

**Enhancing the Discovery of
(Digital) Collections –
The View from the UK**

David C Prosser
Executive Director
RLUK

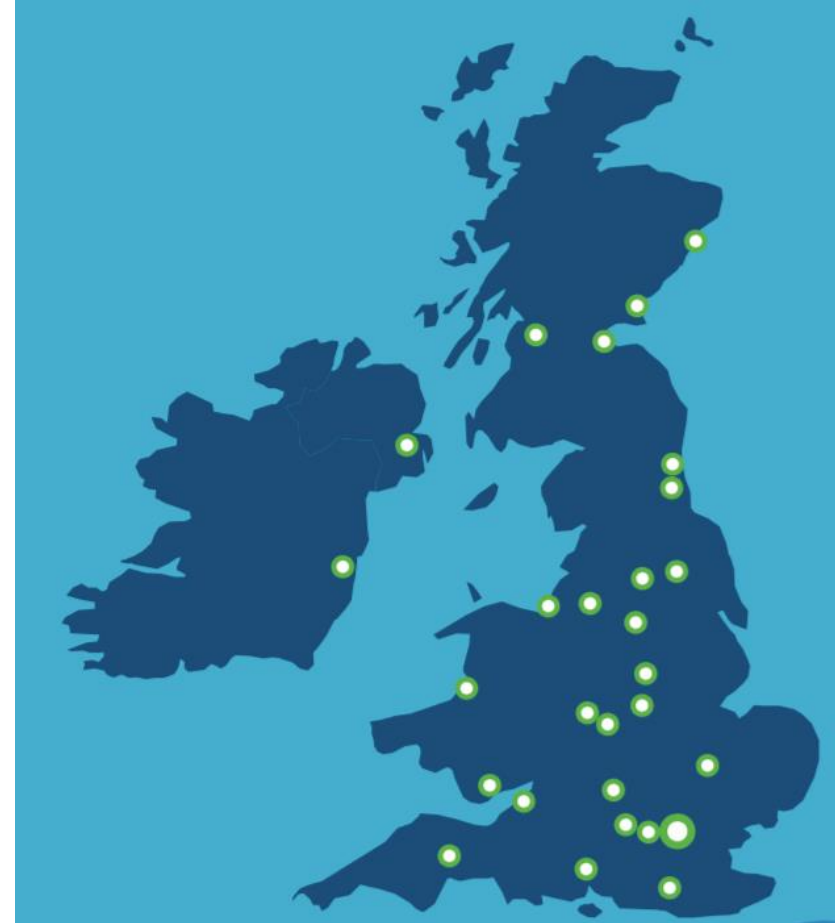
Enhancing the Discovery of Collections

- **Who We Are**
- **Hidden Collections**
- **The 21st Century's Hidden Collections**
- **Mapping Collections**
- **Sharing Collections**
- **Describing Collections**
 - **Owned / non-owned material**
 - **Walled gardens / open commons**

RLUK

- Membership organisation of 37 libraries
 - 33 university libraries
 - 3 UK national libraries
 - Wellcome
- Vision:

Our network of distinctive research libraries will be recognised as influential leaders in and valued contributors to the advancement of research and scholarship through their outstanding collections and services



Hidden Collections

HIDDEN COLLECTIONS

Report of the findings of the RLUK Retrospective Cataloguing Survey in
association with The London Library

- 2010 survey carried out by RLUK and the London Library
- 77 responses were made to the survey
- Including 38 academic, seven public and 32 specialist libraries
- Report written by Mike Mertens, RLUK, and Dunia García-Ontiveros, the London Library
- <http://www.rluk.ac.uk/strategica ctivity/strategic-strands/udc/hidden-collections-report/>



Survey Responses

- **Cathedral Libraries:** Ely, Norwich, Worcester, York Minster,...
- **Museums and Galleries:** Courtauld, Horniman, London Transport, National Gallery, National Portrait Gallery, National History, National Trust, Tate, V&A, Zoological Society of London,...
- **Independent Libraries:** Dr Williams's Library, London Library,...
- **Public Libraries:** Bristol, Exeter, Hull, Manchester
- **Society Libraries:** Oxford Union Society, Royal Asiatic Society, Royal College of Music,...
- **University Libraries:** Cambridge, Cardiff, Edinburgh, Glasgow, Lancaster, Leeds, LSE, LSHTM, Manchester, Middlesex, Oxford, Newcastle, Royal Holloway,...

