

# Content, collaboration and aggregated connectivity

**PATRICK FLEMING**

**Head of Operations and Services, London**

# Our 2020 vision

http://www.bl.uk/aboutus/stratpolprog/2020vision/2020A3.pdf - Windows Internet Explorer provided by The British Library - Inter

http://www.bl.uk/aboutus/stratpolprog/2020vision/2020A3.pdf

File Edit Go To Favorites Help

http://www.bl.uk/aboutus/stratpolprog/2020vision/2...

Page Tools

## Our values

Our values represent the aspirations of our individual and collective behaviour. They guide our day-to-day decision making. In order to achieve the vision we have set out for 2020, we will endeavour to put these values at the heart of everything we do.

**We inspire** We aim to inspire people around the world through our content and expertise.

We are passionate about sharing information and are inspired by our users and their contributions.

**We innovate** We nurture leading-edge innovations and promote an entrepreneurial spirit.

We take measured risk and bring about change through informed and agile decision making.

**We collaborate** We work in partnership to achieve more by working together than we could do on our own.

We strive to be an open, responsive and reflective organisation.

**We enable** We aim to work with our users to create new insights from our content and to develop world-leading services.

We strive to encourage and empower our staff to seize new opportunities.

**We respect** We promote respect for our users and stakeholders and for one another.

We care for the collection and respect legal and institutional frameworks.

**We learn** We will continue to learn from others. We integrate the expertise of our users into our networks.

We work to share our ideas and expertise.

**\* FORECAST**  
The online landscape will increasingly resemble the 'semantic web' – in which computers become capable of extracting, classifying, categorising, and analysing data to create context and new uses for content.



## Our mission and vision

**Our mission:** Advancing the world's knowledge.

**Our vision:** In 2020 the British Library will be a leading hub in the global information network, advancing knowledge through our collections, expertise and partnerships, for the benefit of the economy and society and the enrichment of cultural life.

**Our vision is supported by five key themes which set out the strategic priorities for the Library:**

- 1 Guarantee access for future generations.
- 2 Enable access to everyone who wants to do research.
- 3 Support research communities in key areas for social and economic benefit.
- 4 Enrich the cultural life of the nation.
- 5 Lead and collaborate in growing the world's knowledge base.

Previous page (1) Dame Lynne Brindley, Chief Executive, British Library  
Left (2) Image from our Growing Knowledge exhibition

3

Unknown Zone

start 7 Internet ... 3 Microsoft ... 5 Microsoft ... C:\Documen... Microsoft Po... Broadcast N... 15:54

## Our 2020 vision

**In 2020, the British Library will be a leading hub in the global information network, advancing knowledge through our collections, expertise and partnerships, for the benefit of the economy and society and the enrichment of cultural life**

- For newspapers we see a world of fully searchable aggregated multimedia resources, a mix of free and pay to access content
- The Library's role is to open up resources and point researchers to where they can find them around the world

# The Newspaper Collection Today



# The British Library Newspaper Collection

- Collection dates from 16th Century
- 53,198 separate print titles
- 270,000 volumes
- 370,000 reels of microfilm
- 50 km of shelf space
- Legal deposit collects 95% of 1400 titles from UK and Ireland
- 200 overseas titles
- 800 million pages
- 30,000 readers per year

# The British Library Newspaper Collection







# A stark contrast



# The Kings Library



# The British Library reading rooms



# Reasons for change

- British Library Act 1972 requirement to store legal deposit content for the nation
- 1930s building with poor storage conditions
- 15% of the collection is 'unusable'
- 19% of the collection is 'at risk'
- The building is full
- Newspaper Library is not located with other collection items, at St Pancras reading rooms
- Basic catalogue and indexes

# The answer is digital

In February 2007 the British Library Board reconfirmed 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 deliver the vision the Library had to persuade the Government to invest £33m to fund a new newspaper storage building at Boston Spa, Yorkshire to store all hard copy newspapers

Access would then revert to St Pancras through surrogates microfilm and digital with hard copy newspapers only being accessible in very controlled circumstances

Some previous digitisation experience: The Burney collection of 1800 to 1900 content and three million pages of 19<sup>th</sup> century local newspapers

But just a drop in the ocean against a newspaper collection of 800 million pages ...and growing every year by 300 metres

## The Burney Collection

- The largest single online collection of English news media from the 17th and 18th centuries
- The Burney Collection offers unique insights into two centuries of history through access to over 1,270 newsbooks, newspapers, pamphlets and a variety of other news materials published in England, Ireland and Scotland, plus papers from British colonies in Asia and the Americas
- Digitised through a partnership between the National Science Foundation and the Library the content was developed and hosted online by one of our partners Cengage Learning.

