

The British Library Newspaper Collection: Long Term Storage, Preservation and Access

Patrick Fleming

Head of Reader and Reference Services for the British Library,
96 Euston Road, London, NW1 2DB, United Kingdom,
patrick.fleming@bl.uk

Phil Spence

Director, Operations & Services, British Library,
Boston Spa, Wetherby, West Yorkshire LS23 7BQ,
phil.spence@bl.uk

Abstract

The British Library has recently unveiled plans for improving the collection care of, and digital access to, one of the world's greatest newspaper collections. The collection, which dates from the 16th century, contains UK titles as well as 200 overseas titles and is at risk. 15% of the collection is unusable because of its deterioration, and there are pressures on space and a need to develop greater digital content for access.

The solution is to place all of the physical newspapers into a new, state of the art Newspaper Storage Building at Boston Spa, the British Library's main storage and logistics site, and home to its document delivery service. The building will comprise of 160 km of high density shelf space, fully automated under special environmental conditions including low oxygen, regulated 16 degree Celsius temperature, and 52% humidity.

Future access to the collection will be at the BL's main Library site at St Pancras, in a specially equipped reading room, so that researchers can integrate other collection items into their studies. However, access will be by surrogate means, initially microfilm, stored in London, and then, increasingly, digital access. By March 2009, the British Library will have digitised 3 million historic newspapers with JISC funds, and this content will be available to readers for free in reading rooms and on a paid for basis online, or via subscriptions. The Library is investigating a similar model going forward with a private sector partner who is willing to digitise and sell online out of

copyright newspapers from the BL's collection for a fixed period. This will provide digital content for customers in reading rooms, free at the point of access.

Key Words: digitisation; British Library; newspaper collection

1. The British Library Newspaper Collections: Facts and Figures

The newspaper collection dates from the 16th century and is made up of newspapers, serials and periodicals. It contains over 53,198 separate print titles and 370,000 reels of microfilm on nearly 50 km of shelf space, spread across the two buildings at Colindale in North London, 10 km from the St Pancras main reading rooms.

Through legal deposit the Library collects an edition of most newspapers published in the UK and Ireland. This accounts for 95% of the 1,259 titles currently available. These comprise:

- 111 national and regional daily titles
- 17 national and regional Sunday titles
- 511 paid-for weekly titles
- 637 free weekly titles.

The Library also collects more than 200 overseas titles. The total of UK, Irish and overseas titles amounted to 167,401 issues in 2006/07. This equates to more than 750 million pages. The collection grows annually by seven million pages.

British people are among the most avid newspaper readers in the world. 84% of all British adults (41 million people) read a regional newspaper, compared with 65% who read a national newspaper. Since 2000 total readership has increased by 1.5 million readers. Regional press has a high readership; 28% of those who read a regional newspaper do not read a national. Readership of weekly paid-for titles alone has grown by 14% since 1994 (Newspaper Society Research, 2007).

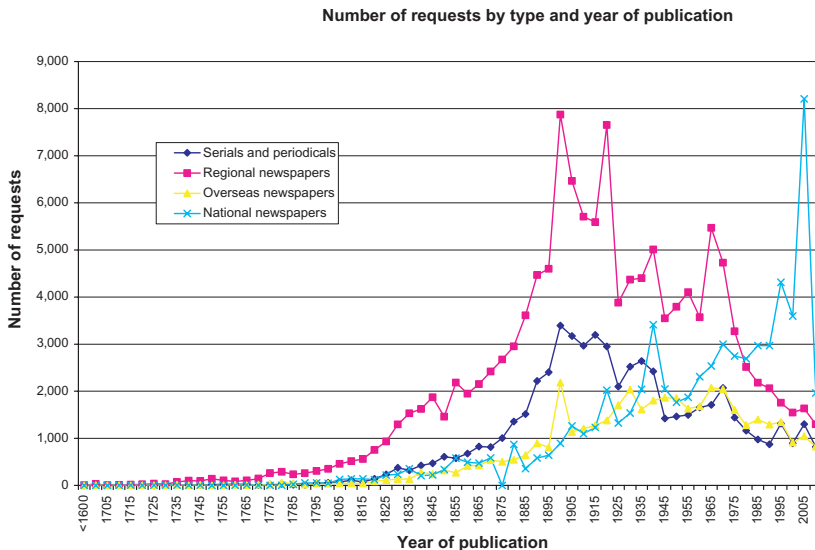
The Newspaper Library attracts an average of 30,000 reader visits per annum and is able to satisfy 100% of all reader requests based on items identified in

the Integrated Library System (ILS). Readers come from diverse communities, all of whom have an interest in the rich content of the collection. Popularity of the newspaper collection continues to grow, particularly with genealogists.