Key Findings

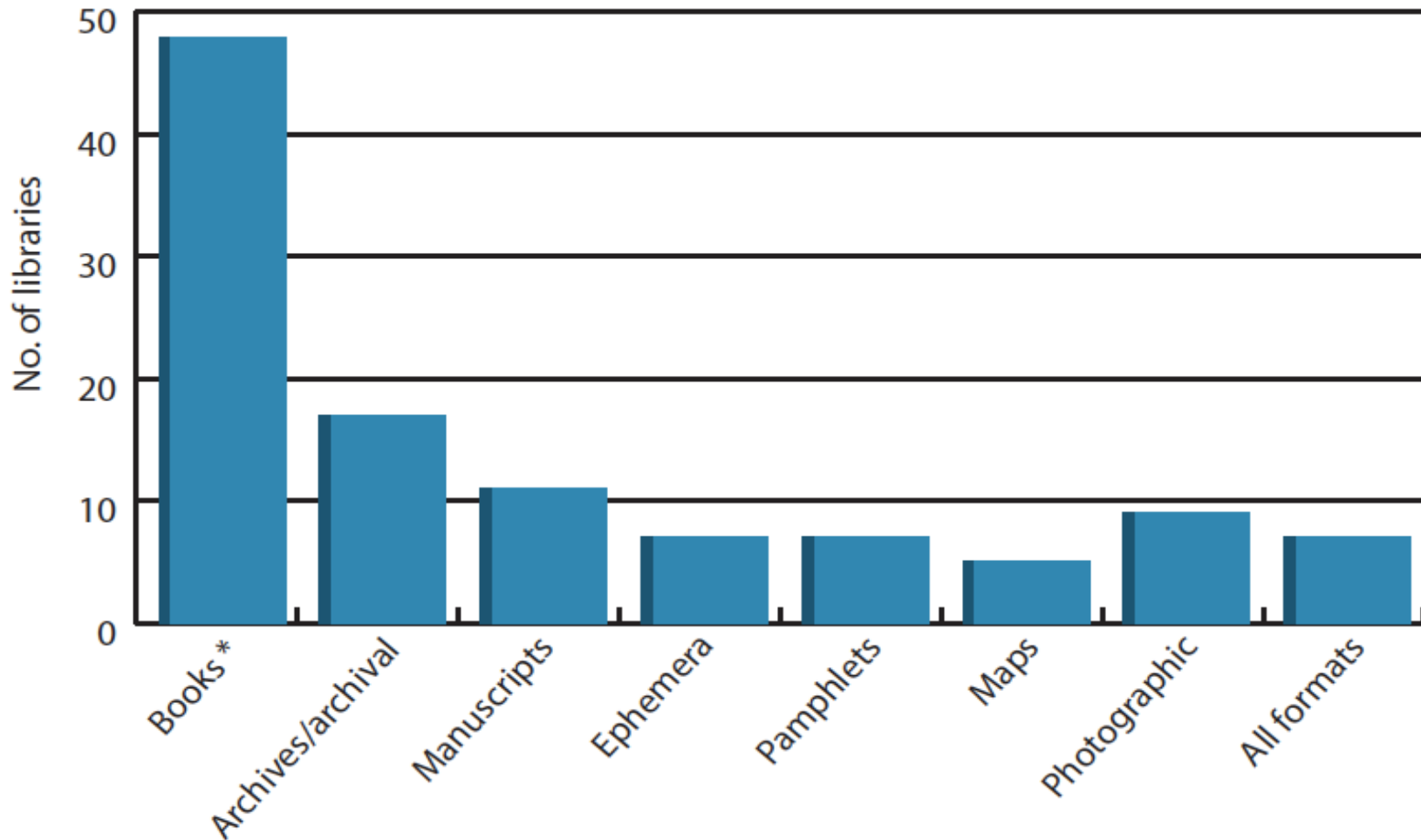
- Over 13 million volumes are uncatalogued in the libraries that responded, 18.5% of the total number of volumes held by those libraries.
- Over 4 million more (in a smaller number of libraries) have unsatisfactory catalogue records.
- Museums, public libraries and independent libraries have a higher proportion of collections which are invisible online.
- While research libraries have better coverage of printed collections, their hidden archival collections often remain vast.
- Modern material is being added to the backlogs. The presence of 21st century materials in the backlogs suggests that some libraries are unable to keep up even with current acquisitions.

Key Findings

- Foreign language material and formats which require particular skills and expertise (maps, music, archives) are heavily represented.
- There are serious problems in collating and comparing metrics for materials other than printed books.
- Librarians are aware of the problem and are actively trying to tackle the backlogs; over 60% have retrospective cataloguing projects under way. However, the scale of the problem is often beyond individual institutions.
- Respondents support an online register of retrospective cataloguing and are interested in exploring national initiatives and technical solutions to bring this about.

Formats

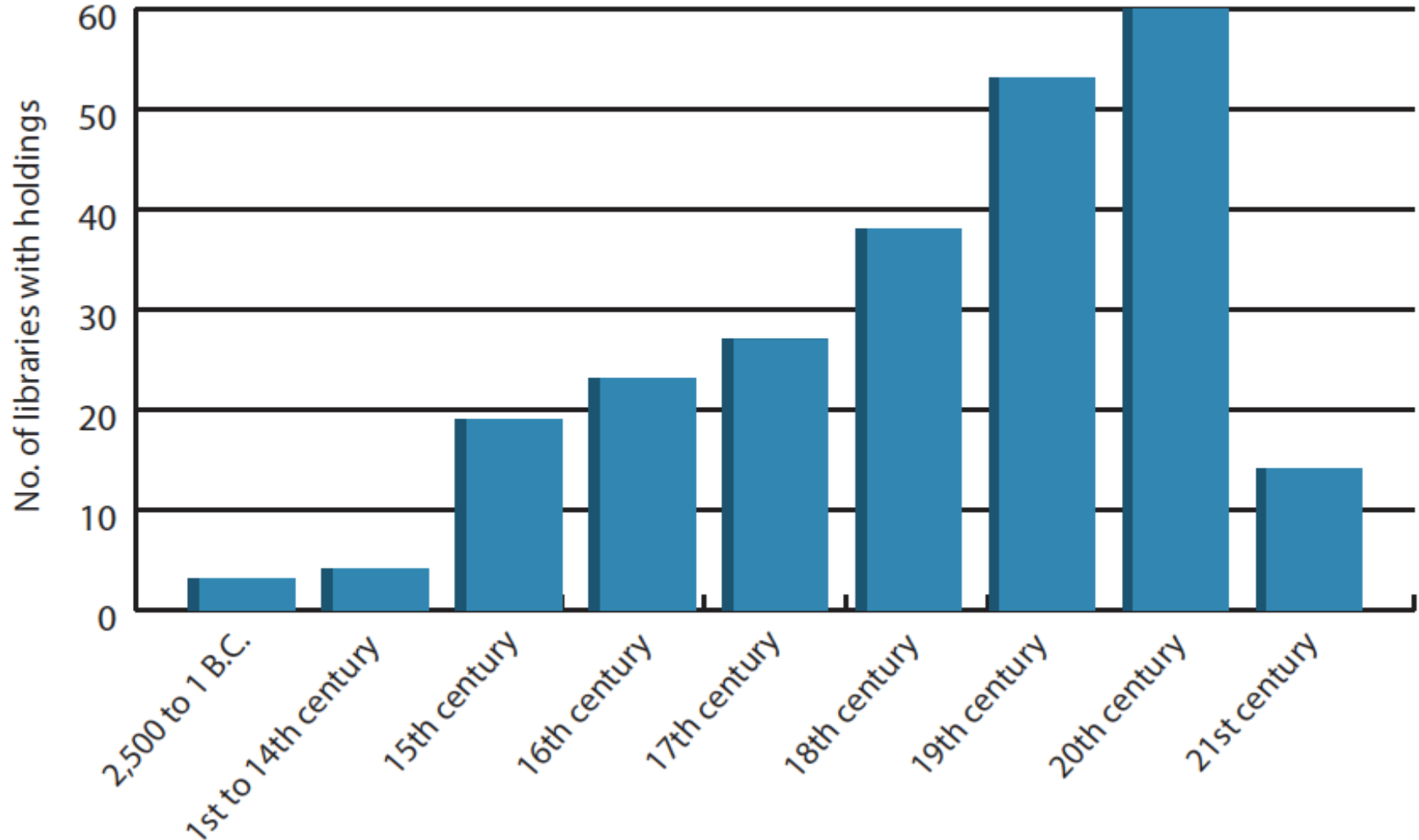
Question 22: Formats in hidden collections



