

John MacColl*

Towards A Shared Print Collection in UK Research Libraries

<https://doi.org/10.1515/bfp-2017-0048>

Abstract: This paper¹ describes a recent project by Research Libraries UK to analyse the ‘collective collection’ of its member libraries, in order to understand the implications for the community of a collectively-managed print resource in the future. It discusses the work of OCLC Research in using the OCLC WorldCat database for this analysis, taking account of inaccurate data matching and its effects, and considers how the RLUK analysis feeds in to broader work across the UK, led by Jisc, to create a UK National Bibliographic Knowledgebase. It compares the findings of the OCLC Research study to those of an earlier similar analysis of the collective Association of Research Libraries collection in North America. The governance and funding complexity of the UK is described to account for the challenges inherent in taking a national approach to the problems of managing a collective collection. The UK Research Reserve is described as an example of a shared print approach, thus far only applied to journals, which has over the last 10 years been a successful initiative for a number of participant libraries in allowing them to free up shelf space by removing duplicate holdings. The collective collection work of a subset of RLUK, the White Rose University Consortium in Yorkshire, is described as an exemplar of an implementation of the findings of the RLUK-wide study within a regional context.

Keywords: RLUK; Jisc; OCLC; WorldCat; COPAC; Sconul; UK National Bibliographic Knowledgebase; ‘collective collection’; HathiTrust; UK Research Reserve; union catalogues; White Rose University Consortium; GreenGlass; e-books

Auf dem Weg zu einer verteilten Sammlung von Drucken in Forschungsbibliotheken in Großbritannien

Zusammenfassung: Vorliegender Artikel beschreibt ein jüngeres Projekt der Forschungsbibliotheken von Großbri-

tannien, um die ‚kollektive Sammlung‘ ihrer Mitgliedsbibliotheken zu analysieren und die Implikationen einer gemeinschaftlich verteilten Sammlung für die Gemeinschaft der Zukunft verstehen zu können. Die Arbeit von OCLC Research, die die WorldCat-Datenbank der OCLC für ihre Analyse verwendet, wird diskutiert. Dabei werden sowohl fehlerhafter Datenabgleich und als auch dessen Auswirkungen untersucht. Des Weiteren wird im Artikel sorgfältig geprüft, wie die RLUK-Analyse durch Jisc landesweit genutzt wird, um eine ‚National Bibliographical Knowledgebase‘ zu entwerfen. Es werden die Ergebnisse dieser Studie mit einer früheren, ähnlichen Studie über Forschungsbibliotheken in Nordamerika verglichen. Die Komplexität der Regierungs- und Forschungsförderung in Großbritannien wird beschrieben, die Herausforderungen eines nationalen Ansatzes für die ‚kollektive Sammlung‘ werden aufgezeigt und die UKRR wird als ein Beispiel eines gemeinschaftlich verteilten Ansatzes für Drucke beschrieben, dass in den letzten zehn Jahren nur erfolgreich bei Zeitschriften funktionierte, da diese Initiative es den beteiligten Bibliotheken ermöglichte, Zweitexemplare zu entfernen. Überdies wird im regionalen Kontext beispielhaft die Arbeit der ‚kollektiven Sammlung‘ am ‚White Rose University Konsortium‘ in Yorkshire, einem Teilprojekt der wissenschaftlichen Bibliotheken von Großbritannien, bei der Umsetzung der Forschungsergebnisse vorgestellt.

Schlüsselwörter: RLUK; Jisc; OCLC; WorldCat; COPAC; Sconul; UK National Bibliographic Knowledgebase; ‚kollektive Sammlung‘; HathiTrust; UK Research Reserve; union catalogues; White Rose University Consortium; GreenGlass; E-Books

1 Introduction: Research Libraries UK

Research Libraries UK (RLUK) is a membership organisation that consists of 33 of the leading research university libraries, including those of Oxford and Cambridge, the four ancient Scottish universities (St Andrews, Glasgow, Aberdeen and Edinburgh), the English ‘red brick’ universities (Birmingham, Bristol, Leeds, Manchester, Liverpool and Sheffield), and various of the Colleges of the Univer-

1 With thanks to my colleagues Janet Aucock (University of St Andrews Library) and Rosemary Stenson (University of Glasgow Library) for their comments and advice.