The trend is for publishers to create more news pages and to enhance their products through value-added supplements and colour magazines. The current (2007) average pagination for a British newspaper is 64 pages. The Library ingests more than six million pages of UK and Irish newspapers per annum. This figure excludes value-added supplements and magazines which are included in the products and are also collected, microfilmed and stored at Colindale. It is difficult to project future growth rates given increases in pagination, but the current trend continues to be upward, putting ongoing pressure on storage.

Figure 1 shows current researcher usage of The British Library's collection. The majority of requests were for material which was published post-1850. The number of requests increases significantly for material published after 1900. This material is of great interest to academic historians and family researchers. Therefore, if the Library concentrated on surrogate production for material post-1900 the highest number of requests could be met.

Fig. 1: Reader usage profile.



Management of the Colindale sites centres on two core functions: access and storage.

- Access is the method by which readers access the collection, either by the provision of hard copy or surrogate. Current demand is met by providing hard copy and microfilm access to the collection at Colindale. There is also a range of eProducts: either licenses to Newspaper Publisher sites, or the BL's Historic Newspapers Online service.
- Storage is the provision of the best archival conditions for the long-term preservation of hard copy newspapers, serials, periodicals, microfilm and digital surrogates.

2. The Rationale for Change

There is a compelling strategic need for change, to both protect and preserve the collection for future generations and to continue to meet the ever changing digital demands of the current and future researcher.

The key drivers for immediate change are:

Insufficient storage

The smaller building, 120 Colindale Avenue, is full. The existing lease expires finally in December 2010. The Newspaper Library at 130 Colindale Avenue will be full by 2012. There is a critical need for a new long-term storage facility.

Poor storage conditions

Newspapers need to be stored in a temperature and humidity controlled environment to ensure long-term preservation. The Newspaper Library fails to reach the minimum archival conditions set by British Standards (BS5454: 2000). Our collection care experts have prioritised newspapers as having the most urgent need for preservation compared with any other part of the Library's vast collection.

Deterioration

15% of the existing hard copy collection is 'unusable'. A further 19% is of 'high risk'. 'Unusable' means the preservation and conservation needs are so high that the material can not be given out to readers. 'High risk' is categorised as material which if given out without proper care and supervision, runs the risk of becoming unusable (BL Stock Condition Survey, 2004). The collection will continue to deteriorate if kept in the existing conditions. Continued handling of the hard copies contributes to deterioration.

Fragmented reader experience

The reader experience at Colindale is inferior to St Pancras. Research shows that readers require the Colindale collection to be integrated into the main library at St Pancras so that they can research simultaneous items in one collection.

Inefficient business processes

The existing legal deposit and storage functions provided at 120 Colindale Avenue are in a separate building to the Newspaper Library. Effective business processing would bring legal deposit and the NSB closer together on one site. The existing resource discovery, cataloguing of titles, administration and financial management of other Library collections is all provided mainly at Boston Spa.

Collection security

The provision of collection items to customers at St Pancras is underpinned by an electronic system which tracks and monitors usage and users. Colindale relies on a paper-based system. Short-term measures to provide electronic data capture are currently being implemented. The level of site security provided at Boston Spa is higher than that provided at either of the Colindale sites. To deliver the programme, the Library will have to change the current business model.

Fig. 2: The existing Reading Room at Colindale.



Fig. 3: Colindale. The main building is shown at the top, 120 Colindale at the bottom of the photograph.



Fig. 4: An example of existing storage conditions.



The changing needs of the researcher

The majority of titles currently produced by the UK and Irish newspaper publishing industry are now created in digital format and printed in full colour. Today's researcher wishes to access rich digital content in a structured manner that allows them to move from the Library catalogue to the particular item of interest as quickly as possible. Hard copy and microfilm, whilst being excellent resources, do not enable such flexibility.

The Library and JISC-commissioned report *Information Behaviour of the Researcher of the Future* (Information Behaviour, 2008) outlines the enormous changes taking place in the information landscape that are transforming teaching and learning, scholarly communication and the role of 'traditional' research library services. Many of these changes have been brought about by technology and the explosion of electronic 'content' made possible by electronic publishing, mass digitisation projects, and the internet. Library users have rapidly become information consumers who can switch instantly between commer-

cial search engines, social networking sites, wikis, bookmarked resources and electronic services provided by their library to satisfy their information needs. Library users demand 24/7 access, instant access at a click, and are increasingly looking for a particular format: a newspaper, a research monograph or a journal article, for instance, so they can scan, flick and 'power browse' their way through digital content, developing new forms of online information.

