* – well, perhaps it is getting clearer in the UK, but who knows?
> *'Counting the cost of on-line optimism'* – how do you tell the difference between hype and reality?

These are just a few of the issues facing us all, every day of our working lives. I want to talk about some of the key drivers for change and then put up a straw man for discussion. I will then discuss three major areas where the British Library is focusing its development efforts.

Key drivers for change

I suspect I don't need to tell most of you about changes in research behaviour, but what is becoming clear is that big science has become even bigger, with many international researchers co-ordinating specific projects. This is particularly noticeable in physics and biomedicine. Activities here are also moving towards global collaboration. Interdisciplinarity is also increasingly common. I am a member of EPSRC (the Engineering and Physical Science Research Council) User Panel and the Council has re-labelled itself the IRC, the 'Interdisciplinary Research Council'. I think that the way in which disciplines are developing is critical for libraries. Another trend we are seeing is the development of e-communities and portals, offering new types of information services, built on improvements in technology and the social need for communities and kinship across disciplines.

The days of the library of record have, I think, long since gone for any single library - the Library of Congress is, perhaps, an exception. Libraries are being forced to operate within a decreasing budget allocation, which is something most of you will be familiar with. But there is still a frustration gap, despite site-licensing activities cutting the unit cost of information. Resource sharing and consortia developments have arisen to optimise the group holdings, but that does not help those millions

world-wide who do not have an affiliation with a major library. One of the questions for us is how we reach that potentially disenfranchised group, including some of the professions, distance learners, scientific hobbyists and so on.

Changes in the publishing scene will also be very familiar to you. We are seeing a lot of acquisitions and mergers, with concentration particularly on science. However, there remains a very long publishing tail and much of the important material is being published by small, distinctive publishers. Concerns over current business models are shared by everyone. There appears to be an increasing mismatch between the creation of information, growing at rates largely determined by R&D spend, and the purchase of information, determined largely by local institutional circumstances and by overall government funding of an institution. The two quite simply do not relate.

Policy-makers and librarians are concerned at the haemorrhaging of academic and research funds out of the system into the hands of large commercial publishers; this has led to ventures such as SPARC and the Public Library of Science initiatives. New business models are being experimented with, initially through subject specialist e-print servers, moving on to open access and, more recently, to institutional repositories of open archives.

Copyright remains difficult. Whether the author, the institution, or funding organisation has ownership of the research output remains an area of dispute, as does the distinction between what is fair use and fair dealing. The European Directive brings an added complexity of determining what is actually commercial use. Authentication and authorisation become more important as the network of textual and non-textual resources continues to expand. New ways of identifying the rights of users, current and historical, need to take into account all of their rights, whether they are society membership, university, faculty, editorial board membership, or others.

These are just some of the many challenges facing us today and it is within this environment that I am leading and positioning the British Library. A message I would like to emphasise is that we all need to work across traditional industry boundaries to solve the challenges we face.

Information environment 5-10 years hence

But let me try out a straw man on you. How might our information environment look in 5 to 10 years' time? These are some of the characteristics I think we will see:

- extensive networking of electronic access
- heavy reliance on digital data sets
- content providers offering greater multimedia packages
- over 50% of refereed scholarly material born digital
- fixed price, rather than pay-per-use business models
- libraries as electronic gatekeepers and navigators
- publisher consortia larger, but reduced margins
- digital archiving/preservation taking on critical role.

You might think this is more like the present day, not five to ten years hence; certainly every day this 'vision of the future' seems to get much closer. We already have extensive networking of electronic resources. The estimated 22 million internet users in 1997 will have risen to 162 million in 2002 and 2.2 billion by 2010, and much of this is increasingly high speed and broadband networking. There is a huge dependence on digital data sets and manipulable software, which is a trend I think will continue. New multimedia resources are being developed in rapidly moving STI areas and I envisage this spreading to other disciplines over the coming years. I think we will also see a significant increase in the amount of published refereed material which is born digital, rising to over 50% in five years' time.

Business models, I might argue, will follow the pattern of other industry sectors in which a fixed-price strategy is employed, rather than being based on usage target numbers; an analogy here is the telecoms industry. The role of electronic gatekeepers and navigators is one that many libraries already perform, but will become increasingly important in the future. Looking to the publishers, I imagine we will see an increase in the development of large scholarly publishing consortia as commercial publishers absorb more of society publishing activities. Margins will reduce, leading to intense competition. The final challenge, digital preservation activity, will, I think, move from being an experimental, fringe activity into an embedded

reality, unlocking the e-potential of both academic publishers and libraries.