*Books including "other printed material", "monographs", "serials", "periodicals", "rare books"

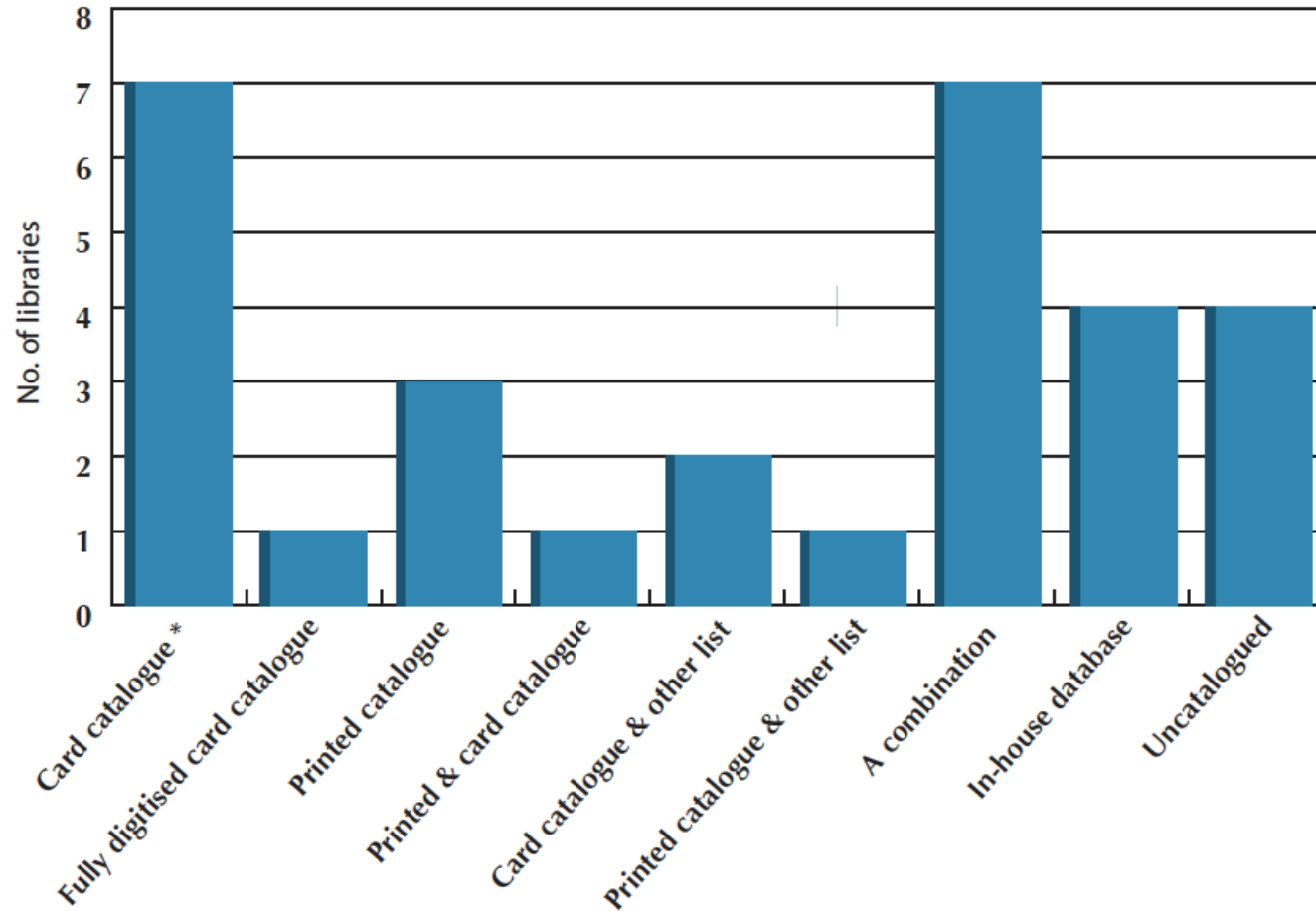
Age of Material

Question 23: Dates of hidden collections



Degrees of Hidden

Question 77: Visibility of special collections



*one of these is partially digitised.