*Kontaktperson: John MacColl, jam31@st-andrews.ac.uk

sity of London. It also includes the UK's three national libraries (the British Library, the National Library of Scotland and the National Library of Wales), and – perhaps oddly – Trinity College Dublin (which is a 'legal deposit' library under the terms of UK law). There are also some newer university members, and the Wellcome Trust Library, a specialised biomedical library.

RLUK has for many years maintained its own catalogue (Copac), which aggregates records from its member libraries and some 60 or so other research collections across the UK and Ireland. The service is hosted and managed by Jisc, on behalf of RLUK. It currently contains some 40 m records. As an adjunct service, RLUK has its own record supply database. This record retrieval service provides records in MARC21 format and is open to any non-profit organisation.

2 Evaluation of UK bibliographic infrastructure

Both Copac and the RLUK Database were subject to review in 2015-16. The aims of the review, agreed jointly between RLUK, Jisc, SCONUL (the Society for College, National & University libraries) and the British Library were:

- To help inform shared collection management decisions across RLUK.
- To recognise the implications of an OCLC Research analysis of aggregate bibliographic holdings across RLUK, and understand the preservation fragilities in our collective aggregation.
- To compare the RLUK system-wide aggregation with a recently completed study of ARL holdings, also conducted by OCLC Research.
- To link up with other UK-wide collection management work (Jisc's 'National Monograph Strategy' and the UK Research Reserve).
- To move from piecemeal collection analysis based on patchy holdings metadata in WorldCat, to system-wide analysis.
- To provide a step towards identifying the 'long tail' of publications which may have no formal preservation arrangements, and to consider our collective responsibility for its preservation.

RLUK's report led to Jisc issuing a tender for what it called a 'National Bibliographic Knowledgebase' service in the autumn of 2016.

All of this activity fitted well with the RLUK Strategy – which is currently being refreshed. The previous strategy

(2014-17) has as a theme: *a collective approach: re-shaping the modern research library collection*, and under that the objective to work towards a shared approach to the management of print, manuscript and archive collections across RLUK. A related objective was *coordinate and rationalise digital collections across the UK*.

The fit was also appropriate more generally. Jisc commissioned a study which appeared in 2015 to examine the implications of pursuing specific bibliographic data strategies. This work was undertaken as part of a set of activity within Jisc as it adjusted its focus and shape for the future, in the light of reduced government funding. The report recommended that Jisc should focus on ensuring that bibliographic data services are available and building relevant partnerships, rather than continuing with the relatively labour-intensive task of operating Copac. In its paper summarising the report, it was stated quite boldly that: "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".¹

3 Charge to OCLC Research

RLUK commissioned OCLC Research to use the WorldCat database in order to provide the RLUK community with intelligence on its bibliographic profile. OCLC Research was asked to:

- Explore the characteristics of the aggregate RLUK bibliographic resource in the context of members' strategic priorities.
- Use WorldCat to obtain a multi-scalar perspective on the RLUK profile, at local, group and global levels.
- Complement and advance recent work it had done on the 'collective collection' and shared print.²
- Refine its analysis via consultation with an RLUK Advisory Group that was established to work with the OCLC Research team.

Fundamental to the analyses that OCLC Research would conduct was the currency of the data with which they were working. As a catalogue of catalogues, WorldCat has great potential to be used for bibliographic analysis of this type –

¹ *Bibliographic data services and the National Monographic Strategy – next steps* (2015).

² See for example *Understanding the Collective Collection: Towards a System-wide Perspective on Library Print Collections* (2013) and *Stewardship of the Evolving Scholarly Record: From the Invisible Hand to Conscious Coordination* (2015).

but the picture it provides is only as accurate as the data within it. In the UK, the convention for a long period of time has been that most libraries upload their own catalogue records to WorldCat on a batch basis, which may or may not be regular. They do not catalogue directly into WorldCat, as is more common in the US. This particular study therefore revealed, not surprisingly, that quite a number of RLUK member libraries had not performed batch uploads to WorldCat for some time. The first part of the project therefore involved a number of libraries in undergoing WorldCat ‘reclamations’ – a process which took several months. Once finally complete, however, the analysis could be performed, and it led in May 2016 to publication of the report by Constance Malpas and Brian Lavoie of OCLC Research.³

4 The RLUK Bibliographic Profile

The RLUK collective collection encompasses 29.4 million distinct publications of all types, including 20.9 million distinct print book publications. These titles are represented by 61 million holdings across the RLUK group – an average of two holdings for every title. Scaling this up to WorldCat, the representation of these 29.4 million titles is in the form of 1 billion holdings, which implies that every title held within the RLUK collective collection has an average of 34 holdings across the WorldCat community.