Implications for the British Library

All this has major implications for the British Library. New roles will be required of us, and must be backed up by our own strategic developments. I would like to share three of these with you: digital acquisition, digital article services, and digital preservation and archiving.

Digital acquisition

The challenge for the national library, more than for any other kind of library, is to guarantee the integrity of the national published archive. At the British Library, we are doing an enormous amount towards that end in the digital world. We have had a voluntary scheme for depositing electronic material since January 2000, administered by the Joint Committee on Voluntary Deposit and supported by a publisher code of practice. This has been relatively successful, attracting the participation of over a hundred publishers who are increasingly depositing online material. We estimate that roughly 90,000 items have already been deposited, but there was general agreement, both with the libraries and the publishers, that a statutory scheme was necessary to guarantee the long-term integrity of the national published archive. It is these discussions which have led to various stages of legislation for the UK.

Extending legal deposit

A Private Member's Bill to extend legal deposit successfully passed its second reading on 14 March 2003. The Bill goes to committee stage in June and then, we hope, to the House of Lords for a third reading in July, with Royal Assent potentially scheduled for September. Whilst I am not sanguine about the work which remains ahead, I have been heartened by the level of support we have received so far – from all parties and all sectors. This legislation, if we succeed, will enable the national and copyright libraries to require deposit of digital materials, and thus to fill a long-term role in providing preservation and access.

In addition to voluntary deposit, the BL purchases a significant amount of commercial digital material. Purchase and licensing of e-publications for the reading rooms is consistent with our development policy and we are already spending some £1.5-2 million each year in this area. We also have major licensing agreements with a range of publishers to hold parallel e-formats of high-demand publications already held in print. This is primarily for service reasons, to serve both our reading rooms and our remote users. We hold well over 3,000 of the highest-use titles, with another 1,500 in the pipeline, so will shortly have approximately 4,500 titles for direct supply under licence.

Web harvesting

We are also engaged in the area of web harvesting, albeit with a degree of trepidation. Last year we undertook a pilot project to harvest a representative sample of 100 UK sites spread across the Dewey classification. We obtained rights-holders' permissions and crawled the sites regularly, using Blue Squirrel software, to gain first-hand experience of the technology, the selection and the ensuing management issues. We have also submitted a bid to Government, and will be commencing a wider UK harvesting programme, with the proposals based on combining snapshots of the UK domain and selective harvesting of up to 10,000 research interest sites.

We are working closely with international partners including the Library of Congress, the Bibliothèque Nationale de France, and colleagues from Scandinavia and Canada, to agree distributed global approaches to web archiving. The consortium is likely to be working with the Internet Archive to develop next generation web crawlers and acquire retrospective content from the mid-nineties. We are also working with UK partners, such as JISC, the National Archive and the Wellcome Trust to refine our UK website collecting policy, taking into account important UK collaborations.

In addition to the acquisition of digital materials, we are digitising analogue items from our collection. The BL has, for example, created a new resource as part of a major NOF lottery-funded project ('Collect Britain'- www.bl.uk/collectbritain), and we have also just received from JISC funding of around £3 million to digitise material from our sound archive and historical newspaper

collection. I believe these initiatives will be extremely useful and important to both academics and the wider public.

Digital article services

Let me move on to the area of digital article services. Acquiring digital materials and agreeing publisher licences to store and use those materials is clearly the first step, but we must also establish new types of services and the supporting infrastructure required to develop and deliver digital materials to users. Users expect to have instant e-access to articles and services. Publisher licence agreements now give us the mandate to move ahead, but we want also to provide the foundations for electronic delivery by modernising our document supply operations. We are in the process of installing new equipment, such as scan-on-demand, on a major scale. This will support new processes enabling us to deliver items directly to customers' desktops. Whilst the demand for traditional paper-based document supply may well decline over time, we envisage strong continuing demand for the services in electronic form.

There have been plenty of studies and library experience which suggest that document delivery does not compete with subscription sales. The view we are taking with the publishers licensing with us is that it offers them new opportunities to reach audiences outside the current concentrated academic research environment. Copyright fees are agreed, set and transparently projected by the publishers, with a potential new revenue source and a new commercial market.

A recent exciting development for the BL is a partnership with Adobe and Elsevier, enabling us to provide users with copyright-protected pdf versions of Elsevier articles on a strict pay-per-view basis. Commercial customers get the service they want and publishers' rights and revenue streams remain protected. Secure electronic delivery allows quick delivery to the user, with no onward transmission. We very much hope that more publishers are going to follow suit and take advantage of this important opportunity.

