Defining National Solutions for Managing Book Collections and Improving Digital Access

Neil Grindley

Head of Resource Discovery, Jisc
Neil.Grindley@jisc.ac.uk, orcid.org/0000-0001-9808-3032

Paola Marchionni

Head of Digital Resources for Teaching, Learning and Research, Jisc
Paola.Marchionni@jisc.ac.uk, orcid.org/0000-0002-9544-5410

Abstract

In 2013 the National Monographs Strategy (NMS) project in the UK explored the potential for a national approach to the collection, preservation, supply and digitisation of scholarly monographs. The resulting NMS Roadmap recommended seven components, believed to be critical for the provision of a national monograph infrastructure. This paper discusses how Jisc prioritised three of the recommendations and started planning for the development of a National Bibliographic Knowledgebase (NBK) in association with key stakeholders representing UK Higher Education and the British Library. In parallel, Jisc also explored recommendations around a national digitisation strategy and national licensing approaches by establishing a 'Digital Access' strand of activities.

Key Words: monographs; collection management; books; ebooks, metadata; bibliographic data; digitisation; licencing; digital surrogates

1. Introduction

The work set out in this paper refers to activity over the last 2–3 years and sets out forward plans for building a new national service and designing digital access strategies. The organisation that is leading and managing the work is the UK charity, Jisc, which provides digital solutions for UK education and research.[1] The scope of the work originates from a Jisc-led initiative called The National Monograph Strategy (NMS) which convened a large group of relevant stakeholders from libraries and academia to examine and formulate recommendations for the UK academic sector in relation to monographs.[2] The NMS Roadmap (Showers, 2014) was published in September 2014 and described seven components that the group believed were critical for the provision of a national monograph infrastructure. They were as follows:

1. A national monograph knowledgebase
2. A national digitisation strategy
3. A 'systemic changes think tank' group
4. New business models for monograph publishing
5. A negotiated national licence for access to digital scholarly monographs
6. A shared monograph publishing platform
7. An impact metrics framework to demonstrate the value of monographs

One of the challenges of convening the original NMS discussion and of following it up is that it is very difficult to constrain a conversation around the concept of 'the monograph.' From whatever point you begin, the discussion quickly expands to include a very wide array of issues, including library collection management, metadata quality, diversity of formats, availability of digital surrogates, publishing processes and platforms, appropriate infrastructures, governance, trust, and so on and so forth. However, from the present vantage point it is now much clearer to Jisc that some of the seven recommendations were a higher priority than others and there is now much more clarity about what specific actions should be taken. This is partly due to an inevitable evolution over time in the user community's requirements; but it is also due to the work that has been done to turn a set of conceptual recommendations into an actionable plan that has solid stakeholder support and is able to justify the required levels of investment.

On consideration of an NMS phase 2 work plan, Jisc took an early decision to deprioritise two of the seven recommendations. The first of these was (3) the 'systemic changes think tank,' on the basis that such strategic thinking would take place in a devolved way across various stakeholder and governance groups as a matter of course. The other one was (7) the 'impact metrics framework' due to the fact that other entities or collaborative partnerships (particularly those with better incentives and/or a clearer mandate to quantify the value and impact of research monograph publishing) would be better placed than Jisc to lead on such a topic.[3] Further discussion across teams at Jisc and ongoing work (Collins & Molloy, 2016), clarified that (4) 'New business models' could be subsumed into the Open Access Monograph work that was being taken forward by Jisc Collections in collaboration with initiatives such as OAPEN[4] and Knowledge Unlatched.[5] It was also established that (6) the 'shared monograph publishing platform' was still at an early conceptual phase and could be separated off and managed as an R&D initiative by staff in Jisc Futures.[6]

2. Defining the Problem Statements

Despite shrinking the actionable NMS recommendations from seven down to five and then devolving responsibility for two more, defining the remaining activities into coherent strands of work was still a significant challenge. The problem statements that were eventually alighted upon were the product of much additional discussion between Jisc and relevant stakeholder groups throughout 2015. In support of these discussions, Jisc commissioned a consultation and path-finding report to consider the future of bibliographic data services in the UK. The resulting report – the *Bibliographic Services Implications Study* (Hammond, Kay, Schonfeld, & Stephens, 2015)—contained an influential set of recommendations which crucially elicited the support of the leading academic library membership groups in the UK (RLUK[7] and SCONUL[8]). Building on this report and the broader consultation, it was feasible by early 2016 to present the required solutions as a response to two prioritised and broadly agreed assertions.

A: 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

B: 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

Boiling down the challenges into two broad categories was extremely useful in terms of being able to divide up responsibility and allow tangible progress to be made. In general terms, statement A focuses on issues to do with metadata, metadata quality and the aggregation of metadata on a large scale; and statement B focuses on content and access to that content. A two-pronged strategy was agreed whereby the metadata issue would be addressed by the specification and development of a new service – the *National Bibliographic Knowledgebase (NBK)*; and the content issues would be tackled by a series of actions characterised as addressing '*Digital Access.*' The rest of this paper sets out what has been taken forward in those two areas of work and provides an update on the progress that has been made since the presentation at the LIBER conference in early July 2016.

