

Copac Collection Management Tools Project

Synthesis of the case studies

July 2012

Summary

In Phase 1 of the Project, Leeds, Sheffield and York university libraries had set out a series of use cases which illustrated how the Copac Tools might be used. One of the deliverables for Phase 2 was, therefore, to test a number of these scenarios in real life and develop them into case studies which documented both process and results. Joining later in the project, four “associate libraries” – those of University College London, the Victoria and Albert Museum and the universities of Manchester and Warwick - were given access to the Tools, principally to assist with the testing of the User Interface. Even though time was limited, they also offered to develop use cases or case studies based on their own experience which was of great value in developing a broader understanding of how the Tools might be used in practice.

The 10 case studies focussed on three main areas of activity: stock withdrawal, collection profiling and decision-making in conservation and digitisation. They addressed the following specific situations or issues:

- During stock editing of low-use titles, to avoid discard of those which are unique nationally or only held by a small number of other libraries. (*Sheffield*)
- Enhancing understanding of the collections through collection profiling. (*York*)
- Profiling collections in order to identify those of potentially national significance (“heritage collections”). (*Leeds 1*)
- Supporting decisions in conservation and digitisation. (*Leeds 2*)
- Stock withdrawal – a detailed consideration of unique and rare items. (*Leeds 3*)
- Developing an integrated and more cost effective approach to assessing material in a store where catalogued and uncatalogued items are interfiled. (*Manchester 1*)
- Reducing the staff time required to assess stock for withdrawal. (*Manchester 2*)
- Use of the tools for assessment of collections in non-Roman scripts. (*V&A*)
- Developing stock editing procedures which allow the uniqueness of material to be taken into account and thereby make the process more acceptable to academics. (*Warwick*)
- Support for collaborative projects and bidding for funding in the context of specialist collections. (*UCL*)

The key points and messages that emerge from this work are as follows:

1. The underlying data and data structures within Copac are all-important to the effectiveness of the tools in a practical setting. The case studies have been very important in enhancing our understanding of the issues surrounding duplicate records in the database and how the current database can only reliably return results based upon ISBN. The new Copac database, due for release later this year, is set to address these issues and improve the reliability of results returned based upon control number queries. It will be important to verify this by re-running some of the case study operations using the Tools once the new database goes live.
2. The case studies succeeded in demonstrating how the Copac Tools can be integrated into the processes of an academic/research library and how they can deliver added value and efficiency across a broad range of collection management activities.
3. By using the Tools and by integrating them into automated workflows, it is possible to save significant staff time, aiding decisions, and making possible tasks that previously were not considered viable. These workflows, that have already been demonstrated and tested within the Project, can be used as exemplars for the benefit of new users of the Tools.
4. Even with the current data limitations, all the libraries involved in the project want to continue to have access to and actively use the Tools after the current project. The value experienced by incorporating the Tools into their workflows has convincingly demonstrated the benefits of having access to the Copac Tools.
5. Help and advice is essential to ensure that the tools are used effectively and that results can be interpreted appropriately. The Copac Collection Management Tools website requires enhancing with on-line help and examples which outline best practice and can inform user expectations of the Tools.

A synopsis of all 10 case studies is given as the final section of this document and full reports on the work undertaken at Leeds, Sheffield and York can be found via the project website at <http://copac.ac.uk/innovations/collections-management/>

Understanding the data

One common theme runs across nearly all of the work undertaken, namely the difficulties caused by the presence of duplicate records for the same bibliographic item within the Copac database. This is very widespread and arises from the history of the database, variations in quality of the records and the limitations of the current software platform. Copac will migrate to a new platform in the second half of 2012 and it is expected that this will eliminate the majority of duplicates for material published post 1800. Once this change takes place, it will then be possible to use a local record number search (derived from the library management system of one of the holding libraries) to identify all holdings of the title in question. Up to present, it has only been possible to reliably identify all holdings of a title by using an ISBN search (which provides a mechanism capable of bringing together multiple records for the same title). In consequence, reliable results have only been available for titles where an ISBN is recorded locally i.e. titles published from the 1960's onwards (and even after that, many titles were issued without an ISBN). It also proved surprisingly

challenging to extract from the MARC records in the Library Management System a “clean” list of ISBNs to submit to the Copac Tools.