Resourcing Discovery

- For some HE establishments, special collections are seen as unique selling points for the institution
- They are used in community engagement and teaching, not just as research resources
- Special collections can bring kudos and attention to institutions
- If we are to gain large-scale funding for exposing the hidden then we need to make a wider societal case for the collections
- This is especially true in periods of ‘austerity’

Leveraging the Unique

BBC

Sign in

News

Sport

Weather

iPlayer

TV

Radio

NEWS

Home

UK

World

Business

Politics

Tech

Science

Health

Education

Entertain

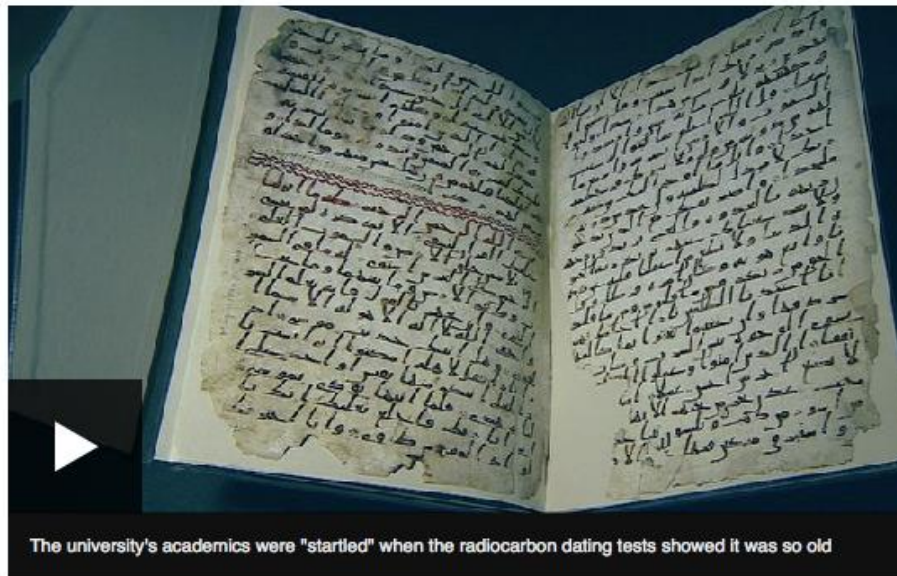
Education

School Report

'Oldest' Koran fragments found in Birmingham University

By Sean Coughlan
Education correspondent

🕒 22 July 2015 | Education & Family



Leveraging the Unique

HOME » NEWS » UK NEWS » LAW AND ORDER

Magna Carta worth £10m found in council archives

Forgotten version of historical parchment that established principle of rule of law discovered in files of history department of Kent County Council in Maidstone



3K



458



0



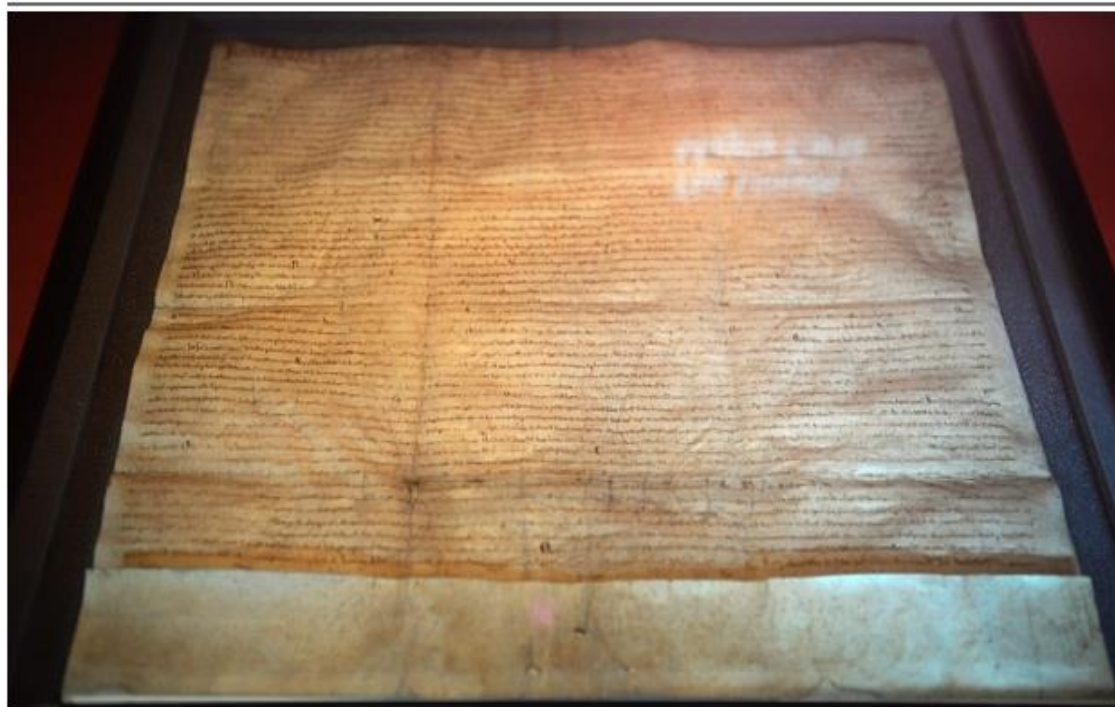
13



4K



Email



The Lincoln Magna Carta on display at The British Library in London Photo: EPA/FACUNDO ARRIZABALAGA