The titles ‘boil down’ to 19 m creative works once subjected to FRBR (Functional Requirements for Bibliographic Records) clustering, which creates work sets of related manifestations. Three quarters of these were represented by a single manifestation.

Print books in RLUK member collections reflect a rich global diversity, with 467 languages and 254 countries of publication represented. Books (all formats) comprise the largest segment of the RLUK collective collection, presenting – in the words of the OCLC ‘Highlights’ section: ‘a significant opportunity for re-imagining library services to both maximise the value of the aggregate resource as a shared asset and improve efficiency in local library operations.’ Books represent 87 % of the total catalogued collection in WorldCat. The remainder of the collection is represented by serials (5 %), musical scores (5 %) and maps, visual resources and other materials (3 %). We might ask why books are so dominant in the RLUK collection. The figure of 87 % of the total catalogued collection represents

71 % of overall distinct works. In contrast, the same study performed on the Association of Research Libraries group only accounted for 49 % of that total collective resource.

The OCLC report explains the variation thus:

“It should be noted that ARL libraries have been registering their holdings in WorldCat for decades and have adopted more comprehensive record-loading practices than are in place for some RLUK libraries. For example, most if not all ARL libraries catalogue rare books, manuscripts and archival collections in WorldCat, whereas only some RLUK libraries do so. Consequently, the RLUK collective collection may appear to be more print book intensive than the ARL collection simply because a less diverse range of library holdings are represented.”⁴

This does indeed seem to be the likely explanation, revealing cultural differences in cataloguing practice across different geographies. The ARL proportions would therefore be likely to be more representative of the contours of an average research library collection (irrespective of its country) than are those of the RLUK study. However, for the purposes of assessing the collective monographic collection of the UK – its degrees of overlap and rareness – the approach is nonetheless likely to be a very reliable guide.

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. Ninety per cent of the collection was deemed scarce or distinctive at the RLUK scale. At the WorldCat scale, naturally that figure will reduce – to 60 %. These figures might imply that the UK scale is not robust enough to guarantee preservation of all distinct copies for the UK’s collective collection just within the UK, and that full preservation will require the backup of the ARL collection, to increase the resilience of the aggregate research collection. But moving from these inferences to putting mechanisms in place to maximise the preservation capacity of an aggregate research collection is a big task, requiring collaborative infrastructure at the human, system-wide and data levels that does not currently exist.

Nearly 460.000 distinct subjects are associated with the print book publications in the RLUK collective collection, with a variety of particular subject strengths distributed across the RLUK membership. The authors note that collective collection of a group of institutions is rarely subsumed within a subset of member collections. Full coverage of the collective collection requires full in-group participation in the stewardship effort.

The RLUK collective collection is both similar to and distinct from the collective collection of the Association of

³ *Strength in Numbers: The Research Libraries UK (RLUK) Collective Collection.*

⁴ Cf. *Ibid.* 4.

Research Libraries (ARL). What at first sight seems like a significant portion (42 %) of the RLUK print book collection overlaps with the ARL collection, but the fact that 58 % therefore does not might seem surprising given that we are comparing research library aggregations in two Anglophone countries.

5 The HathiTrust Digital Library

One of the interesting analyses done by the report was the comparison of the RLUK collective collection against the HathiTrust e-book database. Here again, the information generated by this comparison can indicate preservation possibilities – and also point to opportunities for enrichment of local library catalogues by linking records for print resources to available e-book versions in HathiTrust. This analysis revealed that the median duplication for titles in the aggregate RLUK collection with the HathiTrust database was 25 % (compared to 34 % for the ARL collection). This figure is likely to rise over time, as Hathi adds more and more content from e-book digitisation projects across the world.