Where a control number search is used, this may or may not identify the majority of the holding libraries. For such searches:

- If 8 holdings, for example, are identified, then we can be confident that at least 8 libraries hold that item. This may be sufficient information to permit withdrawal as part of a stock editing programme.
- However, if fewer than 8 holdings are identified, other records may exist within Copac which would bring the total number of holdings above 8.
- If a title is identified as unique to your library, this may not be the case: it is more often because your record in Copac has not been merged with those from other libraries, for whatever reason.

This caused a number of issues in the case studies:

- Many of the processes (e.g. stock withdrawal, collection profiling) are based on a mixture of more recent and older titles. Libraries needed to either pre-select material which had ISBNs (cf. Leeds’ approach to collection profiling) or accept that results would be incomplete (cf. Sheffield’s stock withdrawal exercise). Both these approaches prevented a full exploration of the value and limitations of the Copac Tools.
- This limitation was not initially recognised. Considerable effort was therefore expended in comparing results for searches by ISBN and record number for the same set of titles (cf. York, Warwick).
- The associate libraries in particular found this limitation very difficult to understand as they had no prior experience of using the Tools. This severely restricted their ability to complete the work they had originally intended, especially as they mostly wanted to work with older materials.
- De-duplication options were introduced, partly in response to the need to merge duplicate records after an ISBN search. However selection of an appropriate de-duplication option introduced a degree of complexity which further hindered understanding of how best to use the Tools (cf. V&A).

Key themes and results

Despite these fairly fundamental difficulties, the case studies did succeed in demonstrating how the Copac Tools can be integrated into the processes of an academic/research library and how they can deliver added value and efficiency across a broad range of collection management activities. This section summarises key work undertaken in each of the three main areas of activity to which they were applied, namely stock withdrawal, collection profiling and decision-making in conservation and digitisation. It then draws out the factors which seem to be required for successful exploitation of the Tools’ potential, and the key benefits identified by staff as arising from their use.

Stock withdrawal

According to participants in the Project workshop in December 2011, this is the single most important benefit to be gained from using the tools, in that those titles which are rare or unique nationally can be identified and retained. The case study from Sheffield, in particular, demonstrates an automated workflow which saves substantially on staff resource, with data derived from and flowing back into the library management system. Manchester too was able to document in detail the cost savings achieved through use of the tools. Both York and Warwick discuss how engagement with academics can potentially be improved through this process. Leeds documented some of the issues which arise once unique titles have been identified while Manchester employed the tool to assist with stock assessment in the somewhat different context of reviewing material in a basement store that included quantities of uncatalogued titles.

Collection profiling

Leeds and York both exploited the Copac Tools to deepen their understanding of their collections, profiling individual subject collections and with the long-term aim of developing a comprehensive categorisation to support prioritisation and decision-making. Leeds worked with over 70 separate collections and sub-collections and their case study tentatively suggests a framework for identifying those of national significance. York focussed on just one collection and combined data from various sources including the Copac Tools in order to present a comprehensive profile which could be shared with staff in the relevant academic department. UCL and the V&A considered collection profiling as a tool which can assist with stock management, identify opportunities for collaboration with other institutions and potentially underpin bids for funding around specialist collections.

Conservation and digitisation

Leeds undertook a small case study in this area. Though very restricted because of the search limitations discussed above, it does demonstrate how Copac holdings can be used to inform real-life decisions in an operational context and how integration with the local LMS can facilitate certain parts of the workflow.

Conditions for Success

Across all these areas, we can discern common themes. In terms of the conditions for success, these include:

- A good understanding of the tools themselves, the limitations of what they offer and the data flow which will yield the desired outputs.
- Integration with existing workflows and, if possible, the technical ability to establish smooth data flows from and back into the local library management system.
- A range of appropriate and interlocking policy documents to guide decisions once raw data has been gathered through use of the Copac Tools.
- An understanding that the data generated by the Copac Tools can be of value in informing decisions, but only if used appropriately.