3. The National Bibliographic Knowledgebase (NBK)

The *Bibliographic Services Implication Study* (Hammond et al., 2015) was published in September 2015 and set out much of the strategic and tactical framework for going into a new phase of Jisc service provision in the area of bibliographic data. A summary of the most pertinent recommendations from the study are as follows:[9]

1. The UK has a fundamental need for a new national-scale service to drive a range of required functions
2. The new service should consist of an aggregated database and its management should be outsourced to an organisation that is capable of delivering the service as core business at scale
3. The primary focus of future effort should be on supporting UK academic libraries with collections management. Resource discovery and records delivery are of secondary importance
4. The data contributed to the new system must remain shareable and reusable by all contributing organisations and by other relevant organisations that support discovery and records delivery
5. The route to greater impact for contributed library data is through exposure to global search engines and other high impact web-scale

channels rather than through reliance on Jisc-funded discovery interfaces

6. The new system should combine knowledge about both print and digital publications for services to be efficient and effective

As of early 2017, it is now possible to describe progress against all of these objectives and to set out the ambitious plans that Jisc and its strategic partners have put in train to address what was set out in the National Monograph Strategy.

3.1. A New National Scale Service

The NBK will be a new service but will replicate components of existing services. It will supersede the current Copac and SUNCAT services that Jisc provides and will work at much greater scale, with more diverse data sources, more functionality and greater flexibility. The goal of the NBK is to help transform how libraries manage their collections, provide access to resources and collaborate with each other. It will provide a sustainable fit-for-purpose next generation national data infrastructure that practically supports libraries to make the transition from a print-first to a digital-first paradigm.

It will build on and surpass the functionality of Copac which is the nearest current equivalent service that Jisc provides and which currently aggregates data from around 90 libraries. The NBK will, over time, include catalogue data from more than 225 academic and specialist libraries, and by doing so it will more effectively support the management of library collections so that they are optimized for contemporary research and learning needs. By drilling down into the 'long-tail' of holdings across the UK it will support the formulation of a more joined-up national strategy around the retention of print materials. It will aggregate bibliographic data with availability and usage data and will facilitate more efficient access to eBooks, digitised books and journals.

Another high-level objective of the NBK is to make a positive contribution to the overall quality of data that circulates around what might be referred to as a bibliographic data 'ecosystem.' The NBK will act as a positive agent of change in relation to the accuracy and effectiveness of metadata; the standards that are adhered to and promoted across the sector; and the development of a national approach to the use of authority controls and identifier frameworks in relation to bibliographic resources. There would seem to be broad agreement that the

library sector probably needs to move on from a focus on MARC format data and legacy workflows; but there is also acknowledgement from bodies seeking to be progressive that practice within libraries is very slow to change.

> *Librarians used to working with full MARC records may not easily grasp that a move to the more atomic level of individual statements will make possible innovation in areas like new services, localization, and distributed data improvement. Outside of libraries, these activities are building and taking shape, but most librarians aren't yet monitoring those activities, mostly because they have yet to appreciate the connection with the library world.* (NISO, 2014)

The NBK will also support and facilitate the most unhindered flow of data possible in order to maximise the prospects of users encountering data that will lead them to library resources wherever they may be looking for it. This may be via a Google search; or within a commercial discovery system environment; or via another specialist library aggregator system.

3.2. Outsourced Service Management

Following an extensive procurement and competitive dialogue process, Jisc selected OCLC as their service provider and partner to build and deliver the NBK. OCLC are uniquely positioned to make library data globally available via their WorldCat service and to connect library data-hubs at scale. They are a known quantity and already provide national and regional bibliographic infrastructure in a number of countries, including in Australia, France, Germany, Switzerland and the Netherlands. They have also worked collaboratively with RLUK libraries in the UK to undertake analysis to explore the concept of the 'Collective Collection' (Malpas & Lavoie, 2016). Jisc has entered into a multi-year agreement with OCLC to work closely together to develop the solution that UK libraries need. Jisc will ensure that the service is owned and controlled by the community of libraries that contribute data to the aggregation and will share data management responsibility on the OCLC-provided CBS platform (Central Bibliographic System).

3.3. Focus on Collection Management Functionality

Collection management has emerged as a much higher priority for libraries over the past 3 or 4 years and this has largely been driven by the need for

universities to focus closely on managing the space they have available for learning and teaching purposes.[10] Libraries are now carefully considering their print monograph holdings in the same way that they do for journals which have been under scrutiny for some years and there is an urgent and well-defined requirement to provide an authoritative source of data that will support library decision-making to transfer, relegate or withdraw titles.

Jisc currently offers the Copac Collection Management Tool to UK HEI's (higher education institutions) and it supports a number of use cases (Jisc, n.d. b):

- Identifying last copies among titles considered for withdrawal
- Identifying collection strengths
- Deciding whether to conserve a book
- Reviewing a collection at the shelves
- Prioritising a collection or item(s) for digitisation
- Subject search—collection development and marketing

The NBK will replicate or exceed the functionality of the Copac Collections Management (CCM) tool either by adopting the CCM toolset and integrating it with CBS; or by developing a native CBS tool.