Nonetheless, these figures reflect the duplication in the print content across libraries, and only 13 % of the print book publications held in the total RLUK collection is available in Hathi. However, the question of copyright is an important one in any consideration of use of the Hathi database to substitute for print equivalents in research collections. At the time of the comparison, Hathi contained 6.8 m e-book titles, of which only some 2.2 m (32 %) were public domain titles in the US. Of the total RLUK content duplicated in Hathi, 78 % of it is not available due to copyright – unsurprisingly, since 60 % of the collective RLUK content has been published since 1950. The figure of 78 % prompts the authors of the OCLC report to comment that there is ‘ample scope’ for the UK to do a lot more book digitisation on a concerted basis.

This is undoubtedly true, but to make effective use of Hathi the UK would also have to commit to working with the Trust, and inputting e-book content – an activity which has been noticeably lacking in the UK. We might at this point raise the question of willingness within the UK to be involved in international-scale collaborations generally, thinking back to the earlier observation about the way in which the UK has tended to use WorldCat for cataloguing over the years, doing so indirectly in comparison to the normal North American approach. And of course, this is not a phenomenon unique to the UK. What is perhaps being revealed is a reluctance to engage in international-scale collaboration that is considered too US-centric, and

we may need to reflect upon that fact, and how we address it if it is preventing us making the most of our collective resources. The fact that the UK’s new National Bibliographic Knowledgebase (discussed further below) will harvest public domain records from Hathi is very welcome, and should stimulate the use of this major international aggregation by libraries in the UK.

A key question in managing collections as a shared resource is whether the planning perspective should be restricted to group membership only, or if it should be elevated to higher scales, recognising that the group collective collection is embedded in a broader system of collective collections.⁵

6 Practical value of the analysis

The reflections upon this analysis up to this point have occasioned speculation about how the RLUK community could take steps to make any repeat of this analysis in future more valuable – by, for example, having a higher share of the total content catalogued and uploaded to WorldCat, by participating substantially in HathiTrust, or by engaging in a collaborative e-book digitisation effort. We would urge RLUK to consider each of these challenges in its strategic planning for the years ahead. But we should also ask about the value of the data that has been generated for immediate use. Is it helpful to RLUK or the UK more widely? Will it lead to any immediate activity?

To answer that question, we need to consider the overall bibliographic context of the UK. In numerical terms, RLUK represents only a small – if important – fraction of the total number of libraries in the country – some 37 institutions. All members of RLUK are also members of Sconul, the Society of College, National and University Libraries – which has 175 member libraries. Also significant, of course, is the British Library – a member of both RLUK and Sconul – with its huge collection of printed books in WorldCat. Indeed, the UK has three national libraries (the British Library, National Library of Scotland and National Library of Wales) and these – together with the libraries of the University of Oxford, the University of Cambridge, and Trinity College Dublin – make up the six copyright (or ‘legal deposit’) libraries of the UK. The RLUK collection therefore, which encompasses all of these, may reasonably claim to represent a UK union catalogue, and

⁵ See <http://www.oclc.org/research/publications/2016/oclcresearch-strength-in-numbers-rluk-2016.html>.

so to be a platform for strategic development of library services across the country.

The difficulty, however, in moving forward with a national library development agenda, is in the very fact of there being several different loci of responsibility to different user constituencies. We have the national libraries: the British Library and the 'home nation' national libraries; we have the set of copyright libraries; we have the academic libraries – represented by RLUK and Sconul; and then we have Jisc, the digital services and solutions provider funded by the UK government's higher and further education and skills sector. In addition to these, we have the UK's most successful venture into collaborative collection management to date, the UK Research Reserve.

7 UK Research Reserve

The UK Research Reserve (UKRR) has become a prominent part of the collections landscape for universities in the last few years, and represents an example of active collaboration for collection management. It is a partnership between the British Library and 29 UK Higher Education libraries, which started in 2007 and – to date – has been focused on print journals. Key to its operation is coordinated de-duplication, which gives those responsible for managing collections in the 29 member libraries a validating framework for decision-making surrounding print de-accession, and an assurance for academics and the whole research community that print copies are secured for the future. RLUK member libraries find that academic staff – at least in the Humanities and Social Sciences – can be highly resistant to the suggestion that books should be discarded as the libraries find themselves challenged to provide more space either for books or – more commonly – for study space. One of the most valuable roles that UKRR has played has been to provide a trusted depository environment for academic library print journal content UK-wide, thus reassuring academic colleagues that the material that may no longer be present on the shelves of their university library is nonetheless still available somewhere within the system. While their acceptance of this may at best be grudging, given a natural preference to have everything ever owned by their library (and quite likely recommended by their colleagues or predecessors) close to hand, it does usually permit them to agree to the removal of physical stock. UKRR is a member of EPICo,⁶ the European Print Initiatives Collaboration.

