

The NBK and the UK Distributed Print Book Collection (UK DPC)

Rozz Evans, Head of Collection Strategy, UCL Library Services
Bethan Ruddock, Senior Product Owner, Jisc Library Hub

For many years, the UK's academic library community (including special and national libraries) has been talking about the increasing urgency of a shared approach to ensure the preservation of, and preservation of access to, increasing numbers of scarce print books held in libraries across the UK.

The ongoing move towards digital first provision, pressure on physical spaces, and the challenges posed by ever-increasing demands on staff time and financial resources make this time critical. There is precedent in the form of the UK Research Reserve (UKRR),¹ which addressed the need for a collective approach to the management of print journals and began life as a project in 2007 before transitioning to a 'business as usual' service at the BL in 2019. It has long been agreed that something similar for monographs was needed.

The UK Distributed Print Book Collection (UK DPBC) seeks to address this need and is supported by Research Libraries UK (RLUK) under their 2022 – 2025 strategy, The Library Transforming.²

The RLUK's Collections Strategy Network (CSN)³ Print Storage Group⁴ is leading this work, but it is built on the back of many years of discussion, and some significant pieces of work previously completed are listed at the end of this article for further reading. Credit is due to all those involved across the sector that have brought us to this point.

While this vision was endorsed In May 2022 by the RLUK Board of Directors, it must be emphasised that this shared print collection will need to extend far beyond RLUK libraries, with SCONUL, Legal Deposit Libraries (LDLs), special libraries and Jisc all playing a vital role.

What is it?

In a nutshell, it is a nationally distributed book collection based on a minimum number of retained copies, enabling individual libraries to make local decisions about the retention or disposal of their print items which aren't part of the national collection, knowing they can do so without affecting wider long-term access to that content.

The objectives of the UK DPBC are:

- To maximise the value of the shared national collection
- To preserve access to print books at a national level
- To benefit from the shared collection
- To provide awareness of rarity and scarcity within Library collections
- To give confidence to decision making in the local management of print Collections
- To facilitate decisions on space allocation
- To secure 'at risk' print book material
- To provide reliable access to scarce material

¹ <https://spiral.imperial.ac.uk/handle/10044/1/73162>

² <https://www.rluk.ac.uk/the-library-transforming-strategy/>

³ <https://www.rluk.ac.uk/rluk-csn/>

⁴ The Print Storage group consists of Jane Saunders (University of Leeds), Sandra Bracegirdle (University of Manchester), Hannah Mateer (University of Edinburgh), Joseph Marshall (National Library of Scotland), Michael Williams (University of Cambridge), Sarah Thompson, (University of York), Stuart Dempster (University of Southampton), Rozz Evans (University College London)



Complexities

The aim is simple, and the need is clear, but there are many complexities, and this is why consulting with stakeholders and making sure concerns are discussed is vital. The following are some of the areas that have been identified – this list is not exhaustive!

- **Accuracy of data**

Clearly accurate, quality, shareable metadata is crucial, and the whole project is dependent on this. Work has been ongoing with colleagues at the Jisc National Bibliographic Knowledgebase (NBK) – of which more later in this article.

- **Number of copies:**

The final agreed number needs to be robust enough to accommodate e.g., lost or damaged copies without being over-cautious and thus cancelling out the benefits of such a collective approach. At time of writing, a number between 5 and 7 copies is looking to be most supported option. We hope to make a final decision soon, following some final data analysis of RLUK libraries as well as some volunteer SCONUL libraries.

- **Preserving access to content**

Legal Deposit Libraries (LDLs) are likely to play a crucial part in terms of retention of copies, given they are already committed to permanent retention for a large percentage of their stock. Questions have arisen in terms of how far loan status needs to be a consideration.

Taking loan status into account adds a layer of complexity which may not be feasible to support. This includes the fact that for most libraries loan statuses are subject to change, the NBK doesn't store loan status data, and adding additional data added to retention statements considerably increases the burden on libraries. On balance it seems pragmatic to ensure the national scarcity threshold is set to take this into account.

We don't expect the UK DPBC to drive ILL usage significantly. It's important to remember that we are talking about safeguarding low use items, so by their nature, it's unlikely that a library would be inundated with requests to borrow. However, access **is** a key issue, and we hope that where possible a holding library will be able to lend (although we recognise it may not always be feasible).

Future digitization initiatives, ongoing developments in mechanisms for sharing resources etc. will all have a part to play in the future.

- **Frequency of NBK updates and use of Retention Statements**

Keeping the data in the NBK as current as possible is vital and support will be given to help libraries to implement this where necessary.