3.4. Reusability of Data

The data that finds its way into the NBK will be managed and licensed so that wherever possible it will be available for discovery and re-use by other systems. As well as the CBS metadata management system, OCLC will provide the CBS publishing platform which will contain the enriched de-duplicated master records. This will provide a mechanism for the syndication of data, either freely or according to a fee model depending on circumstances and sustainability requirements.

Reuse of data will be provided via batch export of files on request; a scheduled 'push' mechanism; or a 'pull' service using OAI harvesting or other API mechanisms. It will be possible to select data using the indexes that have been defined on the CBS platform and the log-files that are maintained for database updates. Common selections will be per library, group of libraries, dates, material types and subjects. For the creation, maintenance, scheduling

and monitoring of export jobs a web-based application will be used called CJM (CBS Jobs Management). With CJM, Jisc staff will be able to create, maintain and schedule jobs. Exports that need to be produced regularly can be automatically launched daily, weekly, or at other frequencies.

3.5. Exposure of Library Data to Global Search Engines

One of the gains that UK libraries can expect from a Jisc/OCLC collaboration is that NBK data will be visible in WorldCat. Through WorldCat, NBK data will gain greater visibility in OCLC discovery applications and in third-party applications which take part in the OCLC Web Syndication program. If, for any reason, libraries would prefer their data not to be published in WorldCat, then it will be possible to exclude records from that synchronisation process.

In addition to OCLC Web Syndication, the CBS publishing platform enables website creation specifically for search engine crawlers that will give access to all records that should be made available for search engine harvesting. The data would be represented using schema.org mark-up, which is usually preferred by any such web service. In this special web site, filters for available records and represented data elements can be applied.

Data that is syndicated through WorldCat will be presented in web services (such as Google, Bing, Wikipedia, etc.) in a way that allows linking back to WorldCat. The user will be directed from the initial web site to the respective WorldCat entry for that title or author. The authority data will also be syndicated (the Wikipedia pages of many authors contain the VIAF number, which also makes it available on the Google knowledge card).

3.6. Knowledge about Print and Digital Publications

OCLC will load eBook vendor collections onto the NBK and the expectation is that these will be regular and automated using data feeds from eBook vendors according to agreements that are in development by Jisc in association with libraries and third party organisations. The solution will be extended to eBook data that comes directly from libraries, thereby providing a platform for community supported management of shared collections.

CBS functionality will include the ability to regularly test the availability of an eBook and identify 'broken' links, and log the results. Analysis of the causes of broken links will be undertaken and batch change functions will be applied where it is possible to make corrections. CBS supports dynamic FRBR clustering and this could be used for the creation of FRBR work records. The NBK will seek to connect with international sources of eBook metadata/content wherever possible and integrate them as data sources. For example, both Jisc and OCLC have strong connections with HathiTrust and the NBK will incorporate data and links to content from their openly available files.

The last of these six objectives is a critical goal for the NBK and is one of the key differentiators between the capability that current systems provide and the functionality that the NBK aims to deliver. Copac is primarily geared towards searching library holdings records to identify the location of print materials in libraries. The NBK intends to extend and expand that scope by sharpening the focus on the availability of e-resources, wherever versions and copies may be available for use in the UK. This whole area of work was the subject of a parallel activity within Jisc as the specification for the NBK was being considered and assembled. The next section describes the complementary analysis that was undertaken during 2015/2016 to identify and analyse institutional requirements for eBook resources and to get a clearer picture of demand and supply issues.

4. Digital Access

Over the last year or so, the Digital access strand of activities[11] has progressed in tandem with the development of the NBK addressing two of the recommendations in the original NMS report: a digitisation strategy to support the building of a national digital research collection and a national licence to support access to digital monographs through a negotiated national agreement.

A key driver for both recommendations was the ambition to increase access to monographs in digital form, i.e. monographs that are not already currently available digitally, for the benefit of academics and students, and also enable collections managers to make more informed collection management decisions based on information on what is available where, and in what format. Another aim the two recommendations shared was to ensure that any strategic approach to increasing access to monographs was evidence-based, and founded on analysis of requirements from institutions and their patrons.

As both recommendations focused on increasing digital access, the scope of our work was from the beginning focused on print collections, rather than born digital monographs, and on "monographs" rather than textbooks. It soon became apparent, however, that the term "monographs" was too narrow for the purpose of our work at this initial stage, and also that the borderline between what constitutes a monograph or an academic book and a textbook is not always clear and self-evident. There is more on this below.

4.1. Engaging with the Community

If we were to tackle recommendations about a digitisation and a licensing strategy, the first questions the team was confronted with were, what should a strategy achieve? Where should we start from? What kind of titles should a digitisation and/or a licensing strategy focus on? What would be useful criteria to consider? Should we privilege Public Domain material? Make use of hard-won copyright exceptions such as the one on Orphan Works[12]? Prioritise out-of-commerce books? What about books still in-copyright given the difficulties of clearing rights for 20th and 21st centuries publications (Freire, Scipione, Muhr, & Juffinger, 2013)? What would a business model and a service model to support digitisation and licensing at scale look like? How should we balance the immediate needs of practitioners within libraries in satisfying the day-to-day demand of academics and students with more strategic and ambitious aspirations of the HE community?