Benefits

In terms of the benefits which accrue, the following have been either realised in practice or identified as likely if appropriate systems can be set in place:

- More efficient workflows, leading to significant cost savings.

- Improved understanding of the collections, their strengths and how they fit within the national context of holdings.
- Better use of space as a consequence of better informed stock retention decisions and more efficient processes. Many libraries are currently looking to consolidate space as part of a wider programme of service development.
- Better information to support prioritisation and decision-making across all areas connected with collection management, including conservation, digitisation and bids for external funding.
- Evidence (especially statistical data) to support advocacy across the wider institution.
- Greater buy-in from users/academic departments, particularly around withdrawals and space management generally.
- Examining uniqueness of collections with a view to supporting bids for funding.
- Identifying areas of overlap with other Copac library holdings with potential collaborative ventures in mind.

Synopsis of each case study

University of Sheffield

Issue addressed. During stock editing of low-use titles, to avoid discard of those which are unique nationally, or only held by a small number of other libraries.

Summary. A largely automated workflow was set up which exported record numbers and ISBNs from Sheffield's library management system (LMS), ran the data through the Copac Tool and identified those for which fewer than 8 holding libraries were recorded in Copac. Data for the remainder was then loaded back into the LMS and lists generated for withdrawal of these titles. 2500 books were actually withdrawn from an initial list of 6836 titles with a saving of at least 60% on staff time.

Problems identified.

- Care is required in selecting an appropriate de-duplication option within the Copac Tool.

Key benefits.

- Efficient workflow for stock withdrawal.
- Risk removed of discarding material which is rare or unique nationally.
- Integration with the LMS.
- Significant saving of staff time.

University of York

Issue addressed. Enhancing understanding of the collections through collection profiling.

Summary. The Copac Tool was used, along with other data such as circulation statistics derived from the LMS, to present a profile of the collection of monographs on Language and Linguistic Science. The resulting profile benefits from including not only statistics about usage, age profile, etc., but also how the collection at York fits within the national context. This forms part of a wider programme of collection categorisation which, it is hoped, will improve decision making in areas such as space management, stock editing, digitisation, conservation and bids for external funding.

Problems identified.

- Lack of full catalogue records for some York titles led to them being identified as unique when in fact they were not.

Key benefits.

- Ability to create a profile of the collection which describes it in both its local and national context.
- Strong quantitative basis (which also allows the report to be strong on visual representation of data) to increase awareness among library staff and in academic departments.
- Improved basis for prioritisation and decision-making across a broad range of library activities, including stock withdrawal.

University of Leeds 1

Issue addressed. Profiling collections in order to identify those of potentially national significance (“heritage collections”).

Summary. A number of subject collections were selected for analysis, broken down by local classmark. Subjects were analysed at both discipline and sub-discipline levels. For each such collection, ISBNs were submitted to the Copac Tool and various key metrics tabulated in order to build a profile for each collection. Because the searching was limited to ISBN, it was not possible to fully profile some of the more interesting collections in the arts and humanities, although some data was derived for recent acquisitions (i.e. those with ISBNs). A fairly consistent pattern emerged for most collections, allowing some tentative suggestions regarding how a heritage collection might be identified from such data.

Problems identified.

- The Copac Tool cannot take account of material which is not recorded in Copac. In consequence, results may sometimes overstate the rarity of items.
- Similar issues may be caused through poor quality records.
- Collections with a high percentage of literary works may generate profiles which are different from those in other subject areas because of differing patterns of publication and acquisition. Further work is required in this area.

Key benefits.

- Provides an objective and statistical basis for categorising certain collections as “heritage”, rather than relying on subjective and sometimes ill-founded opinion.

- Improved basis for prioritisation and decision-making across a broad range of library activities, including stock withdrawal.

University of Leeds 2

Issue addressed. Supporting decisions in conservation and digitisation.

Summary. This small case study looks at the decision-making process associated with just 3 items which were in poor condition and therefore being reviewed for repair, digitisation or withdrawal from stock. As part of the study, a mechanism was implemented which allowed a staff member to view a title in the LMS and then click through to a display of the same title in the Copac Tools interface, so that decisions could be based on the combined evidence of local data (e.g. usage) and the broader context of national holdings.