There are guidelines for using retention statements that have been agreed with the community.⁵ However, please be aware that these will be revisited over the next 6-12 months to ensure they are meeting the needs of UK libraries, and the UK DPBC work.

Work is ongoing with the NBK to establish methodologies to make it as easy as possible for libraries participating in the UK DPBC to upload retention statements, and to use the Library Hub Compare tool to understand how many copies are available nationally as a routine part of stock editing activities. Ideally retention statements will be added where a decision to retain has been made, but libraries are likely to work at different rates on this, so the key ask will remain to ask libraries not to discard their copy if it is one of x or fewer across the NBK.

⁵ <https://libraryservices.jiscinvolve.org/wp/2020/06/recording-retention-commitments-lhcab-confirms-recommendation/>.

- **Physical condition**

There is scope within the retention statement for libraries to add a condition statement if they wish to. However, we don't propose making this mandatory, as many libraries won't have the capacity to do this. We propose that we set the minimum number of copies at a level robust enough to mitigate the fact that some extant copies may be in poor condition (or even missing).

- **Scope**

We don't propose limiting the scope of the printed books included. We feel that this would add too much complexity. We note that for foreign language material in particular this may leave libraries with items, or collections, that they no longer wish to hold. We propose that to take account of this we develop a mechanism for flagging where the holding library cannot retain items it should, under the terms of the UK DPBC, retain. Further consultation on this will be undertaken (but we don't feel this needs to delay the current work).

- **Sustainability**

This is clearly a key benefit for the UK DPBC. There is the potential for all participating libraries to not only save on storage costs and reduce carbon footprints without disadvantaging users.

What will participating in the UKDPBC involve?

It is understood that different libraries will have different capacities to commit and that will be reflected in various levels of membership under an agreed Memorandum of Understanding (MoU). We have access to the UKRR MoU and have considered some interesting examples of similar international projects such as the EAST distributed shared print program in the US.⁶

However, the following draft guidelines give a flavour of what will likely be expected.

1. To **agree to retain** a copy of a print book if there are fewer than x copies held within the NBK.
2. To **add retention statements** to the metadata for items you have committed to keep for the foreseeable future or until a set date.⁷
3. To **upload metadata on a regular basis** including retention statements to the NBK.
4. To **facilitate access** to items with a retention statement, ideally by ILL where this is feasible.
5. To identify **last copies** and retain where possible (work under way to clarify what to do if a library is physically unable to commit to retain).

Governance and administration of the UK DPBC is being considered and further information will be available in due course.

⁶ <https://eastlibraries.org/about-us/about-shared-print/>

⁷ <https://libraryservices.jiscinvolve.org/wp/files/2020/06/LHCAB-Retention-Statement-Final-0620.pdf>

Role of NBK and metadata

We believe that Jisc and NBK will play a key role in the UK DPBC. Success will rely on good quality, reliable holdings data: consistent use of retention statements; shared commitment; minimal administrative burden for individual libraries; continued development to support.

How Library Hub is supporting the UKDPBC

Jisc currently runs three Library Hub service: Library Hub Discover, a freely available service for resource discovery (discover.libraryhub.jisc.ac.uk); Library Hub Compare, which facilitates collection development and management (compare.libraryhub.jisc.ac.uk); and Library Hub Cataloguing, a MARC records download service (cataloguing.libraryhub.jisc.ac.uk).

The Jisc Library Hub services are built on a key piece of infrastructure, the National Bibliographic Knowledgebase (NBK). The NBK is an aggregation of 51,612,311 records created from 152,616,854 records contributed by 204 institutions. These are a mixture of UK HE and specialist research institutions, plus the national libraries and the library of Trinity College Dublin. A full list of contributing libraries can be found at <https://discover.libraryhub.jisc.ac.uk/about/libraries/>.

Bibliographic records from these libraries are supplied to the Library Hub team, usually in MARC format, and they are then loaded into the NBK data lake. Out of this data lake, a deduplication process is run which merges records for the same item. This dataset is used to provide the Discover and Compare services, while an unduplicated database providing permissibly-licensed MARC records as provided by the libraries is used for the Cataloguing service.

This database is kept updated by regular data uploads from contributing libraries; however the lag inherent in the system mean that the NBK is never an entirely comprehensive or current reflection of the UK national collection.