We needed to start with identifying the use cases, who needed what and why, and drill further into the problems that a digitisation and a licensing strategy needed to address.

We embarked on a series of preliminary informal conversations with librarians and collection managers from a range of different universities who provided us with some initial insights into their problems with regard to the day-to-day management of book collections and the provision of digital resources to their users. We found that key issues were space and the need to weed collections based on usage levels. This is certainly not a new issue as libraries have been struggling with the management of high volume-low usage books for years (Bracegirdle, 2012). Satisfying readers' growing expectations on quick and easy access to (digital) resources was also a recurrent concern, in the words of some of our interviewees,

> *"it is an important issue, as being able to provide access to digital copies of content for reading lists is a key priority."*

and

> *"… so much of our purchasing is based around reading lists. So if we cannot provide a book digitally, then potentially a large cohort of students will be unable to access it in that format."*

Many institutions have a "digital first" policy when acquiring new books, although print is still purchased when no suitable alternative is available. Time and budget are also major constraints and book availability information is hard to find.

This preliminary research informed the design of what we called the Digital access pilots project. Following an open Call for Participation (CfP)[13] we enlisted 10 institutions, a mixture of research intensive, specialist universities and teaching and learning focused institutions, to help us define the problems and identify the type or categories of books that libraries most needed to provide access to in digital form, and why. At this stage, we didn't know whether demand might be primarily for out-of-copyright books, in-copyright, in or out-of-commerce, or if there might be any publishers or subject disciplines emerging as most "in demand."

Guided by the informal conversations with senior librarians and practitioners, and discussions with the BIBDOG[14] group, we took the view that at this stage of the project we would have a more inclusive view of what constitutes a "monograph." The CfP adopted a broad definition:

> *At this stage we are adopting a broad definition in order to get more information about libraries' needs. We include most types of academic book but it must not be a core textbook. A core textbook is defined as something written specifically to serve the needs of students and lecturers following a course. There are no other restrictions since we would like to know what sort of books libraries would prioritise above others.*

One of the key requirements for institutions to participate in this project was that they would provide us with a list of up to 100 titles of books each that the library had been requested to supply but which they couldn't fulfil, for

whatever reason, so that we could have real data to work from. We supplied libraries with an initial template to gather the titles and assembled a list of over 1200 titles. However, it soon became apparent that there were going to be a number of challenges in analysing the data.

Libraries reported that the data they supplied for this pilot project had not, by necessity, been collected systematically, the bibliographic information wasn't always accurate in the way it was available to them, and a great deal of cleaning up and standardisation of the data had to be done by the team against the Nielsen BookData Online database (Nielsen, n.d.). This in itself posed its own challenges as we found that the bibliographic information on the database was not always reliable or expressed in the way that was useful for our research. In addition, it appeared that a number of titles had been included which were more likely to be textbooks, but we couldn't be absolutely certain simply judging from the bibliographic information. This resulted in us working off two sets of data, a larger one (N=1216), which was the aggregation of titles originally provided by the ten libraries, and a slightly smaller and "cleaner" one (N=1117), where we removed about 100 titles which we recognised as most certainly textbooks – there may have been more.[15] Despite the two data sets, however, the overall picture of results did not change by any significant amount.

4.2. What we Found: A Bird's Eye View

A health warning before delving into the data: this is a sample of data gathered from 10 institutions and therefore is likely to contain a certain degree of bias and may not be representative. In addition, although best effort was applied to the checking and cleaning up of the bibliographic metadata through the Nielsen database, there may still be inconsistencies where information is simply not available or in cases where titles have changed their availability status over time. Our objective, however, was to look for high level patterns, rather than achieve complete accuracy, to see if there were some key messages emerging, and we feel that the project achieved that.

4.2.1. Main Availability Problems
We asked libraries to tell us the main problems they had experienced in accessing the titles they submitted to us based on the following categories that we provided, as shown in Figure 1:

Fig. 1: Library categorised titles (N=1216).
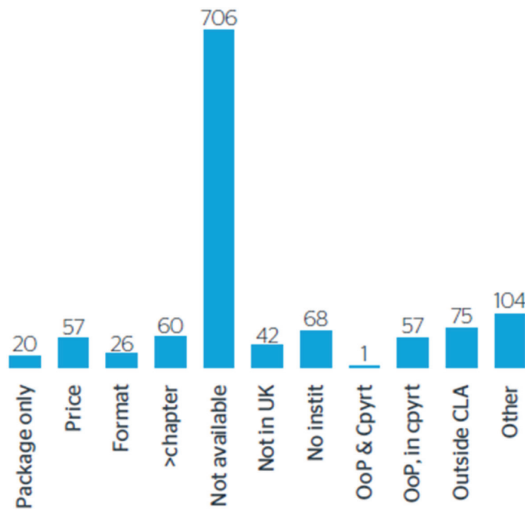
Key