# British newspapers 1620-1900

- With a grant of £3m from JISC and online support again from Cengage Learning we digitised three million pages of complete runs of 19th century local, regional and national British newspapers
- Titles were selected to reflect both the full chronological span of the era and the geographical spread of the UK
- Contains major historical events from the Battle of Waterloo, the Great Exhibition of 1851 to the opening of the Suez Canal, factual reporting of the Battle of Trafalgar in the *Examiner* and the gory details of the Whitechapel murders in the melodramatic *Illustrated Police News*
- Many famous authors of the 19th century period wrote for newspapers, and their work is represented, Dickens and Thackeray are included
- The corpus of Burney and the two JISC projects led to a step change within scholarly communities, vastly expanding the range and variety of what can be used in research, at all levels of education.
- But at just over three million pages this was a drop in the ocean of more than 800 million pages of hard copy. We needed a bigger, more dramatic solution.

# A mad idea!

**Find a partner who would invest many millions of pounds to digitise many millions of pages at no cost to the Library**

**The Library decided that a public procurement under a new process called Competitive Dialogue would be the best solution.**

- **This had to be a process which was:**
- **Transparent**
- **Unbiased and giving best value**
- **Enable market and technical knowledge to be gained from the suppliers – calling on as wide a pool of potential providers as possible**
- **Commercial issues were undefined and the Library was seeking a supplier to take on the project at their own financial risk**

## Our desire

- **Within the Library's overall Newspaper Programme the digitisation project was required to meet the need for bulk digitisation and:**
- **Grow the use of surrogates (to meet a KPI for the Library of 80% of reader requests from the newspaper collection to be met from a surrogate by the end of 2017)**
- **To "improve" the user experience following the completion of the Programme and the move of the physical newspapers to a new Newspaper Storage Building at Boston Spa**
- **Test the market for solutions where the supplier could be "self-funding" as experience with JISC had shown digitisation of newspapers to be expensive and time-consuming and requiring specialist skills and equipment**
- **Provide the Library with a clear exit strategy if it didn't work or no supplier could be found to undertake the work**

Home | British Newspaper Archive - Windows Internet Explorer provided by The British Library - Internet

http://www.britishnewspaperarchive.co.uk/

File Edit View Favorites Tools Help

Home | British Newspaper Archive

## The British NEWSPAPER Archive

Email address  Password    Remember me

[Home](#) [Search](#) [Help](#) [About](#) [Print Shop](#) [Blog](#)

[Register](#) | [Purchase Packages](#) | [Forgotten Password?](#)

The BNA will be attending a [#familyhistory](#) event at Hospitalfield House in Arbroath on Friday 7 September [#genealogy](http://t.co/KCXqPvCp)

### History at your fingertips

 [Getting started guide](#)  
[Free example pages](#)  
[Frequent questions](#)

### Gift Vouchers

New! Send to a friend.  
[Purchase a Gift Voucher](#)

Received a Gift Voucher?  
[Redeem your Gift Voucher](#)

### High Quality Prints

Purchase a high quality A1 print of any page from the archive. Framed

*One place, millions of stories*

[Tweet](#) [Like](#) 51

## Search The Archive

Learn more about Britain's history, search The British Newspaper Archive's vast treasure trove of historical newspapers.

Enter keyword(s) eg, Locomotive OR 'Florence Nightingale'  [+ Advanced](#)

### Olympic stories from The Archive

 JULY 13<sup>TH</sup> TO 25<sup>TH</sup> 1908  
IN THE GREAT STADIUM

We've found some great Olympic stories here on The British Newspaper Archive - what will you discover?