The Library has adopted this digital trend and has seized many of the opportunities new technology offers to inspire our users to learn, discover and innovate. Turning the Pages 2.0 and the mass digitisation project to digitise 25 million of pages of 19th-century English literature are examples of the pioneering work. The Library is a trusted and independent source, both in cyberspace and through its vast printed collections, with more than 67 million hits on the BL website in the past 12 months and 500,000 readers passing through our doors every year.

The Library has a strategic need and desire to bring the newspaper collection into the digital world as well as to preserve it as part of our legal mandate.

3. A Vision for Long-Term Storage, Preservation and Access

In February 2007 the British Library Board established the following vision:

'The Library will offer an integrated newspaper service based on digital surrogates at St Pancras with hard copy stored to help preserve it for future generations.'

To meet the future demands of the reader and satisfy the preservation demands for protecting hard-copy newspapers, the British Library must change the existing storage and access model in the short to mid term.

To achieve this, the strategic aim is to:

- build a dedicated new NSB at Boston Spa to provide 25 years ongoing storage for historic and contemporaneous hardcopy newspapers
- create a facility to enable the ingest of legal deposit and non legal deposit newspapers

- provide the digital infrastructure to enable the collection of both digitised and born digital newspapers. Digitised newspapers will be from either existing hard copy or microfilmed newspapers. Born digital newspapers are those provided by publishers and vendors as electronic digital files
- move to digital as a preservation medium when reliable storage standards come into existence
- integrate the newspaper, serials and periodicals with the rest of the British Library collection, with low access serials and periodicals going to Boston Spa and high usage serials and periodicals going to St Pancras
- house all historic and contemporaneous microfilm surrogates together with access to future digital surrogates at St Pancras (apart from master negative microfilm, which is stored in cold storage at a separate location for preservation and contingency purposes)
- create a newspaper reading room at St Pancras to provide long-term access
- build a new newspaper storage building at Boston Spa, to be available in 2013
- relocate legal deposit ingest from Colindale to Boston Spa
- create a digital production facility for content ingest at Boston Spa and to prepare newspapers for direct digital delivery onto the Library's DLS in 2009
- disaggregate the collection to separate newspapers from serials and periodicals
- store high-use serials and periodicals at St Pancras
- store low-use serials and periodicals at Boston Spa
- create a newspaper reading room at St Pancras
- move all access microfilm assets for storage and access in St Pancras
- create future access to the newspaper collection through microfilm and digital surrogacy at St Pancras
- create temporary access to hard copy newspapers in the reading room at Boston Spa until 80% of reader requests for access to the newspaper collection can be reached through surrogacy at St Pancras
- this will be achieved within 5 years of opening the NSB in 2012, i.e. by 2017
- relocation of staff at Colindale to St Pancras or Boston Spa between 2009 and 2012 depending upon the transition of functions and/or the availability of suitable alternative employment.

4. Future Storage and Preservation

Specification of the Newspaper Storage Building

The Library needs to obtain the most dense storage possible consistent with equal levels of accessibility for all stored collection items. As the demand rate is low compared to books, the speed of available handling systems can be used to perform complex moves to retrieve items from dense storage while using handling technology that is much simpler than that employed in the mainly book store that is currently being built.

The basis of the store will be two sizes of totes that will both accept the different sizes of newspaper and form palletised stacks of consistent unit loads. The greater part of the facility will be occupied by a double-deep automated crane bulk store, containing pallets of totes. These pallets will not leave the building. A robotic device, capable of de-stacking and restacking totes, will receive pallet loads of totes from the crane system, dismantle them to extract those required by the operator stations, rebuild loads with those returned by the operator stations and send the rebuilt pallet loads back to the bulk store. Alongside this will be a small, agile, miniload buffer store that will handle only individual totes of a smaller specification. This will be capable of absorbing any fluctuations in the workload, holding totes that have a known short waiting time and allowing the build-up of totes of fast accumulating periodicals that would otherwise place undue strain on the system by frequent pallet demand.