Package only – the title is only available as e-book as part of a larger package	**No instit** – no institutional licence available
Price – the price is too high	**OoP & Cpyrt** – out of print and copyright
Format – the type of digital format (pdf, epub, etc.) is unsuitable	**OoP, in cpyrt** – out of print and in copyright
>chapter – more than one chapter is needed from this book	**Outside CLA** – the title is outside the current CLA licence
Not available – no e-book is available at all	**Other** – a different problem to the ones listed
Not in UK – there is no e-book available in the UK	

By far the biggest problem was that the titles libraries wanted as an e-book were simply not available—or at least this was the libraries' perception based on the availability information they had access to, typically through a select number of aggregators or book vendors they use as suppliers. This is an important point to note. As it turned out at the time of writing during phase

2 of this study[16], a number of titles may in fact be available as ebooks, for example as different editions. However, based on the information they had, as far as the librarians were concerned, they did not have knowledge of, or could access, those titles in a digital format. This resonates with other recent surveys on key concerns librarians have in relation to publishers' provision of books/ebooks such as around publishers' pricing strategy and libraries' budgets, licensing models and accessibility (Folan & Grace, 2017).

The other problems, although each much smaller, did account together for a large portion of the data, as the pie chart in Figure 2 below illustrates:

4.2.2. Reasons and Mode of Access
When we asked libraries to tell us the reasons for wanting a given book, this was mostly to fulfil reading list requests (80%), as shown in Figure 3, while 17% were for research, and 3% for preservation purposes, accessibility or other purposes.

Libraries also told us that they require a certain degree of scale in the number of concurrent users, as shown in Figure 4, with the great majority of titles requested being for between 5-50 concurrent users and greater than 50.

Libraries also require flexibility with remote access from out of campus within and outside of the UK. Two-thirds of the titles were needed for remote access

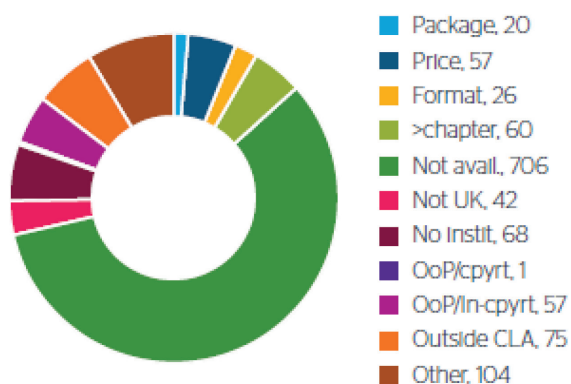*Fig. 2: Titles categorised by problem (N=1216).*

- Package, 20
- Price, 57
- Format, 26
- >chapter, 60
- Not avail., 706
- Not UK, 42
- No Instit, 68
- OoP/cpyrt, 1
- OoP/In-cpyrt, 57
- Outside CLA, 75
- Other, 104

Fig. 3: Primary use (N=1216).

Reading list, 972
Research, 210
Preservation / accessibility, 22
Other, 12

Fig. 4: User access (N=1216).

Single user, 71
2-5 concurrent, 151
5-50 concurrent, 502
>50 concurrent, 374
Undeclared, 118
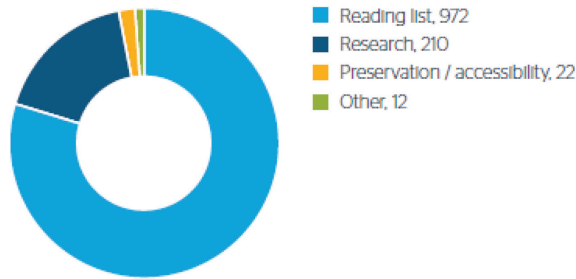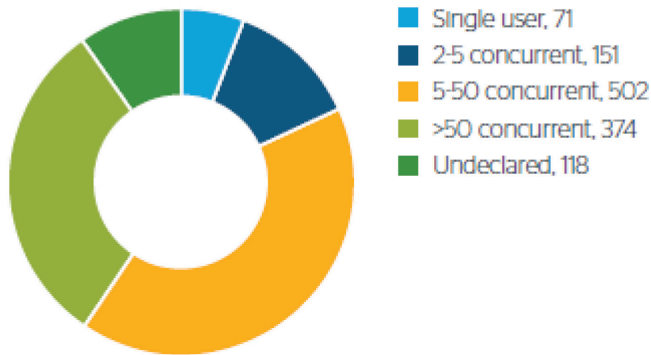
in the UK and the rest of the world, and one third for remote access in the UK only. The project's final report identifies a number of scenarios based on real library workflows typical of situations that libraries find themselves in when trying to satisfy requests for books for reading list purposes or research projects and the kind of barriers they come up against, as highlighted in Figure 1.

4.2.3. Availability and Status of Titles

We conducted further analysis on the smaller "cleaner" data set (N=1117) to gain a better understanding of the availability and copyright status of the books and to see whether there were any publishers that may emerge as dominant. The chart below (Figure 5) shows the distribution of titles by decade of publication cross-analysed with the availability status.