Featured

Convert-a-Card

Transcribe-a-Card



Indonesian Card Catalogue: Drawer Three

Transform this printed card catalogue into a searchable online database



✓ Completed: 70%

☰ Tasks Remaining: 314

👥 Volunteers: 99

🕒 Last Active: an hour ago

CONTRIBUTE

Hidden Collections - Recommendations

- Create a (self-described) national register of hidden collections. The register could have inter alia the following functions:
 - Describe the extent of uncatalogued material
 - Describe the extent of uncatalogued material in broad subject categories
 - Depict a sense of institutional priorities for cataloging and digitisation
 - Show the extent of catalogues presently not available digitally or online
 - Demonstrate the impact of retro-cataloguing/retro-conversion funding

Hidden Collections - Recommendations

- Collection holders should adopt a community licence for any bibliographic records
- Lobby for the funding of retro-cataloguing and retro-conversion of material
- Investigate how Copac can be made more easily integrated into cataloguing workflows, such as through a separate, freely-available cataloguing module targeted at smaller-scale collection holders, to help accelerate, and reduce the barriers to, the creation of online records.
- Promote the vision of the 'National Research Collection', comprising ideally collection holders from all sectors under which activity around hidden collections, collaborative collection management, and shared print could be discussed and pursued cohesively.

Hidden in the 21st Century

- Jisc/RLUK/SOCNUL 'Spotlight on the Digital' Project looked at digitised collections.
(http://repository.jisc.ac.uk/5569/1/JISC_REPORT_spotlight_on_digital_WEB.pdf)
- Out of 217 online collections surveyed, almost 20% (40) had become lost to the web and were no longer discoverable
- Of those that were still available, many had descriptions only at the collection level – individual items were not discoverable

High-Level Recommendations

- **Build institutional capacity** - Using national support to develop local institutional and practitioner capacity and capability, while recognising the value-added and fall-back roles of aggregation
- **Exercise collective watchfulness** - Engaging experts to ensure the sector is not caught out by technology changes and implementing automated checks to identify assets that are becoming undiscoverable
- **Leverage expertise at scale** - Identifying demanding lifecycle activity that can be optimally delivered and repeated at scale and therefore above-campus, such as software tool development, high profile promotion and licensing advice

Practical Recommendations

- Make Google searches work for the collections
- Use social media effectively
- Use content aggregators
- Using popular websites (Wikipedia, YouTube and Flickr) to reach broader audiences

<https://www.jisc.ac.uk/guides/make-your-digital-resources-easier-to-discover>

Figure 1: Below are five possible starting points for research in academic literature. Typically, when you are conducting academic research, which of these five starting points do you use to begin locating information for your research? *Percent of respondents who indicated that each option is the starting point for their research.*