Conveyor systems will move pallets or totes between these components, which will all exist in a sealed void containing a low oxygen atmosphere. Totes extracted from the system will be conveyed through airlocks to operator stations and similarly returned to store after items have been removed. The design also includes provision for the robot to send larger type 2 totes directly to two dedicated large tote workstations. Returns from this area will be rebuilt into pallets by the robot and returned directly to bulk store. A separate area is provided for goods in, with loading docks and air-locked conveyor spurs providing full pallet loading, a problem resolution point for pallets rejected by the profile checker and an empty pallet store to feed the system during growth.

Where possible, the Newspaper Building will draw on the experiences of the British Library's Additional Storage Building (ASB) which is nearing completion and is shown in Figures 5 and 6.

Fig. 5: External view of the Additional Storage Building.



The environmental conditions for the internal storage void will be $16^{\circ}\text{C} \pm 1^{\circ}\text{C}$ and $52\% \pm 5\%$. This retains compliance with BS5454: 2000. This specification is designed for material for long-term retention.

Fire prevention and suppression systems in the facility will use low oxygen levels which have the benefit of being preventative rather than suppressive and for enabling a significantly smaller building, as space is not required for sprinkler systems.

Moves Methodology

For planning the moves of items and the amount of storage required, the following criteria have been modelled:

Fig. 6: Racking in the Additional Storage Building.



- growth rates;
- staffing levels;
- methodology — i.e. Library staff to undertake planning and preparation work, contractor staff to physically move the material.

The Digital Library System (DLS) for Preservation and Access

The DLS is the Library's strategic response to the long-term transition from acquiring print to acquiring digital information sources. The system has been created to design, plan, develop and implement a secure, reliable and scalable management system for the Library's digital content assets. The resulting DLS will provide a single location to ingest, store, preserve, manage, discover and provide controlled access to digital content assets. This approach will ultimately eliminate the 'technology islands' and create a single environment to manage our digital items.

A management system for digital objects will be established that will:

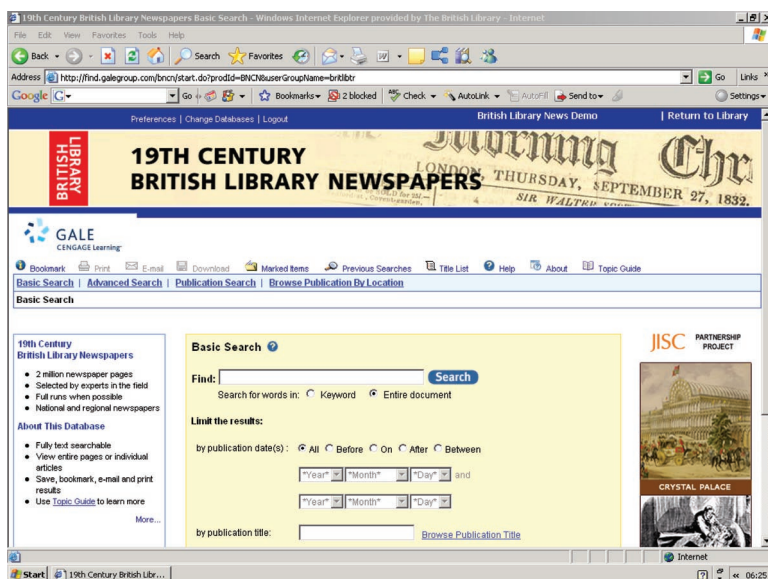
- store and preserve any type of digital material in perpetuity
- provide access to this material to users with appropriate permissions
- ensure that the material is easy to find
- ensure that users can view the material with contemporary applications
- ensure that users can, where possible, experience material with the original look-and-feel.

At the present time the DLS holds eBooks, and newspaper content will be stored in the system during 2008 and 2009, both from born digital files from volunteer publishers and digitised files taken from microfilm and hard copy out of copyright content.

5. Creating a Sustainable Digital Access Newspaper Strategy

The Library already has a number of digital projects underway and has considerable experience in creating digital newspaper sites. In October 2007 the Library launched a beta website as part of a partnership programme with JISC to digitise two million pages of historic newspapers to produce a fully searchable resource. This is an exemplar programme for the Library. The website contains a number of London and English regional titles including *The Morning Chronicle*, *The Examiner* and *Illustrated Police News* and Chartism newspapers. Regional

Fig. 7: Screenshot of the British library's Historic Newspaper Service.



newspapers include: *The Northern Echo*, the *Western Mail*, *The Newcastle Courant*, *The Ipswich Journal*, *The Bristol Mercury*. The site is illustrated in Figure 7.