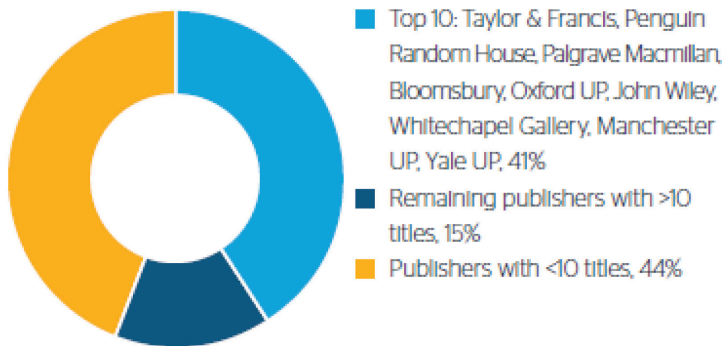*Fig. 5: Titles by date and availability, excluding textbooks (N=1117).*
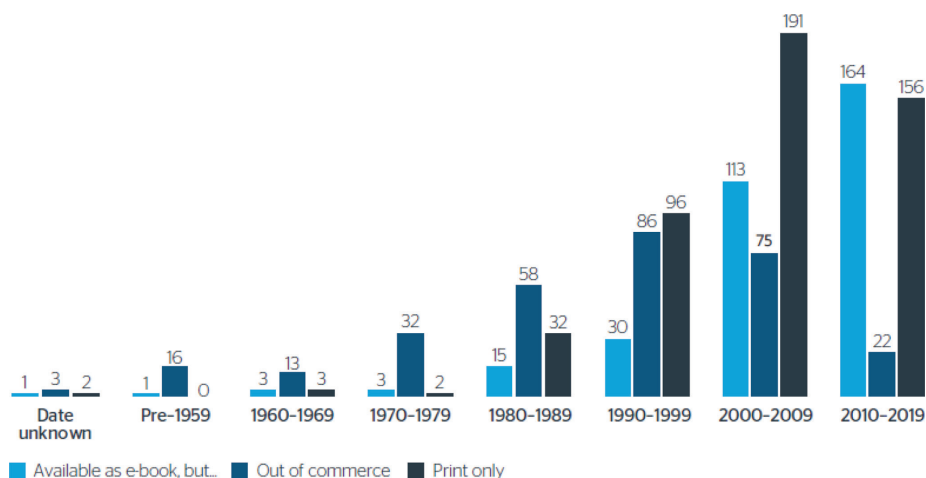
One of the most interesting findings was that there was hardly any overlap of titles requested by the institutions, only six were requested by two institutions, and one by three institutions. However, as this was a skewed sample, a larger aggregation of titles at national level might reveal a different picture.

Another finding that we hadn't anticipated was that most of the requested books were published from the 1960s onwards and therefore likely to still be in copyright, hardly any public domain book was included. Having checked the titles against the Nielsen database, we were also able to estimate with some degree of confidence whether the books were available as in-print only (no ebooks), out of commerce (i.e. no record or information existed in the Nielsen database), or were indeed "available as eBook, but" not in a way that was useful to libraries. Again, it is worth remembering that there may have been different editions/ISBNs of these titles available as ebooks, but based on the availability data that the libraries could check against, they didn't seem to be available.

Finally, we looked at the breakdown of publishers, as exemplified in Figure 6.

Our estimate (after checking imprints and sales of companies as far as possible) is that the 1117 titles were published by 291 different publishers. A large

Fig. 6: Publishers breakdown (N=1117).

Top 10: Taylor & Francis, Penguin Random House, Palgrave Macmillan, Bloomsbury, Oxford UP, John Wiley, Whitechapel Gallery, Manchester UP, Yale UP, 41%

Remaining publishers with >10 titles, 15%

Publishers with <10 titles, 44%

proportion of just over 41% are accounted for in the top 10 publishers, each with more than 20 titles. Within the top 10 Taylor & Francis Ltd. and Penguin Random House accounted for 150 of the titles, 32%. A further 165 titles, 15% can be accounted for by the next 13 publishers, each holding more than 10 titles per company. The remaining 268 publishers had ten or fewer titles each, 44%. Within this long tail of the 268 publishers, 185 companies held one title only and one was untraceable.

4.3. High Level Requirements

Despite the fact that libraries are on the whole able to satisfy the majority of book requests, the participating institutions stressed how critical it is for them to provide access to digital versions of books to ensure they meet the requirements and expectations of their staff and students and in particular for reading lists purposes. One of our pilot institutions stressed how:

> *"The titles passed to you for investigation were the tip of the iceberg, sourced from reading lists submitted by our History Department. Doing a systematic trawl through reading lists from all departments would reveal a great many more titles where the demand for an e-book in recent years has gone unsatisfied."*

In planning for a potential solution, or sets of solutions, to these problems, key requirements that have emerged from this project are the ability to:

1. aggregate at scale problem titles mainly from libraries' reading lists
2. check reliability of bibliographic data (publishers, ISBNs..) from libraries against authority source
3. obtain more permissive licences to produce digital versions of books to satisfy access needs
4. cater for a "long tail" pattern of requests from libraries
5. keep cost of digital copy to no more than a print copy, if available, max £100
6. deliver an on-demand service for digitisation/provision of digital copy. This seemed to be the most appropriate route to satisfy "just in time" requests, possibly through existing mechanisms such as the British Library document delivery service or services provided by universities that might have spare capacity in digitisation
7. create ebooks in an appropriate format: searchable pdf as minimum but epub or HTML5 are preferred. Accessibility for users with disabilities is still a big problem.