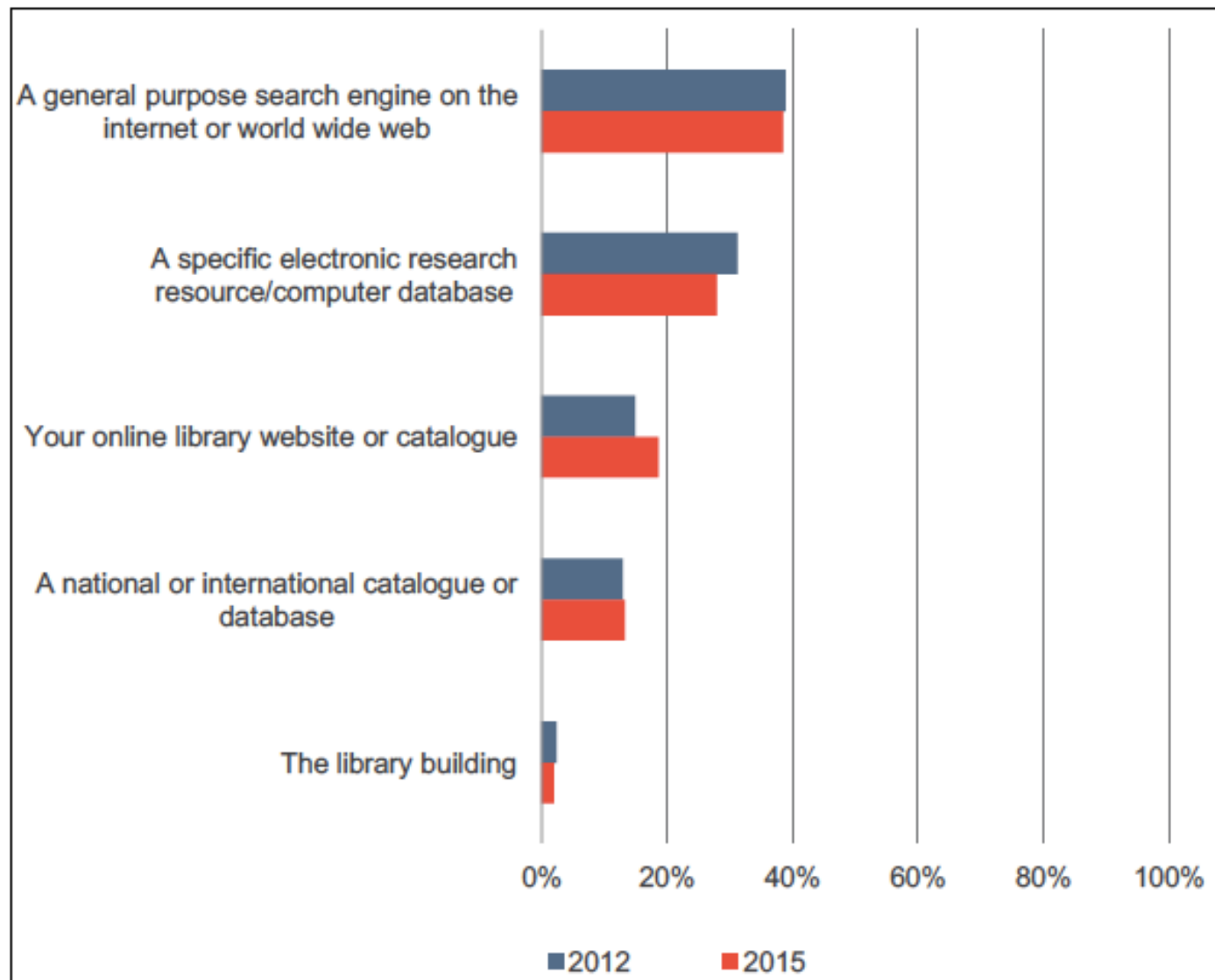
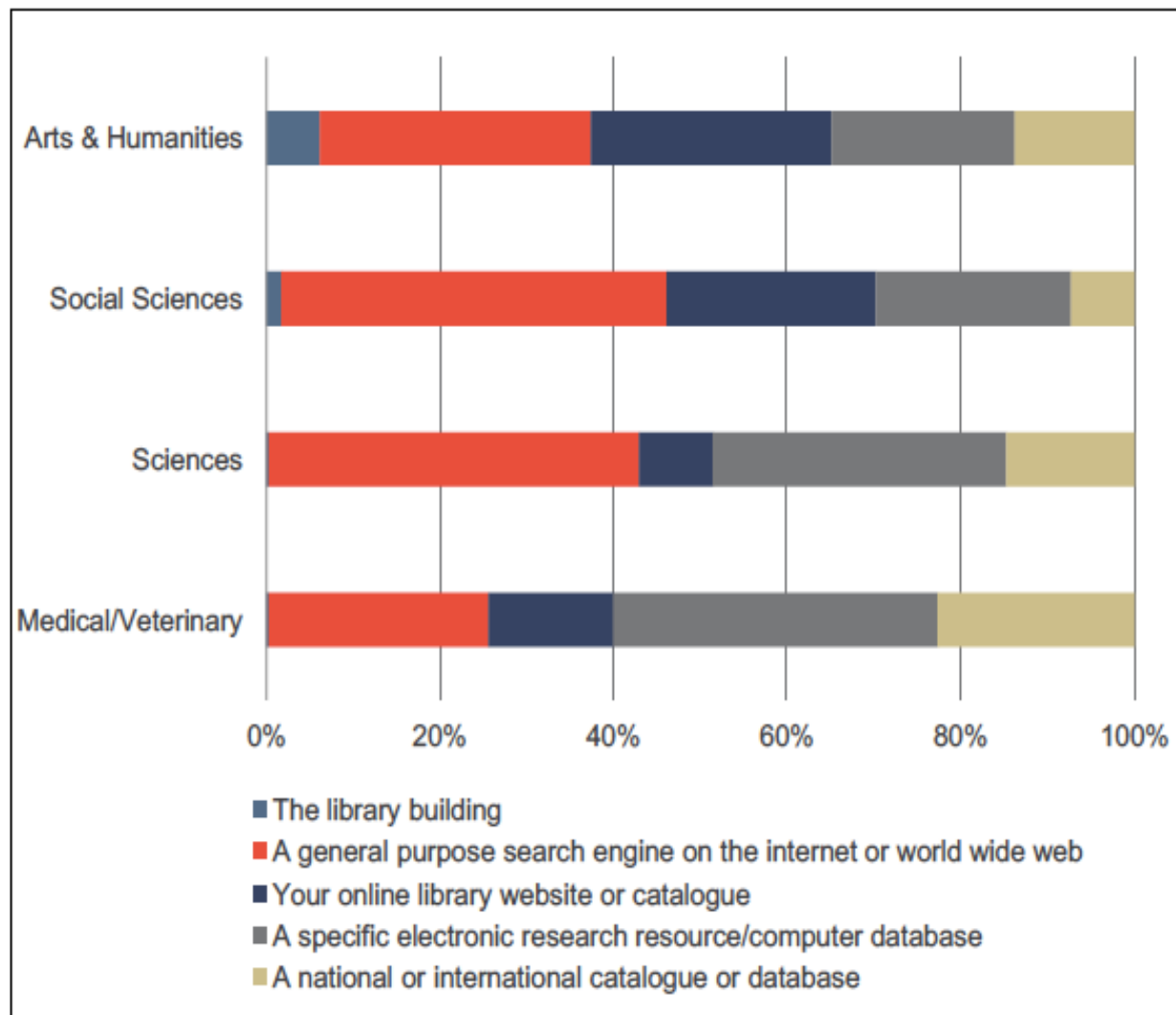


Figure 2: Below are five possible starting points for research in academic literature. Typically, when you are conducting academic research, which of these five starting points do you use to begin locating information for your research? *Percent of respondents who indicated that each option is the starting point for their research.*



Practical Recommendations

- Make Google searches work for you
- Use social media effectively
- Use content aggregators
- Using popular websites (Wikipedia, YouTube and Flickr) to reach broader audiences
- Make your collection available for learning and teaching
- Improve the user experience
- Reach academic researchers
- Create collection champions
- Integrate with your organisation's systems

<https://www.jisc.ac.uk/guides/make-your-digital-resources-easier-to-discover>

Mapping Collections

- UK has successfully rationalised print journal holdings nationally through UK Research Reserve (UKRR - <http://www.ukrr.ac.uk>)
- Looking now to investigate possibility of similar shared service approach for print monographs.
- Benefits could include:
 - Saving space
 - Addressing de-duplication, cataloguing, preservation, and digitisation
 - More effective document supply
 - Guaranteeing the collection for the nation