### Register today, and receive 15 FREE credits!

 **Become a Registered Member**  
Receive **15 free credits** (up to 3 page views!) to help get you started with your

 **Purchase a Package**  
Continue your journey in The British Newspaper Archive. Four great value

Done Internet 100%

start 8 Internet ... 3 Microsoft ... 3 Microsoft ... C:\Documen... British Librar... Microsoft Po... 15:03

## Key contractual points

- **brightsolid will digitise content from the Newspaper Archive to populate a commercial offering: brightsolid is responsible for the selection of material to be digitised based on their business plan**
- **brightsolid is operating entirely at their own commercial risk with no direct costs to the Library**
- **The Library receives:**
- **Copies of all scanned materials for ingest (with accompanying OCR text and metadata)**
- **Free access from the Library premises to [www.britishnewspaperarchive.co.uk](http://www.britishnewspaperarchive.co.uk)**
- **For the duration of the contract the out of copyright wall was set at 1900**
- **A royalty on all paid-for accesses to out-of-copyright materials: this revenue is designed to recoup the additional costs to the Library (for example for the provision of materials to the scanning operation)**

## The brightsolid studio at Colindale scans 8000 pages a day



## What does success look like?

- More than 100 million pages have been viewed
- brightsolid has more than 50,000 unique subscribers
- The Library is receiving royalties from brightsolid
- brightsolid is receiving significant revenues
- The project has already opened up a treasure trove of content
- It has transformed search for researchers all round the world
- It has proved that where there is a will there is a way
- It has transformed access to the newspaper collection
- It has enabled the library to create digital surrogates

## Non Print Deposit

- April 2013, legislation which will permit the Library to collect the entire UK web domain of over 10 million websites, born digital free newspapers and paid for newspapers published as digital surrogates of their print editions
- We will also be permitted to collect ipad and tablet editions of newspapers.
- But what about live radio and TV news? In the true spirit of our vision of aggregated connectivity we want to move on from just collecting newspapers to enable researchers to view and research newspapers alongside multimedia platforms.

# Television and radio news



# Cengage Learning-Nineteenth Century Collection Online

Nineteenth Century Collections Online - Home - Windows Internet Explorer provided by The British Library - Internet

http://ncco.galegroup.com/gdc/ncco?u=acd\_ncco

File Edit View Favorites Tools Help

Nineteenth Century Collections Online - Home

Nineteenth Century Collection Online Trial [Change Resources](#) [Title List](#) [Return to my library](#)

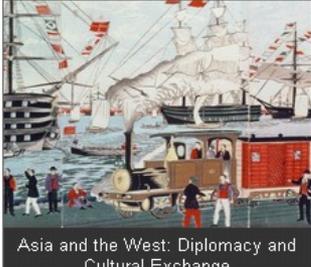
Resource Links [Sign In](#) [Saved Items \(0\)](#) [Search History](#) [English](#) [End Session](#) [Help](#)

## Nineteenth Century Collections Online

[Home](#) [Explore Collections](#) [Graphing Tool](#) [About](#)

Searching 3 of 3 Archives   [Advanced Search](#)