Problems identified.

- One of the items was a multi-volume work. Each volume may be catalogued separately or on one record as part of the set. Different Copac contributors have adopted divergent approaches, resulting in multiple records, which makes it more difficult to establish how many holdings there are.
- There was also some difficulty in separating out the records for later reprint editions.

Key benefits.

- Ability to make appropriate decisions regarding conservation and digitisation, taking into account the rarity of an item.
- Targeting of staff effort on items where there is maximum benefit.
- Potential for integration within the local LMS.

University of Leeds 3

Issue addressed. Stock withdrawal – a detailed consideration of unique and rare items.

Summary. A small number of titles in the subject area of nursing had been identified for potential withdrawal, based on low use. A list of ISBNs was extracted from the LMS and submitted to the Copac Tool. Surprisingly, nearly a quarter of these had fewer than 5 holdings recorded (including Leeds) and 4 were unique to Leeds. A detailed analysis was carried out on these (including checks in WorldCat) to validate the results and explore the nature of the material but no clear patterns emerged beyond a number of titles coming from one publisher.

Problems identified.

The following were not so much problems as questions and issues which emerged from the exercise.

- Why did so many “rare” titles come from the one publisher, Springhouse?
- For modern commercial publications, how many holdings nationally are required? Particularly for books published in the United States, should we also take account of copies in American libraries, and is it enough to have a single copy in the UK?

- Might these titles be in other libraries which do not contribute to Copac (many nursing libraries would be in this category)?
- For scientific works, is it important to retain multiple copies of each edition within the UK?
- How willing will libraries be to retain modern titles which are no longer relevant to their needs? And would they continue to house them as part of the standard subject collection, where their inclusion could be misleading to users?
- If libraries using the Copac Tool do identify “unique” titles which they might otherwise have discarded, they may need to spend some time on investigating those titles and relocating them for permanent retention.

Key benefits.

- Without the Copac Tool, it is unlikely that these titles would have been individually checked against Copac since there was no thought that any might be rare. 4 titles apparently unique within the UK would then have been lost.

University of Manchester 1

Issue addressed. A basement store contains a large amount of material with catalogued and uncatalogued material interfiled. The aim was to test a method of assessing the unbarcoded material for either retention or withdrawal at the shelf using the Copac tool. If the book is to be retained a record will be imported into LMS *in situ*. It was hoped that use of the Copac Tools would allow an integrated and more cost effective approach to assessing this material.

Summary. A list of all the material shelved in that store and recorded in the local LMS was generated in spreadsheet format. A cataloguer checked the actual shelf material against the spreadsheet using a laptop. Any items found that were not on the spreadsheet were then checked using the web-based Copac search to assess the number of copies at other Copac libraries. Any non-English script/pre-1900/multivolume titles were excluded with details being noted on the spreadsheet for future possible cataloguing work. If there were more than 8 copies at other libraries they were put to one side for potential withdrawal (after being checked over by a relevant subject specialist). If fewer than 8 copies were found a suitable record was imported onto the LMS, adding a barcode to the item and the item details onto the imported record.

Problems identified. No major problems.

Key benefits.

- Uncatalogued books can be efficiently assessed at the shelf for retention/disposal with an imported record being immediately brought in where applicable. This provides both time and cost savings.
- Uncatalogued books are revealed for future cataloguing e.g. potential Special Collections material or foreign language material; helping to produce a full inventory of the library collections and potential areas for prioritisation for Collection Management activity and resources

- This process identifies and corrects errors of catalogued material e.g. incorrect Shelfmark or missing/lost items on the LMS.

University of Manchester 2

Issue addressed. The University of Manchester Library stores contain large amounts of catalogued material that has never been included in the previous library weeding projects. This case study examined how the Copac Tools might reduce the staff time required to assess this material for withdrawal.