⁶ See <http://www.varastokirjasto.fi/epico/>.

Through UKRR's de-duplication process, abundant material is discarded and scarce material is identified and retained in a shared national collection, distributed among the member libraries. Since 2007, UKRR has freed up more than 104 kilometres of space occupied by print journals on behalf of its members. Formerly led by the British Library, the UKRR is now led by the Library of Imperial College London. Its funding base is not a permanent one, however, which has limited its ability to move on from offering a very welcome service of print journal downsizing, to tackling the much more complex challenge of monographs. However, it has recently surveyed its members, and looks likely to be moving forward within the next couple of years with a service offer extended across the UK, and a determination in due course to tackle the question of collective management of monographs.

Across these different elements there is no single authoritative body, and so solutions for our users of bibliographic data (which means practically every person in the country) rely upon collaboration and agreement, and are often represented by short-term projects and initiatives, not least because permanent funding is very hard to achieve since each element is funded for its own constituency.

8 The White Rose libraries

Faced with the strategic difficulty of taking coordinated action across the UK on the basis of the OCLC report, RLUK adopted as its initial response the approach of letting one of its regional sub-groups, the White Rose Consortium, act as an exemplar to the rest of RLUK of how to move forward with a shared print project. The White Rose Consortium is made up of the three Yorkshire universities in the UK's elite 'mission group' of universities, the Russell Group⁷ – Leeds, Sheffield and York.

The White Rose Group conducted their own consortial analysis, based upon the OCLC project data. This produced some interesting findings. The University of Sheffield Library discovered, for example, that 65 % of its titles were held uniquely; 24 % were held by Sheffield and one of the other two libraries; and 11 % were held by all three libraries. They were also able to ascertain that 11 % of Shef-

⁷ See <http://russellgroup.ac.uk/>. The Russell Group's 24 universities are all members of RLUK, which includes in addition several universities from the 'small research-intensive' group that formerly had their own 'mission group', the 1994 Group, which was wound up in 2013.

field's collection was unique within England (which is of course not the same as being unique within the UK – but the 'England' analysis did include the British Library).

What the Group then did was to drill down further within the data, to gauge how accurate these assumptions were in fact. OCLC had used its GreenGlass tool to establish duplicate matching, based upon 'same edition' exact WorldCat number matching. However, in the case of UK OCLC member libraries, which historically have been less likely to source their records directly from WorldCat, different WorldCat numbers can be generated for local records, and this can then lead to an assumed non-match when a match should have been identified, thereby giving a falsely high figure for uniqueness.

When libraries catalogue directly onto WorldCat, matching is a more manual process with human input from a cataloguer, and there is intellectual input into matching local library holdings onto OCLC master records using the key matching points of identifying standard numbers (ISBN and OCLC numbers) and the bibliographic data itself. It is likely that as much care is taken and priority given to accurate matching to create local holdings as would be used to aggregate identical copies onto a record in the contributor's own library database. But when libraries batch-load their holdings into WorldCat, they rely on automated matching which inevitably must throw up mismatches. Standard numbers and algorithms do some of the work, but there will be issues where contributed records do not contain OCLC record numbers, or they contain the wrong OCLC record numbers where contributors have not kept up to date with OCLC record number changes, or have in fact simply in error retained the wrong numbers. Even when algorithmic matching works, contributor libraries do not update their own local databases with the correct OCLC numbers, usually because they cannot resource this, or achieve it technically or do not perceive its value. Among cataloguers in the UK there is a perception that the matching is problematic, which induces a culture in which we do the best we can but do not generally prioritise achieving higher matching rates. In addition, while we use OCLC as a key source of bibliographic records to download, we do not necessarily think about how our holdings are recorded back into WorldCat. We cannibalise records and re-edit them to suit our local requirements – so that, despite cataloguing standards, libraries can take very different approaches to defining what is an edition, a printing, a paperback or hardback edition or how they distinguish e-books provided by different aggregators. Provider-neutral guidelines for the production of e-book records are used in WorldCat for master records, but they do not always fit local catalogue discovery and access

requirements, which we often wish to customise to provide better and more specific services for our users.