Mapping Collections

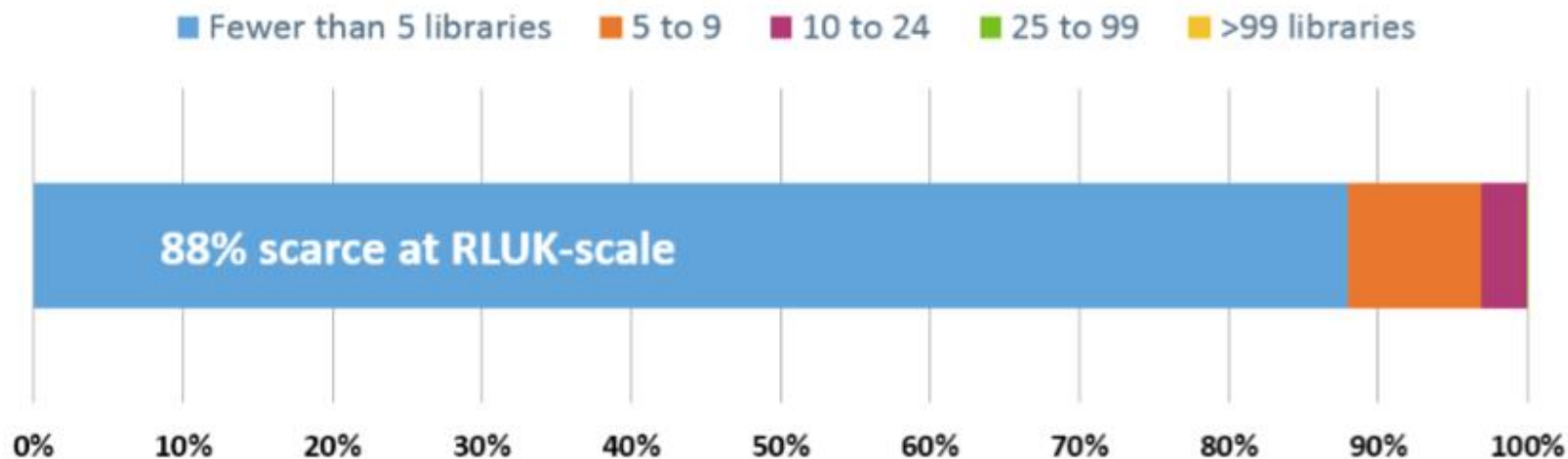
- First step is to map the collections to know what we have
- Robust, accurate data is vital
- Initial analysis carried out with OCLC, using WorldCat data

Mapping Collections

- The RLUK collective collection encompasses **29.4 million distinct publications** of all types, including 20.9 million distinct print book publications.
- Print books in RLUK member collections reflect a **rich global diversity**, with 467 languages and 254 countries of publication represented.
- **Rareness is common** in the RLUK collective collection, with relatively small levels of overlap across RLUK member collections adding scope and depth to the collective resource.

<http://www.oclc.org/content/dam/research/publications/2016/oclcresearch-strength-in-numbers-rluk-collective-collection-2016-a4.pdf>

In-group Duplication of RLUK-held Print Book Titles (N = 20.9M)



WorldCat Duplication of RLUK-held Print Book Titles (N = 20.9M)



Shared Storage/ Shared Collections

- Knowing what we know, can we reduce duplication of low-use print collections
- Allowing shared storage, or even shared collections.
- Can we make use of analysis of collections to prioritise preservation and digitisation targets
- Analysis highlights new unexpected collection strengths and helps to redefine what we mean by ‘special’ collections.

Shared Storage /Collection - Open Questions

- **POLITICS:** Institutional: Disciplinary similarities / differences:
National: Ownership / access issues
- **MODEL(S):** Governance: Legal Deposit material: Just HE or beyond: Number of copies
- **DATA:** Above library scale: e.g. OCLC, Worldcat:
Collaborative services and data: Jisc, COPAC, National Bibliographic Knowledgebase: Collaborative tools:
Greenglass, COPAC Collection Management Tools

Shared Storage /Collection - Open Questions

- **PHYSICAL STORAGE:** BL: Commercial suppliers: Community / consortial-owned/constructed storage
- **DIGITAL SURROGATES:** HathiTrust: Google books: Others: Commercial digitisation projects: Digitisation on demand
- **INTER-LIBRARY LENDING:** BL: Inter-institutional lending/supply now permissible under UK copyright regime
- **INTERNATIONAL DIMENSION:** Examples of good practice: greater efficiencies of scale

Describing Collections – A New Infrastructure for the UK

- Working collaboratively with Jisc and SCONUL we have identified two key needs:
 - Libraries want to make **data-driven decisions** about the management of their print and digital book collections but the data that is currently available does not allow them to do this with confidence
 - Libraries want to ensure that researchers and learners have sustainable and convenient access to digital books but it is currently not obvious what is available or what could readily be made available

National Bibliographic Knowledgebase (NBK)

- To enable libraries to make cooperative collection management and development decisions with reference to the largest possible aggregation of UK library data
- To help libraries rationalize their print collections
- To contribute to the data flows that will make it easier (cheaper and more efficient) to discover, obtain and manage e-books

National Bibliographic Knowledgebase (NBK)

- To support a national level framework that clarifies rights and permissions to access scholarly digital materials (and facilitate the lawful creation of digital copies of inaccessible print materials)
- To support the discovery and delivery of the best quality and most persistently accessible digital version of scholarly materials wherever they are globally available
- To allow *any* provider to integrate with or build more effective discovery, analysis and management tools/services on top of the most comprehensive, open and best quality aggregation of UK library data