Summary. A list of all the material shelved in that store and recorded in the local LMS was generated. Items not used in the past 10 years had their record numbers submitted to the Copac Tool and the results fed back into the master spreadsheet: those with more than 8 holdings nationally were flagged for withdrawal. For this case study the 300s classification was chosen containing 22,906 items, of which 19,006 items (83%) had not been loaned for 10 years or more. The Copac tool identified 3,729 books to be withdrawn; 16% of the total items. It was estimated it would take 4 staff approximately 4-5 days to process these withdrawn items. By comparison, if all 19,006 items were manually checked it would take the same amount of staff 7 weeks to complete the process. That checking would include 15,277 items that would eventually be retained, representing approximately 6 weeks of the total work (86% of the manual checking time). The results from the Copac tool also showed that on average the items identified to be withdrawn were at 15 other Copac libraries.

Problems identified.

- The process does not identify or include any uncatalogued material within the store collections.
- Local record numbers were used instead of ISBNs as the majority of the store stock was pre-1970 material.

Key benefits.

- Substantial staff time and cost savings in checking titles for 8 or more holdings in Copac when comparing manual and batch checking.
- Efficient workflow for stock withdrawal.
- Risk removed of discarding material which is rare or unique nationally.
- Due to the time savings more stock from different collections can be assessed and withdrawn freeing more space in the library stores.

University of Warwick

Issue addressed. Current stock editing procedures are based only on usage. Use of the Copac Tools would allow the uniqueness of material to be taken into account and hopefully result in a process which is more acceptable to academics.

Summary. Various collections were identified for analysis, mainly with a view to stock withdrawals, and record numbers were submitted to the Copac Tool. In some cases, it was possible to also submit ISBNs for the same titles. The results differed greatly between the

two search types, mainly because of the known issue around duplicate records for the same bibliographic title in the current Copac database. However, the potential value of using the Copac Tool as an integral part of collection management was identified. More specifically, by incorporating evidence about the uniqueness of titles into the policy and procedures around stock editing, and by developing a complementary retention/preservation policy, it should be easier to gain approval from academic departments.

Problems identified.

- Complexities in extracting clean ISBN data from the LMS.
- If a library makes substantial use of the Copac Tool (e.g. as part of the stock withdrawal process), there is a need to have retention and preservation policies in place to underpin and guide the associated processes.

Key benefits.

- Better information to guide stock withdrawal.
- Because the Copac Tool allows additional data beyond mere usage to be brought into play in making stock withdrawal decisions, it is potentially easier to win support from academic departments.

National Art Library (V&A)

Issue addressed. To assess collections in the Chinese, Japanese and Korean languages.

Summary. The National Art Library holds discrete collections of Chinese, Korean and Japanese language material, covering all subjects within its collecting scope of art, decorative art and design. This material occupies a large portion of storage space close to the reading rooms but is perceived to be under-used by curatorial staff and external researchers. The intention was to gauge its significance in relation to other available resources, within London and throughout the UK. Because of a lack of technical support, it was only possible to work with a sample of titles. Of these, only a minority carried ISBNs, limiting the ability to gain meaningful results.

Problems identified.

- Issues of compatibility with certain browser versions.
- Difficulties of extracting data from the LMS without specialist support.
- The option in the Tools for “detailed de-duplication” of records seemed to have little effect.
- A broader query emerged as to whether the algorithms for merging duplicate records as part of the Copac migration to a new platform will deal effectively with records for E. Asian materials, given the specific issues involved.

Key benefits.

- Demonstrates the potential for the Copac Tools to assist with evaluation and management of collections in non-Roman scripts

University College London

Issue addressed. Work at UCL focussed mainly on testing the user interface to the Copac Tool and exploring potential uses for specialist collections e.g. material in Yiddish.

Summary. Sample searches were tried from a number of different collections but work was hampered by various practical difficulties, particularly for titles without ISBNs (e.g. most works in the Yiddish collection). In consequence, no firm conclusions could be drawn from the results returned.

Key benefits. Having had some experience of the Copac Tool, staff identified the following areas of activity at UCL where its use would be of particular value:

- Examining uniqueness of collections with a view to supporting a bid for funding.
- Identifying areas of overlap with other Copac library holdings, with potential collaborative ventures in mind.
- Informing stock retention decisions in the light of major plans for consolidation of space and service development within UCL Library Services.