This question of false matching has therefore raised a question mark over the reliability of the findings, and the extent to which it has rendered them inaccurate is still being established. It is unfortunate, since in other respects the cataloguing teams within the White Rose libraries were very pleased with GreenGlass' dataset manipulation tools, and the ease of cross-institution analysis. However, until White Rose has completed its analyses, in conjunction with colleagues from other RLUK libraries who were involved in the overall study, RLUK as a whole is not ready to move forward with any consortium-wide programme of action for shared print and preservation.

This problem with data matching has exposed vulnerability in the WorldCat data, created over a period of 50 years by the worldwide OCLC member community principally for the purposes of providing shared catalogue records. The use of the data for cooperative collection management was not foreseen at the outset, and therefore inevitably different communities have behaved differently from each other in the way that they have interacted with WorldCat. Using WorldCat indirectly rather than directly for cataloguing, as the UK has done, has resulted in the unintended consequence that WorldCat numbers cannot be reliably used for duplicate matching. It may be that OCLC will be able to apply further pattern-matching algorithms to compensate for this deficiency, but for now that is unclear. The further use of the updated WorldCat by the RLUK Group for the purpose of cooperative collection management will depend on progress being made in this area.

9 National Bibliographic Knowledgebase

At the same time as the RLUK study was being undertaken by OCLC Research, Jisc was developing plans for a 'National Monograph Strategy' – in conjunction with RLUK, Sconul and the British Library. This was motivated in part by its own strategic review and a decision to contract out the maintenance of the Copac database, which it has managed jointly with RLUK and its predecessor body, CURL (Consortium of University and Research Libraries), since its launch in 1996. It was announced in February 2017 that the contract to run the National Bibliographic Knowledgebase (NBK) on behalf of Jisc has been won by OCLC, with the launch of a beta service projected for January 2018. It will form part of an intention to create a 'national digital

library' – as yet a somewhat nebulous idea that has been touted for several years by both Sconul and Jisc.

In basic terms, the NBK will replace Copac as the *de facto* union catalogue of the UK. The data underpinning the Knowledgebase will rely – on the first instance – on libraries of all types allowing their bibliographic and holdings data to be aggregated by OCLC into a single system that is capable of working at greater scale and ingesting more diverse library data at much faster rates than is currently the case with Copac.

A key consideration for Jisc and its partners in developing the NBK is that the UK's 'monograph landscape' is becoming more complex as it is increasingly populated with e-book content – and as a consequence, the NBK is likely to contain a more diverse range of content metadata than does Copac, and to be fed from a wider range of sources. Neil Grindley, Jisc's Head of Resource Discovery, makes the point that this upgraded national bibliographic infrastructure is necessary because 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. They want the NBK to be a single point of access to e-books for library users.

The NBK may provide some of the solutions we seek to our problems in using a shared aggregation for reliable information on duplication. It is promising that it will be created anew from data loads from contributor libraries. The launch of a brand new catalogue service for the UK at a time when shared collection management has entered our professional consciousness should encourage cataloguers and their managers in UK libraries to become more involved, see vested interests for collections management, and regard the accurate maintenance of the NBK to be a priority for institutions.

10 Conclusions

The pressure on research library space across the UK means that, from an RLUK perspective, shared print solutions are going to be required for many years to come. Many of our libraries are facing the issue of trading off space for books against space for students, and as a result – even in large library buildings – research libraries with physical collections of typically well above one million volumes, and growing – are having to build or rent new storage facilities. Shared physical infrastructure, such as is on offer to the community from the British Library via the coordination of the UK Research Reserve, is necessary – and we require the shared intelligence to determine what

goes in it. Strategically, collaboration on this activity is obviously desirable, but even if viewed only pragmatically, with so many of us making these decisions at the same time, it clearly makes sense for us to take a shared approach.