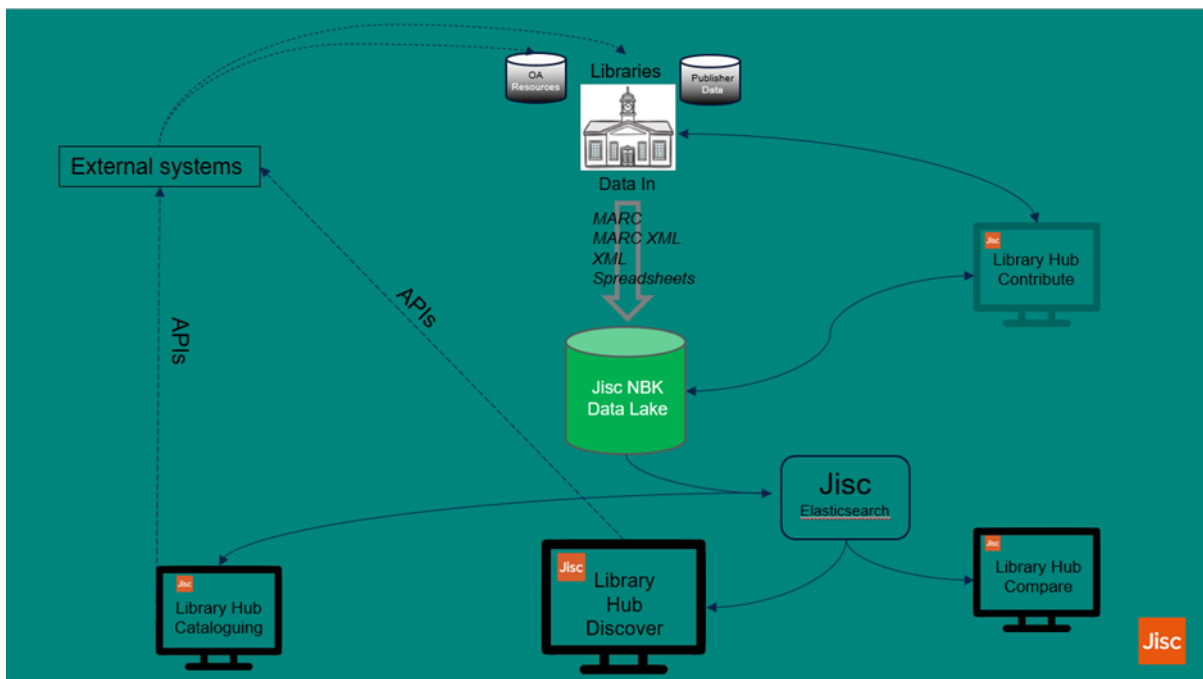












Figure: an overview of the architecture of the Library Hub services

As the NBK is a unique aggregation of the holdings of UK libraries, we are very well-placed to be a data source for the rarity assessment work being undertaken as part of the UKDPBC work.

Following initial conversation with the RLUK CSN members in 2022, we produced a spreadsheet giving an overview of the number of print books held by 10 or fewer libraries (9 or fewer, ..., down to 1 holding library) in the UK, with breakdowns by consortia and individual library. An example section of the spreadsheet can be seen below.

No. holdings institutions	Totals		Consortia: at least 1 of the holding institutions is in the consortium				
	Result	At least 1 record includes NBK-R	RLUK	SCONUL	SCONUL NOT RLUK	Legal Deposit Libraries	Other (not SCONUL or RLUK)
Total deduplicated records in the NBK	50341867	82314	40445825	44481163	4801890	27233665	5107978
Total deduplicated print book records in the NBK	33217075	80964	28402962	29939782	2092010	20814168	2722216
No. of records with 10 or fewer holdings institutions	31524309	52916	26712578	28249373	2091981	19162186	2721836
No. of records with 9 or fewer holdings institutions	31344460	51185	26532735	28069415	2091866	18994016	2721836
No. of records with 8 or fewer holdings institutions	31132169	49217	26320640	27857074	2091620	18792396	2721836
No. of records with 7 or fewer holdings institutions	30852371	47069	26041719	27577236	2090703	18529767	2721836
No. of records with 6 or fewer holdings institutions	30453041	44453	25643107	27177962	2090041	18154166	2721836
No. of records with 5 or fewer holdings institutions	29745526	41819	24936427	26470154	2088910	17476084	2721834
No. of records with 4 or fewer holdings institutions	28955544	38966	24149415	25680085	2085828	16742421	2721827
No. of records with 3 or fewer holdings institutions	28023153	35691	23230237	24747957	2072716	15937647	2721660
No. of records with 2 or fewer holdings institutions	26622406	31374	21853117	23351125	2051292	14784826	2719644
No. of records with 1 holding institutions	23608892	24084	18960607	20392338	1965634	12586242	2684462

Investigation of these figures soon showed that they were not an accurate representation of the print holdings in UK libraries, with many libraries reporting that the figures provided grossly overstated their print book holdings. The issue was that records in the Library Hub are deduplicated regardless of format. Library Hub Discover was designed with the aim of satisfying the end-user need of 'I want this book, how can I get it?'. This means that records for print and e versions are merged into the same record. When you filter the deduplicated dataset by 'print', you won't get any results where all of the items are electronic, but you will get mixed print and e as well as all print. For instance, this record...

