Workshop on Retention and Preservation, March 12th 2012: Notes arising from the afternoon plenary session

Introduction
At the start of the workshop Michael Emly had given a presentation providing background to the project and then set out the objectives for the workshop:

- Understand what information you would like to have access to, and how
- Understand what information you would be prepared to contribute to Copac
- Identify key drivers for Copac contributors and other partners
- Identify benefits for individual libraries
- Identify barriers to implementation
- Think about mechanisms for collaboration
- Begin to think about specifications, formats, workflows and systems

The final plenary session of the day aimed to highlight the discussions of the earlier group breakout sessions and to pull together common themes and points of agreement, ideally identifying practical steps forward, such as:

- A framework of objectives
- Agreement for a common vocabulary
- Stimulate and feed into a broader discussion - leading to a national framework

Common Themes from group discussions
The workshop had split into two groups, A and B, to discuss and consider five pre-set questions (appended to these notes). The plenary session started with the rapporteurs from the groups presenting summaries of the main points arising from the break out groups. (These are available as separate ppt documents.)

Michael Emly led the discussion on identifying key common themes emerging from the groups. These split into three areas:

1. Standards
   This was in the context of allowing confidence and quality assurance so that libraries can interoperate in a common format. It is important to have a common understanding of semantics.

2. Unique and distinctive collections
   The benefits of the idea of “collaborating to compete” were addressed and there was a consensus that these two actions were not mutually exclusive but indeed complementary. The point was raised that libraries have different relationships and roles, for example: as part of an institution; as part of a consortia; a regional grouping; a specialism; etc. Therefore, librarians have to cut decisions in different ways dependent upon the focus at the time. To this end it would be beneficial if the CCM Tools could be reconfigured on the fly to inform the type of decisions that have to be made enabling librarians to interpret information flexibly addressing decisions in appropriate ways.

3. Prioritising information most useful about items held in other libraries
   Each group had been given an initial list of 5 suggestions and then asked to supplement it and then prioritise. A common pattern emerged with the following chosen as the top four:
   i. Intention to retain for the long term.
   ii. Whether an item is damaged or in poor condition. (This conflates 2 of the suggestions reflecting discussion about the definitions of the terms “damaged” v “poor” condition.
iii. Whether an item has been digitised in-house (or is to be digitised). Some discussion about whether “in-house” was relevant as items may be digitised elsewhere.

iv. Whether an item is an e-book and if it is available for ILL.

**Question and Answer session**

The observations above led to further discussion that started with a lively exchange on the topic of e-books. It was noted that there had been a SCONUL collaborative tools project looking into “e”-resource management and at a recent workshop at Warwick University there had been a recommendation for co-development to work towards a definitive source of information for “e”-books in print. This recommendation has fed into a recent JISC call for projects.

The point was made that users are format agnostic and that both “e” and print are managed together. So why not include the electronic material in the retention and preservation decision processes? Not all “e” books are equal. There is a temptation for libraries not to record them. For example, in the case of “Patron Driven Acquisitions (PDAs)” that may be accessible for a relatively short time.

A major concern is how we manage existing print collections. They occupy space so are finite in a way that ‘e’ isn’t. Also, print deteriorates and is therefore at risk. Is digitisation the answer? It can be but access to the physical object is often valued and the digital version just doesn’t satisfy the user.

With regard to preservation and retention question it was put forward that the key driver is access to resources. On this basis it would be possible to use the existence of ‘e’ books to inform disposal and retention decisions. It was noted that ‘e’ book management is diversifying from traditional LMS. Is this a potential problem? RDA (Resource Description and Access) is on its way with the LC going live in March and BL following (at a date currently not known). However, concerns about LMS being able to deal with anything other than flat files were commented upon. The benefits of FRBR are not able to be realised by current LMSs.

Licensing conditions presenting problems of access to ‘e’ books was noted. Some caution was also expressed relating the experience gained through ‘e’ Journal licenses where subscribing libraries

1 A report from the US “Education Advisory Board” comments:

“The Advisory Board report, a thick primer covering a range of trends in digital librarianship, predicts a shift in the way academic libraries provide book content to their patrons that mirrors a broader trend in digital media. (The report is not public.) Academic libraries will jettison “large collections of physical books in open stacks with low circulation,” the report says, in favor of licensing agreements with e-book vendors that will enable libraries to purchase only those books that are highest in demand, while paying short-term access fees for books that students use a little and nothing at all for books they do not use.”


Inside Higher Ed

2 http://www.jisc.ac.uk/fundingopportunities/funding_calls/2011/04/roleofmetadata.aspx

Note that I have found it difficult to locate specific projects. Please let me know of any relevant references.

3 “The primary distinction between RDA and AACR is structural. RDA is organised based on the Functional Requirements for Bibliographic Records (FRBR). These principles identify both the ‘user tasks’ which a library catalogue should make possible and a hierarchy of relationships in bibliographic data. Descriptions produced using the instructions of RDA are intended to be compatible with the large number of existing records created under the rules of AACR2.” From http://en.wikipedia.org/wiki/Resource_Description_and_Access
have found that titles can be removed from collections by a third party provider without notice resulting in entries in the catalogue that are no longer valid.

Some analysis previously done within the White Rose Consortium had pointed to the possibility of establishing a minimum requirement for how many copies needed to be available for the purposes of preservation. Eight copies have been considered, but the question arose as to whether digital copies should be factored in as well. In response to this it was thought that for national collections you should be trying to keep print copies. The quality of the physical copy and its condition all will feed into decisions about retention.