Also digitised is the Burney Collection. This outstanding newspaper collection begins with what are actually parliamentary papers from 1603; the earliest items which are recognisable as newspapers date from the early 1620s. It is in 18th-century London newspapers, however, that the collection is particularly rich. All the major titles are included, such as the *Daily Courant* from 1702 to 1735, the first daily newspaper published in London, and the *London Gazette* from 1665. Periodicals are also included, such as Addison and Steele's famous *Tatler* (1709–1711) and their *Spectator* (1711–1712).

Also represented are English provincial titles from 1712, such as the *Stamford Mercury* of 1728, Irish newspapers (the earliest being the *Dublin Intelligence* of 1691), Scottish titles from 1708 onwards, and many 18th-century American titles, including the *New England Courant* (1721–1723), on which Benjamin Franklin (1706–1790) worked before he moved from Boston to Rhode Island after having published the paper's final issue on 25 June 1726.

Digital technology has enabled newspaper publishers to transform production processes. All national and regional daily newspapers and an ever growing number of weekly newspapers are produced digitally with an output in PDF format. Within five years, the entire UK and Irish newspaper publishing industry will be produced using digital technology. This, together with the growth in colour presses, has enabled newspapers to increase the number of pages and products that they offer.

If the lifetime of storage for the newspaper collection is to be extended, ways have to be found to migrate away from the collection of space-hungry, hard-copy newspapers. The newspaper industry already believes that digital is an acceptable long-term storage and preservation medium.

It is likely that by 2014 the Library will end microfilming as a process and will utilise digital technology as a one-stop-shop for digital preservation, storage and access. To maximise access, the Library must look at having the existing historic collection together with digital content.

Four potential routes are available for the Library in acquiring an enhanced collection of digital surrogates:

- automatic ingest of PDF format copies of contemporary newspapers direct from publishers, initially by voluntary agreements and then by Legal Deposit legislation;
- systematic digitisation of historic newspapers, in partnership with publishers and content aggregators, both in-copyright and out-of-copyright;
- digitisation-on-demand of other material not accessible via hard copy or microfilm;
- purchase existing digital licensed collections. The Library already does this, e.g. with The Times historic archive, The Guardian historic archive and The Daily Mirror historic archive.

Automatic ingest of PDF copies

To achieve this, the Library would work with publishers to create mutually beneficial strategic partnerships to enable the ingest of all editions of published newspapers via PDF and to make them available on a page level, read-only basis in the reading rooms. Fully searchable content would be available via subscriptions or micropayments.

Informal discussions have taken place with a number of publishers. A leading regional publisher, for example, is exploring the possibility of working with the Library to capture several million pages of its historic newspapers and to ingest ongoing regional titles directly into the Library for access in reading rooms.

Until legal deposit legislation is extended to cover digital ingest of newspapers, voluntary ingest is the logical step forward. Some countries such as Norway have legislation in place, and others, such as France, have started voluntary deposit pilots.

Newspaper publishers generally see no competition in providing a digital version for page level, read-only access in the reading rooms or in entering into a mutually beneficial online revenue sharing partnership by which readers would access their searchable and purchasable online databases via an e-commerce platform.

Digitisation of historic newspapers

Content aggregators are also seeking agreements with other publishers to digitise their archive content and to provide access through their discovery platforms. The potential exists for the BL to develop agreements with content aggregators to operate appropriate access and resource discovery platforms to digitise the Library archive where there is a commercial business benefit for them in doing so. This will enable the 'long tail' of 750 million newspapers to be gradually digitised over many years.

For the JISC/Burney newspaper digitisation projects, Cengage Gale has provided its infrastructure development, service delivery platform and ongoing product development in return for a licence to commercialise the content to markets outside UK higher and further education sectors; the Library receives a royalty payment from subscriptions to the Cengage Gale service. The Library could extend this type of model to other historical newspapers. Where the material is out of copyright, there may be the opportunity for the Library to earn a royalty from subscriptions. This would help cover costs of further digitisation of the collection for the nation. Where the material is in-copyright, royalties would be negotiated with copyright holders; but provision exists to support publishers by helping them develop unified archives of their collections.

The defining principles for the British Library are that we do not wish to damage newspaper publishers' interests, but rather wish to work with them to provide access to digital newspapers in reading rooms and develop a significant digital archive of newspapers.

Next Steps

The next steps are to continue detailed planning and secure further funding for the programme.

References

BL Stock Condition Survey, 2004, Library Collection Care.

Information Behaviour of the Researcher of the Future, Centre for Information Behaviour and the Evaluation of Research (CIBER) at UCL, January 2008.

Newspaper Society Research (2007): GB TGI 2007 Q3 (April 2006–March 2007).