- 10 Holding libraries
- [British Library](#) 
 - [University of East Anglia Library](#) 
 - [Institute of Advanced Legal Studies Library](#) 
 - [LSE Library](#) 
 - [National Library of Scotland](#) 
 - [National Library of Wales / Llyfrgell Genedlaethol Cymru](#) 
 - [University of Nottingham Libraries](#) 
 - [University of Oxford Libraries](#) 
 - [UCL Library Services](#) 
 - [University of Warwick Library](#) 

...was being counted as 'print, with 10 holding libraries', despite 1 of them being electronic.

The Library Hub Compare service already has a deduplication system which avoids this mixed-format problem, but because of system demands that can handle a maximum of 50,000 records at a time; not feasible for work across the entire dataset.

What we have done is to create a new development version of the NBK database with a new index which only counts the number of print holdings, and excludes other formats from the count. This also enable us to provide more nuanced reports to the individual libraries, only including consolidations where they hold a print copy.

We are currently in the process of creating a set of dashboard visualisations which we will supply to the RLUK CSN members, along with the individual libraries who are taking part in the initiative.

Next steps after this will be to create a report allowing libraries to ingest and examine a list of their holdings which meet the rarity threshold. Initially these lists will be generated by the Library Hub team, with the intention of eventually making them self-service through the Library Hub Compare service.

While there will always be some errors based on either scarce or inaccurate metadata, especially where the format has not been correctly identified in the record; or rare misconsolidations (from over 45 million searches a year on the Discover we have on average fewer than 10 misconsolidations reported to us), we are confident that the data will allow for confidence in an accurate assessment of print rarity in libraries across the UK, and are pleased to be supporting this important work.

References and Further Reading:

Theo Stubbs et al., *UKRR Final Report*, August 2019, DOI: <https://doi.org/10.25561/73162> (accessed 13/11/23)

The Library Transforming: RLUK's Strategy for 2022 – 25 <https://www.rluk.ac.uk/the-library-transforming-strategy/> (accessed 22/11/23)

RLUK Collections Strategy Network <https://www.rluk.ac.uk/rluk-csn/> (accessed 22/11/23)

White Rose Library staff, *Understanding Collections Overlap: An Investigation Into White Rose Libraries Collections Using Collection Management Tools*, Final Report, July 2017, <https://blog.ccm.copac.jisc.ac.uk/wp-content/uploads/sites/13/2017/09/Understanding-collections-overlap-final-report-july-2017-including-appendices.pdf> (accessed 13 November 2023)

Constance Malpas and Brian Lavoie, *Strength in Numbers: The Research Libraries UK (RLUK) Collective Collection*, 2016, <https://www.oclc.org/research/publications/2016/oclcresearch-strength-in-numbers-rluk-2016.html> (accessed 13 November 2023)

Information Power, *Feasibility Study on Monographs*, Final Report, June 2017, https://www.sconul.ac.uk/sites/default/files/documents/Feasibility_Study_on_Monographs_final_report_June_2017_0.pdf (accessed 13 November 2023)

Theo Stubbs, *An Extension of UKRR Into Low-Use Monographs: Does Appetite Exist?*, November 2018, <http://hdl.handle.net/10044/1/69379> (accessed 13 November 2023)

EAST – Eastern Academic Scholars Trust distributed model shared print program <https://eastlibraries.org/about-us/about-shared-print/> (accessed 22 November 2023)

Library Hub Community Advisory Board – *Retention Statement (Final)* <https://libraryservices.jiscinvolve.org/wp/files/2020/06/LHCAB-Retention-Statement-Final-0620.pdf> (accessed 22 November 2023)

[Library Hub Discover](https://discover.libraryhub.jisc.ac.uk/) <https://discover.libraryhub.jisc.ac.uk/> (accessed 22 November 2023)

[Library Hub Compare](https://compare.libraryhub.jisc.ac.uk/) <https://compare.libraryhub.jisc.ac.uk/> (accessed 22 November 2023)

[Library Hub Cataloguing](https://cataloguing.libraryhub.jisc.ac.uk/) cataloguing.libraryhub.jisc.ac.uk/ (accessed 22 November 2023)

NBK Contributing Libraries <https://discover.libraryhub.jisc.ac.uk/about/libraries/> (accessed 22 November 2023)