An interesting point raised was that of approaches to decisions re space. For example, the UKRR experience showed that when responding to space pressure decisions were made based on collection level considerations. It was felt that it is not so different for monographs and that people would make decisions about retention at collection level and not at item level.

It was agreed that institutional buy-in for conserving material was required as well as for the disposal of material.

Where Next?
Consideration was given to what benefits and what barriers there are to applying a collective approach to retention and preservation and to having CMM Tools that can be applied with confidence.

Benefits
- Space saving and cost saving
- Can assess the quality of a collection in one location
- “Collection” becomes coherent and systematic
- Efficiency
- Strategic planning
- User friendly (managed) shelves – this relates to findings that if you can “de-clutter” shelves making it easier for students to find volumes that are relevant that there will be an increase in issues to students.
- [Reporting circulation data via Copac-SALT]?
- Identify unique and distinctive collections.

Barriers
- Lack of standards – or perhaps lack of agreement on standards. Can standards be enforced? RLUK role?
- Labour cost (Noted that if a library submits data to a union catalogue then they will receive error reports. Currently these reports are not necessarily acted upon and the fact that effectively 3rd party data checking has been provided is wasted. Could RLUK encourage? Agreement that quality data is very important to the effectiveness of the CCM Tools.)
- Librarians not wanting to dispose of material and therefore not motivated.
- Cost benefit analysis for scenarios to support community action is needed. Evidence to support business cases to HEIs locally. (Noted some work in this area – Life Cycle Costs for housing print, Michigan University4.)
- Kudos of the size of collections and the fit to strategy.
- Potential bad publicity - popular perception.

4 Following quote “Paul Courant, (currently Dean of Libraries at the University of Michigan), calculates the annual life cycle cost of retaining books in campus stacks at $4.26 compared to an estimated $0.86 to store them in an off-site facility. Among the Ivies, at least 20% of print collections are housed in high-density off-site facilities, with Columbia and Cornell approaching 40% and Harvard with 50% of its collections off-site, including most newly received acquisitions.” taken from: http://www.library.upenn.edu/docs/publications/ivyleaves/11springivyleaves.pdf
What can RLUK do to help?
1. Put in place a framework agreement for digitisation to address licensing.
2. Standards – sign off and agree these standards.
3. Needs buy-in at a senior level so that the standards can be applied by practitioners.
4. Tell people to use the 583 field. How to use and use in Copac. “Let’s do it!” The BL 583 standards exist – we can agree and move on. Consult with the community and RLUK can then recommend.
5. Standards for digital preservation policy - a template? Some European projects have started this. Starting point would be the dissemination of best practice.
6. Inter-lending between RLUK members would aid decisions about weeding giving confidence about sharing material.

What can the workshop participants do?
1. Look at existing workflows – how they can change these to fit with the existing standards and tools.
2. Consider establishing a registry of UK Digital Masters? Or join the OCLC effort?5
3. Flag collections for disposal so that others can see if they would want the collection – checking it against your libraries holding in the CCM Tools.
4. Look at how retention and preservation works in your institution. Talk to colleagues more widely and get engaged in internal dialogue. See how much common interest in questions there is.
5. Also, feed ideas up to senior management and get them engaged so they will feed these concerns into strategy (both local and national) and support work in your institution.
6. Consider what other organisations (perhaps local to your own) have collections which would usefully be recorded in Copac in order to to increase its coverage. Copac is continuing to expand but suggestions are required about which institutions to approach.

Julia Chruszcz
16 March 2012

5 http://www.oclc.org/digitalregistry/default.htm
For today, please leave to one side questions of defining “long-term retention”. This will be picked up later.

1. What information would you like to be able to access regarding items held in other libraries?
   Various suggestions were made at the December workshop including:
   a) Intention to retain for the long-term
   b) Whether an item is in poor condition
   c) Whether an item is damaged
   d) Whether an item can be borrowed via InterLibrary Loan
   e) Whether an item has been digitised in-house (or is to be digitised)
   You might like to consider other information which is explicitly included in the existing vocabulary for preservation and digitization actions.
   Which categories of information are most useful? Please make a ranked list.

2. Now go back through the list which resulted from question 1.
   a) Discuss how many of these categories are recorded, explicitly or implicitly, within your own institution’s Library Management System (to the best of your knowledge!).
   b) If the information is held external to that system, but could potentially be linked to it e.g. via a shared control number, then please note that also.
   c) If the information isn’t currently available, how feasible would it be to include it? Be realistic! What are the obstacles?

3. Some additional questions:
   a. Would it be helpful if Copac automatically flagged up unique titles to the holding library?
   b. The Copac Tools will allow a Library to view holdings within a consortium or ad-hoc list of libraries – how useful is this in making a decision about items you hold? And what grouping makes sense in your context?
   c. Pre-1800 items have generally been viewed as outside the scope of this exercise because, being hand-printed, each one is unique. Is it valid to exclude them in this planning?

4. It has been suggested that by libraries identifying specific collection strengths, others may:
   i. Choose not to collect in depth in those areas
   ii. Offer unwanted but rare material to an appropriate library that collects in that subject area
   This leads to a number of questions:
   a. Does the model seem valid?
   b. Do you think your library might be interested in “advertising” particular strengths and/or disposing of material in this way?
   c. How might this work in practice?

5. In the past, collaborative collection management projects such as CoFoR put in place mechanisms for stock transfers. In the future the creation of digital surrogates may provide an alternative approach with additional benefits.
   a. What is the role of each, and how could the Copac Tools support either strand of activity?
   b. Which should be prioritised?