We are also using 'e' tools to reach those users who lie outside the subscription model market. Our value-added tools include alerting services, covering 20,000 to 25,000 titles, through to the provision of information experts. We are including

abstracts in many of the 2,000 titles in our electronic table of contents services since this increases the likelihood of their being requested and, therefore, of a copyright fee being generated. We are also linking between aggregated secondary files and document delivery, to provide seamless delivery of full text articles.

Digital preservation and archiving

The third area of our challenge is digital preservation and archiving. As well as providing digital services, we must make our publications available to researchers in future centuries, a responsibility the IT industry finds hard to comprehend! Developing preservation and archiving strategies is a key responsibility of national libraries, so that we can ensure future generations will be able to access the digital materials we collect. The British Library is preparing the ground, with a number of different initiatives. We have already defined a strategy for digital object management, based on perpetual preservation of original items. Our overall aim is to preserve and provide access, with appropriate permissions, to the original bit stream of digital objects and associated metadata. This will enable future researchers to decode or interpret the originals in the same way that they look at original artefacts today.

We are starting with e-documents, defining the migration strategy for each media type and cataloguing items in uniform fashion for consistent retrieval. Collaboration with other institutions is key; we are working with the Library of Congress, MIT, Cambridge and the Dutch National Library to develop and implement standards and systems and to ensure that the world's literature remains accessible to those who have a right to use it. We will be investigating how to synchronise our collection strategy so that we can share the total load amongst other major players in this business.

As part of the Joint Committee on Legal Deposit of e-materials, the BL has worked closely with the Publishers Association and with ALPSP to estimate the size of the problem of storing digital materials. This resulted in a very clear study from Electronic Publishing Services, which fed into the case for the extension of legal deposit. We are also talking to several individual publishers to ensure that their specific needs for a digital archive and

access to it, can be accommodated within the more general framework that we are developing. Our plans now for digital preservation archiving are being built around the assumption that we will be successful in securing legal deposit. Scaling up for legal deposit arrangements will be a huge archiving and storage challenge.

Digital infrastructure development

Building the infrastructure on a scaleable basis to support archiving and preservation work is now a vital step for us. The activities outlined in our digital object management (DOM) strategy need to be supported with the right technical infrastructure, and we are currently developing a digital store using the Ex Libris product for organisation and management. The current programme for DOM includes a range of activities: strategy definition; a research study of best current solutions; migration of voluntary deposit materials into our existing toolkit; the development and scaling up of a legal deposit secure network between the legal deposit libraries; finally, the legal deposit itself of electronic publications, as and when they come on stream. We have also done considerable work on metadata definition so that machine-to-machine search and retrieval techniques will operate correctly. We are now defining interfaces, developing internal workflows and identifying a range of operational considerations.

Another major project which we are running concurrently with DOM is the implementation of a new integrated library system (ILS). We are putting in multiple facets of this system, including major data conversion programmes, work flow development and staff training this year, in order to go fully live at the beginning of 2004. Software for the ILS is also from Ex Libris. All metadata relating to digital object management, except the actual technical data, will be stored in the ILS to ensure consistent delivery of digital and analogue assets. A system of unique identifiers will be defined, linking metadata records to digital items and objects in the digital preservation store. As

well as an enormous amount of work internally, we are working with other organisations on international standards and the development of best practice.

Collaboration and new relationships

This brings me to my final points. It is clear that we are on the cusp of major developments and unprecedented change. In different ways we are all dipping our toes into these uncertain waters. From my perspective it is now, more than ever, important for us to work collaboratively and internationally across traditional boundaries of libraries, publishers, technical suppliers and commercial agencies, to shape this future together. I think we have made good progress over the last few years to build understanding between these groups and to develop working relationships. At the British Library, we are very keen to build on this and to continue to collaborate. In this context, I think the recent report from the Research Support Libraries Group is very important in the UK context, pointing to further and deeper levels of collaboration. Particularly heartening for me, and I quote, is that: "at the heart of our considerations lies co-operation and a vital aspect has been the involvement of the British Library, alongside universities".

Certainly at the BL we are keen to engage, not just with universities, but with publishers, technologists, intermediaries and other parties, seeking new business models and common standards for inter-operability. We cannot expect to get the perfect solution first time round, nor will there be a single solution, but we stand a better chance of success if we work together.

Lynne Brindley,
Chief Executive,
The British Library
London.
E-mail: chief-executive@bl.uk