5. What we Learnt

When we embarked on the digital access project there were a number of unknowns in relation to the drivers for the demand for digital copies of books and the type of books in question. Some of our findings were not what we expected, and key learning points for us have been:

- for libraries, the highest priority is to resource reading list requests regardless of the type of book (monograph, novel, textbook, reference). We focused on the "monograph" in its broader sense ("academic" book as opposed to textbook) partly because of the nature of this work stemming from the NMS report, and partly because we anticipated the solutions to the problem being different for monographs/academic books and textbooks
- having access to reliable availability data with regard to which titles are available and in what format is a big challenge for libraries, coupled with information on who owns the rights to any work. Even finding the current publisher (as a proxy for rights holder) is not always straightforward since availability fluctuates
- titles in demand tend to have been published in the last 20 years. Demand from libraries and their patrons for e-books has increased,

yet the largest category of problem was that no digital version was available to the best of libraries' knowledge

- specialist institutions (particularly in the arts and humanities) seem to struggle more to fulfil reading list requests, probably due to the more niche type of publisher with whom these books are published
- even where e-books are available to libraries, they are frequently unsuitable to meet the needs of their patrons, are often too expensive and follow unsatisfying licensing models

Some of the lessons learnt in relation to the NBK are very much tied in with the digital access work. As stated above, trying to distil the complexity of the issues down into manageable and actionable tasks has been a significant challenge but it has become clear over time that it is possible and helpful to pursue the two strands of work in a semi-autonomous way. What is also clear, however, is that the NBK must not lose sight of the goal of better digital access in favour of concentrating on the monumental task of aggregating bibliographic data at the scale proposed. This is a phasing and planning issue for the implementation team but also an oversight and governance priority.

6. Next Steps

We have made substantial progress since the LIBER 2016 conference in taking forward the vision of the National Monographs Strategy roadmap and in response to the two problem statements on *metadata* and *digital access*.

With regard to the digital access strand of the work, given that the great majority of "in demand" titles is strongly rooted in the in-copyright category (in or out of commerce), we are concentrating on exploring possible alternative licensing solutions with publishers that respond to the requirements of the Higher Education sector. At the time of writing, we have moved onto phase 2 of this work and are digging deeper into the findings of the first phase and consulting with publishers and relevant stakeholders in the UK HE library community. Early interim findings point to the complex issue about the need for libraries to have access to reliable availability data on what is available to them as ebooks, as the challenge of discovery, in this context, might create the perception of non-availability of digital copies when in fact they are

available. Secondly, early conversations with publishers also point to the challenge of terminology in differentiating between monographs, academic books and textbooks, as these categorisations are not objective and fixed in time. A more useful criteria to adopt might be simply refer to "reading lists titles" rather than trying to classify them as academic books or textbooks. The second phase is due to terminate by July 2017 and we will disseminate the results as the work progresses.

At the same time, as Jisc and OCLC work together to build and deliver the NBK[17] we will ensure that the bibliographic records aggregated by this new service will, when possible, link to existing digitised copies of books which are available in the public domain such as from repositories like the Hathi Trust.

Building the NBK represents a substantial investment and in the first instance Jisc will work with its funders and its strategic partners to commit the necessary resources to build the system and to establish workflows and processes. During this initial 'build' phase, discussions with the sector as represented by governance and user groups will ensure that the long-term sustainability model for the service is owned and supported by the UK stakeholder community. It is anticipated that a mix of core services and additional 'value-add' tools and service components will be built around the data, providing ways of designing and designating both cost recovery mechanisms and ways of generating income that can be re-invested into the service to keep it relevant and fit-for-purpose over time.

Specifying and procuring an NBK service delivery partner has been a lengthy and intensive process but really just represents the 'end of the beginning.' The practical work to ensure that the UK has the sort of infrastructure and capability originally set out in the National Monograph Strategy now begins in earnest.

References

Bracegirdle, S. (2012). *Books to the ceiling, Books to the sky, My pile of books is a mile high. What do we do with all the unused books in the modern Research Library?* Retrieved June 20, 2017, from https://rlukrrlm.wordpress.com/2012/10/26/unusedbooks/#_ftn1.

Collins, E., & Molloy, C. (2016). *OAPEN-UK final report: A five year study into open access monograph publishing in the humanities and social sciences.* Retrieved June 14, 2017, from http://oapen-uk.jiscebooks.org/files/2016/01/OAPEN-UK-final-report.pdf.