### Archives



Asia and the West: Diplomacy and Cultural Exchange



British Politics and Society



European Literature, 1790-1840: The Capewy Collection

http://ncco.galegroup.com/gdc/ncco

start Internet Explorer Microsoft Office O... Hardware installation 100% 09:10

# Cengage Learning-Nineteenth Century Collection Online

The screenshot shows a web browser window titled "Nineteenth Century Collections Online - View All - Windows Internet Explorer provided by The British Library - Internet". The address bar contains the URL: <http://ncco.galegroup.com/gdc/ncco/NewspapersFullListPage/NewspapersFullListWindow?failOverType=&query=KE+newsp>. The browser's menu bar includes File, Edit, View, Favorites, Tools, and Help. The page content features a header for "Nineteenth Century Collection Online Trial" with navigation links: "Change Resources", "Title List", and "Return to my library". Below this is a "Resource Links" section with buttons for "Sign In", "Saved Items (0)", "Search History", "English", "End Session", and "Help". The main heading is "Nineteenth Century Collections Online" with a search bar containing "Searching 3 of 3 Archives" and a search icon. A secondary search bar is labeled "Advanced Search". Navigation tabs include "Explore Collections", "Graphing Tool", and "About". The search results section shows "Search results: Keyword= newspapers". On the left, a "Content Types" sidebar lists "Monographs (36)", "Manuscripts (25)", and "Newspapers (2,074)". Below this is a "Search within results" field. The "Analyze" section includes "Term Clusters". The "Limit Search by" section is partially visible. The main content area displays "Newspapers" with a sub-heading "Newspapers Periodicals, Etc." and a "Displaying 1 - 20 of 2,074" indicator. The first result is "The Worker (Hertfordshire, United Kingdom), May 23, 1922, Issue 7, p.[1].", with a thumbnail of the newspaper page and a star icon. The second result is "[We Have Received Upwards of Twenty Other Newspapers Containing Notices of Our Journal, Which Must Remain to Be Noticed in Our Next] The Journal of the Working Classes (London, United Kingdom), n.d., Issue 5, p.60.", with a thumbnail and a star icon. The Windows taskbar at the bottom shows the Start button, Internet Explorer, Microsoft Office, Hardware installation, and Document1 - Microsoft... with the system clock at 09:12.

# Cengage Learning- Nineteenth Century Collection Online

Nineteenth Century Collections Online - Term Clusters - Windows Internet Explorer provided by The British Library - Internet

http://ncco.galegroup.com/gdc/ncco/analyzeSearchResults/AnalyzeSearchResultWindow?failOverType=&query=KE+newsp. Google

File Edit View Favorites Tools Help

Nineteenth Century Collections Online - Term Clusters

Explore Collections Graphing Tool About

Print Download

## Term Clusters

Categories are derived by an algorithm that creates clusters of terms based on the first 100 words of the first 100 search results per content type. By clicking on a category, you will be able to refine your search further.

**Search results:** Keyword= newspapers Content Types: Newspapers

1. [Newspapers Periodicals, Etc.](#)  
Newspapers Periodicals etc EARLY A RELIABLE SERVICE Fenner Mardell 31 Leys Avenue Newspapers Periodicals etc EARLY A RELIABLE SERVICE Fenner Mardell 31 Leys Avenue...

2. [\[We Have Received Upwards of Twenty Other Newspapers Containing Notices of Our Journal, Which Must Remain to Be Noticed in Our Next\]](#)  
We have received upwards of twenty other newspapers containing notices of our Journal which must remain to be noticed in our next We have received upwards of twenty other newspapers containing... [Newspapers]

3. [Tax upon Newspapers](#)  
Tax upon Newspapers it is said that on an average 40000 Newspapers are sent from London every day varying from twentyfive to sixty thousand If these weigh one ounce each it will require one entire mail coach... [Newspapers]

4. [Conveyance of Newspapers](#)  
CONVEYANCE OF NEWSPAPERS CONVEYANCE OF NEWSPAPERS The bill to consolidate and amend the laws relating to the conveyance of newspapers by the post enacts 1 That all newspapers sent by the general post within the... [Newspapers]

Done

start 6 Internet Explorer 7 Microsoft Office O... Hardware installation Document1 - Microsof... Internet 100% 09:14

# Next projects

- **The creation of a UK newspaper resource dataset, locally searchable via advanced text mining and semantic research tools optimised to identify new areas of investigation for digital scholars**
- **To enable experimentation and trialling of new digital research and data mining technologies with developers attempting to identify meaning or patterns in large scale textual datasets**
- **To generate statistics information packages and data visualisations derived from the full text designed to present the data in a way that inspires greater interest and further research**
- **Integration of the metadata/dataset content with external resources (e.g. Wikipedia, British Museum & National Archive linked datasets, search engines etc) in order to present search results from the data in a wider context to researchers and create new perspectives**
- **To link with internal resources such as UK mapping data, Newspaper Library catalogue and UK sound archive recordings to present an integrated UK resource for local historians, genealogical researchers etc**
- **To integrate with other BL resources in the creation of interactive content for exhibitions with a UK theme (e.g. Writing Britain)**
- **To generate metadata to link with wider British Library catalogue data in providing first reviews of authors' published works, historical background for those researching authors' biographies etc**

# ...and of course there are apps!