The OCLC Research report commissioned by RLUK represents a very useful input into this activity. It seems clear that there are valid concerns about some of its statistics, given the problems of duplicate matching that have been unearthed by the White Rose libraries. These reveal the limitations of the WorldCat data aggregation, created originally for one principal purpose – cataloguing efficiency – and now being considered as a potential source for a new requirement – collaborative collection management. We must hope that algorithmic solutions can be found that will refine the data to make it usable, since the report has stimulated significant interest in the potential that exists for the UK to move to this next stage in library collaboration.

In speculating on why the UK took a different route in its use of WorldCat from most US research libraries, we have considered whether there was a resistance to an approach that seemed too US-centric. If that was indeed the case originally, it seems clear now that WorldCat's value is global, irrespective of the headquarters and history of OCLC as its parent, and we should perhaps take a lesson from this experience: where resources are created that have the potential to achieve international collaborative benefit, we should adopt standardised behaviours in our usage of them from the outset.

Decisions need to be taken collectively on how many physical copies of titles we need within the UK collective collection, and then how those are disposed: how many can be permitted to be in circulating collections (where attrition rates are measurable and can generate ratios that we must use), and how many need to be in closed, reserve collections? Do we distinguish the statuses of libraries with preservation responsibilities? For example, UKRR is proposing a system in which the British Library has a primary preservation role, supported by a few 'Principal Holding Libraries' which will ingest items from non-UKRR libraries; then a group of 'Holding Libraries' which will retain some of their own collections on behalf of the UK as a whole; and then a hinterland of non-UKRR 'Contributing Libraries' who will dispose of material only through the UKRR network.

How do we factor in surrogate versions that exist in digital form, as e-books? Do these reduce the physical quantity required by the system as a whole? And indeed, if they do, shouldn't we be pursuing a national strategy to create more of them from within our own collections?

All of these questions require answers, and will occupy our community as strategic challenges over the next few years. Alongside them are questions of appropriate scale. RLUK is happy to let a three-member consortium (the White Rose Partnership) forge a path as an exemplar for the time being, but it does not seem likely that a network so small is at a scale that represents efficient use of national resource. Should we therefore be operating at UK level? The 175 libraries of Sconul would represent a substantial administrative challenge of coordination. RLUK, with its 37 members, is surely more achievable?

However, we have the issue of fractured accountability in the UK: any leadership effort requires that we corral together the governors and funders of RLUK, the British Library, the national libraries, the copyright libraries, UKRR and Jisc. Despite the many collaborative initiatives that have seen subsets of these partners work together successfully, there is no overarching authoritative agency (as exists, for example, in other countries where universities are more significantly funded by government) to require coordination. It is for this reason, perhaps, that RLUK in empowering its regional groups such as White Rose, has been more inclined to 'act down' than to 'act up'. It is to be hoped that the arrival of the National Bibliographic Knowledgebase will provide us with a collectively maintained service whose value to the UK for collective collection management will lead to a form of governance that is considered authoritative by all parties.

In addition to fractured accountability along existing funding lines, the UK is facing increasing tension politically, with devolved governments in Scotland, Wales and Northern Ireland all acquiring enhanced fiscal powers and increased responsibilities devolved from the UK parlia-

ment in London. In a political landscape which the vote to leave the European Union in June 2016 has now rendered highly volatile and unpredictable, the break-up of the United Kingdom itself is a possibility within the next few years.

Research libraries remain stubbornly cooperative, and it would seem likely that any break-up of the UK would still leave the countries of England, Scotland, Wales and Northern Ireland cooperating with each other over library infrastructure, services and data, as has continued with the Republic of Ireland over the past century. But the current volatility creates an unfavourable climate for initiatives to establish coordinating centres for the sort of bibliographic rationalisation that we require. We face the problem of the appropriate scale at which to take collaborative action – and then to translate that decision into cooperative infrastructure across regional, national and international boundaries which are currently in flux. Nonetheless, the questions cannot be dodged, and we require as a research library community to work together more closely than ever at the level of dialogue and diplomacy, to understand our shared needs and to create our own international forums, from which collaborative systems will emerge which will gain enduring acceptance.



John MacColl
University Librarian & Director of Library Services
University of St Andrews
GB-St Andrews
KY16 9TR
Großbritannien
jam31@st-andrews.ac.uk