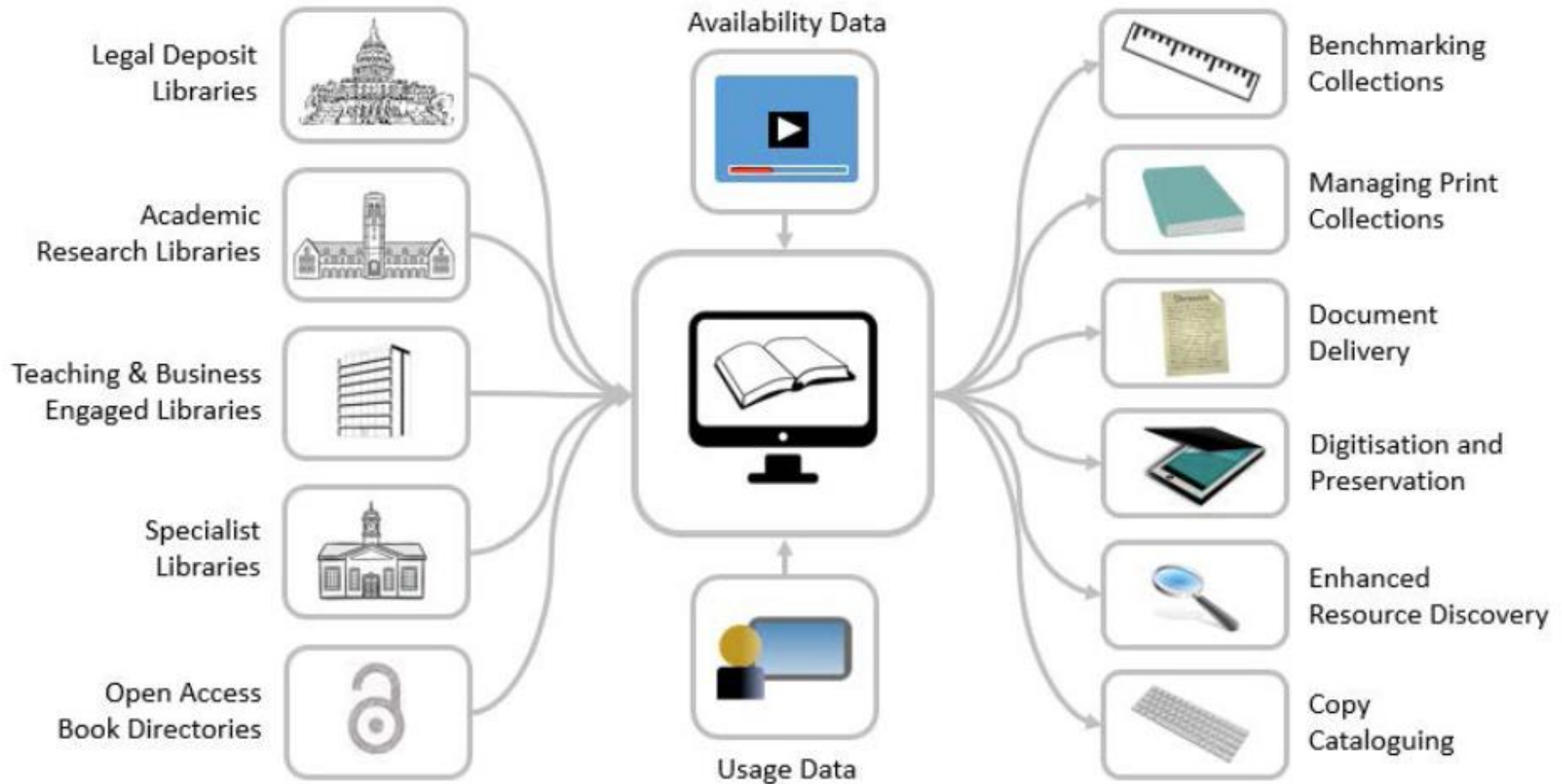
Walled Gardens or Open Commons?

To allow *any* provider to integrate with or build more effective discovery, analysis and management tools/services on top of the most comprehensive, open and best quality aggregation of UK library data

- Increasing vertical integration by suppliers:
 - Research management
 - Repository services
 - Publishing services
 - Bibliographic Database
 - Discovery layers
 - Analysis tools
- How easy is it to get your data out of these systems? How easy should it be?



What is a National Bibliographic Knowledgebase?



'Add-on' Services

- Resource discovery layer on the NBK
- Advice and guidance and community support (particularly around data submission)
- Benchmarking and collection management tools
- Shared print analysis and management services
- Enhanced document delivery functions
- Digitisation and preservation coordination services
- Metadata quality enhancement
- Analysis and dissemination of resource usage and availability data
- Systems to assure data accessibility and reusability
- Copy cataloguing services
- Etc., etc

Shifting Perspectives

- From seeing collections assessed in relative isolation to building tools that allow comparative evidence to be shared and put to work
- From viewing special collections as largely static ones of early-printed books and monographs published before around 1850 to a more dynamic and varied set of criteria that generates value from collections, wherever held in the library
- To a more concrete understanding that a 'National Collection' has to embrace many collections beyond HE to be meaningful

Shifting Perspectives

- From seeing bibliographic data as an ‘unsexy’ and potentially outmoded part of the library landscape to perceiving it as a fundamental building block for the services we want to provide
- From viewing our worth not just in terms of the quantity of ‘stuff’ in our buildings, but also the quality of the services we provide – so valuing both purchased and non-purchased materials if they aid research and teaching.
- Seeing discoverability of ‘bought-in’ and ‘disseminated-out’ material as being two sides of the same coin

Shifting Emphasis?

“Academic libs will increasingly focus on distinctive & unique collections. Rely on coop coll dev for gen material”

Jim Neal, Vice President for Information Services and University Librarian,
Columbia University (via Twitter, 29/04/10)