Folan, B., & Grace, C. (2017, May 19). Librarian messages to publishers. *UKSG eNews 396*, n.p. Retrieved June 20, 2017, from http://www.jisc-collections. ac.uk/UKSG/396/Librarian-messages-to-publishers/?n=121b8a3f-e721-4bc0-8354-7ad4d47f99de&utm_source=UKSG+eNews%3A+UKSGMembersOtherC ontacts&utm_campaign=5fb6da73d1-EMAIL_CAMPAIGN_2017_05_17&utm_medium=email&utm_term=0_091ba63bde-5fb6da73d1-302193797.

Freire, N., Scipione, G., Muhr, M., & Juffinger, A. (2013). Supporting rights clearance for digitisation projects with the ARROW service. *LIBER Quarterly*, 22(4), 265–284. Retrieved June 19, 2017, from https://www.liberquarterly.eu/articles/10.18352/lq.8101/. https://doi.org/10.18352/lq.8101.

Hammond, M., Kay, D., Schonfeld, R., & Stephens, O. (2015). *Bibliographic services implications study: Final report.* Retrieved February 15, 2017, from http://repository.jisc.ac.uk/id/eprint/6550.

Jisc (n.d. a). *Bibliographic data services and the national monograph strategy – Next steps.* Retrieved February 15, 2017, from https://monographs.jiscinvolve.org/wp/files/2015/11/BiblioData_NMS_Next_Steps.pdf.

Jisc (n.d. b). *CCM tools project: Use cases.* Retrieved February 15, 2017, from https://ccm.copac.jisc.ac.uk/use-cases.

Malpas, C., & Lavoie, B. (2016). *Strength in numbers: The Research Libraries UK (RLUK) collective collection*. Dublin, Ohio: OCLC Research. Retrieved June 20, 2017, from http://www.oclc.org/content/dam/research/publications/2016/oclcresearch-strength-in-numbers-rluk-collective-collection-2016-a4.pdf.

Nielsen (n.d.). *Nielsen book discovery – Providing comprehensive, enriched and timely bibliographic data.* Retrieved February 15, 2017, from www.nielsenbookdata.co.uk/.

NISO (2014). *Roadmap for the future of bibliographic exchange: summary report.* Retrieved June 20, 2017, from http://www.niso.org/apps/group_public/download.php/13327/NISO_14007BibliographicRoadmapDevelopmentDoc_FINAL4.pdf.

Showers, B. (2014). *A national monograph strategy roadmap*. London: Jisc. Retrieved February 15, 2017, from https://www.jisc.ac.uk/reports/a-national-monograph-strategy-roadmap.

Ward, V., & Colbron, K. (2016). *Digital access solutions – Report on investigations for possible pilot studies.* London: Jisc. Retrieved February 15, 2017, from http://repository.jisc.ac.uk/6562/.

Notes

---

[1] Jisc website—https://www.jisc.ac.uk/.

[2] https://monographs.jiscinvolve.org/wp/expert-advisory-panel-membership/.

[3] It should be noted that various Jisc services (including the proposed National Bibliographic Knowledgebase) could play a significant future role in providing data or intelligence for monograph impact metrics, once a coordinated and coherent community approach has been proposed.

[4] Open Access Publishing in European Networks—http://www.oapen.org/home.

[5] Knowledge Unlatched—http://www.knowledgeunlatched.org/.

[6] Jisc 'Futures' focuses on Innovation and research & development—https://www.jisc.ac.uk/rd/how-we-innovate. The main thrust of the work described in this paper is the responsibility of the Jisc 'Digital Resources' directorate—https://www.jisc.ac.uk/content.

[7] RLUK – Research Libraries UK—http://www.rluk.ac.uk/.

[8] SCONUL – Society of College, National and University Libraries—http://www.sconul.ac.uk/.

[9] A two-page summary of the report including the 'Jisc response' and a 'next steps' section is available at Jisc (n.d. a).

[10] The SCONUL 2015 workshop: Space planning and the re-invention of the library, is instructive about the type of space planning that libraries are having to consider: https://www.sconul.ac.uk/sites/default/files/Spaceplanningandthere-inventionofthelibrary.pdf.

[11] See the final report of phase one of the Digital access pilots (Ward & Colbron, 2016). Its accompanying data set is available at http://repository.jisc.ac.uk/6563/.

[12] http://www.legislation.gov.uk/ukpga/1988/48/chapter/III/crossheading/orphan-works.

[13] See the original Call for participation at https://goo.gl/o8JznT. The ten participating institutions were: Durham University, Royal Conservatoire of Scotland, University of the Arts, University of East London, University of Glasgow, University of Manchester, University of Portsmouth, University of St Andrews, University of Sussex, University of York.

[14] BIBDOG – The Bibliographic Data Oversight Group, convened by Jisc and consisting of representatives from: Research Libraries UK (RLUK); the Society of College, National and University Libraries (SCONUL) and the British Library.

[15] More information on this is contained in the Appendix of the final report, p. 32.

[16] Following an open procurement process, Jisc has appointed Information Power http://www.informationpower.co.uk/ to carry out phase 2 of this study.

[17] See Jisc-OCLC press release at https://www.jisc.ac.uk/news/new-uk-wide-service-will-transform-library-collaboration-03-feb-2017.