

Retention and preservation: revisiting the CCM Phase 2 recommendations on metadata

Michael Emly, January 2016

1. Introduction

The reports from Phase 2 of the CCM Project (2012)¹ presented a coherent approach for collaborative management of older printed monographs held in the UK. Since that time, further work has been undertaken to provide tools based on the Copac database to underpin this approach, as well as engagement with the wider community to build an understanding and consensus around the associated issues. Towards the end of 2015, Jisc committed to further development as part of the National Monograph Strategy, overseen by the new Bibliographic Data Oversight Group² while RLUK has signed an agreement with OCLC regarding joint work on shared print management³. Thanks in large part to research and advocacy on the part of OCLC⁴, there have also been significant developments in North America since 2012, including the emergence of consortial agreements which seek to actively address the issues relating to print monographs⁵.

The documents prepared for CCM Phase 2 included a report on Retention and Preservation⁶, with very specific recommendations regarding appropriate metadata standards. As work starts on a new phase of activity in the UK, it seems important to revisit the recommendations regarding metadata in the light of these more recent developments so that appropriate standards can be built into the specifications for the “Monobase”. This will also give a better understanding of the issues involved for any preliminary development and trials undertaken by the Copac team or by individual libraries before the main specification is complete.

The 2012 recommendations addressed two main questions:

- How should a participating library record its intention to retain a given print monograph?
- How might a participating library record the physical condition of any such items?

The current document revisits both these questions and recommends next steps for the CCM Board.

¹ <https://ccm.copac.jisc.ac.uk/archive/archive-reports/>

² <http://monographs.jiscinvolve.org/wp/2015/11/20/designing-solutions/>

³ <http://www.rluk.ac.uk/about-us/blog/rluk-collective-collection-in-context-the-oclc-research-perspective/>

⁴ <http://www.oclc.org/research/themes/systemwide-library.html>

⁵ Cf. EAST <https://blc.org/east-project> and work by the Hathi Trust

<https://www.hathitrust.org/files/sharedprintreport.pdf>

⁶ <http://blog.ccm.copac.ac.uk/wp-content/uploads/sites/13/2012/01/Retention-and-PreservationFinal.pdf>

2. Recording intention to retain

The 2012 report made very concrete recommendations on the use of MARC tag 583 for this purpose. Since then, developments undertaken under the aegis of the OCLC Shared Print Management program⁷ have confirmed the basic soundness of that approach. However the working documents published by the Maine Shared Collections Cooperative⁸ around shared management of print monographs highlight the fact that recording data locally in a standard format is not sufficient in itself; it is also important that this data can be stored, managed and exploited by central systems supporting the consortial activity, in that instance OCLC. It is particularly noteworthy in a CCM context that this has raised questions regarding both the type of record in which the 583 tag is stored and also the precise institutional identification code used (in subfield 5).

Given that the OCLC metadata standards⁹ have become a de facto standard in North America, the current agreement between RLUK and OCLC for joint working on collection analysis and the desirability of interoperability between the UK and OCLC systems moving forward, it seems advisable to gain input from OCLC on these 2012 recommendations before proceeding – something which Jisc staff seem currently well placed to do. To facilitate this, Appendix A from the 2012 report (“Examples of 583 content”) has been slightly revised and included as Appendix 1 below: this could form the basis for such a dialogue. Any discussion with OCLC should include (but not be limited to) the following questions:

1. The record type in which the 583 is placed. Given that a number of different systems are in use by UK libraries, the bibliographic record would seem the preferable option, at least for monographs. Any other proposal would require consultation within the UK community to establish its practicability. This would necessitate inclusion of the institutional symbol in subfield 5 to indicate that the tag contents refer to the local holding rather than the title in general.
2. Can existing Institutional symbols be used in 583 subfield 5? Members of North American consortia seem to be using a separate symbol in this context.
3. Subfield f should indicate the consortium/agreement concerned. The Bibliographic Data Oversight Group and the UK community more generally do need urgently to address this question (see recommendation S4 on p.15 of the main 2012¹⁰ report which focussed on the need for a formal framework underpinning shared print management of monographs in the UK) and agree a symbol to be used e.g. UKNMS (for UK National Monograph Strategy).
4. Current OCLC guidelines define subfield d as containing the date the retention requirement expires and state this is a required field. However this does not align with the stated use in the MARC format definition: “Time period which cannot be expressed as a specific date”. Is OCLC’s current requirement really practicable given that such agreements are usually time-limited but would also expect to be renewed? Insisting on subfield d as a required field would also preclude any input by UK libraries until a formal agreement is in place defining that date.

⁷ <https://www.oclc.org/services/projects/shared-print-management.en.html>

⁸ http://www.maineinonet.org/mscs/wp-content/uploads/Teaching-document-Version-1_0.pdf

⁹ <https://www.oclc.org/services/projects/shared-print-management/metadata-guidelines.en.html>

¹⁰ <http://blog.ccm.copac.ac.uk/wp-content/uploads/sites/13/2012/01/Final-CCM2-Tools-Project-Report-010820121.pdf>

5. OCLC seems to recommend aligning the contents of the 583 tag with the PDA standard, highlighted by the inclusion of subfield 2 “pda” in all their examples. However the “committed to retain” statement in subfield a still does not match any of the options permitted by the published standard. We have already noted that use of subfield d also diverges from both PDA and from the standard MARC definition. Given that this is a very distinctive and specific use of 583, it might be that the link with PDA should be broken and a new very simple “standard” published to support shared print management. Conformity with the published standard for MARC tag 583 will ensure broader interoperability with PDA.
6. Can any subfields be automatically supplied when records are loaded into either the “Copac” database or into OCLC in order to simplify the requirements on individual libraries (which might also improve data accuracy)? On the other hand, this might not be a major issue for participating libraries if such data is normally added to multiple records simultaneously through a block change facility in the local system. Again, discussion within the UK community would also be helpful.

3. Recording condition of the item

This was a second area explored in detail by the 2012 report on Retention and Preservation and there seems no reason to change the basic recommendation to follow PDA and to record preservation condition, etc. in MARC tag 583, precisely as documented in both those standards. This also conforms to the OCLC recommendations.

Since 2012, a few UK libraries have begun to do this, and one issue distinctive to the UK has emerged. Any assessment of the condition of an item is time-consuming and expensive, and therefore a relatively rare event. It often occurs within the context of a survey of a representative sample of the collections which will, in the UK, often use the methodology of the former NPO’s Preservation Assessment Survey. This assigns a numerical score from 1 to 4 to represent the condition of each item. Unfortunately, PDA does not offer a way of recording this distinctive coded score except within a general note field, thereby removing the option of reliably undertaking subsequent analysis across the collection. The alternative, which is valid in MARC terms but does not conform to PDA, would be to use 583 subfield l (status) and record a suitable code to reflect this (e.g. UKNPO) in subfield 2 instead of “pda”.

Adhering to PDA would ensure consistency with the international community while use of subfield l for the condition would potentially allow cross-institutional machine analysis of preservation condition within the UK. In practice, it seems unlikely that a significant number of records will carry any such information in the next 5-10 years (especially as support for the methodology is no longer offered by the British Library), so it is probably best to ensure conformity with PDA, as per the original 2012 recommendation. A formal recommendation on this subject could usefully form part of the formal framework for the NMS.

Appendix 1: Examples of 583 content

Level 1. In line with the OCLC Pilot, the 2012 Appendix suggested the possibility of a very basic 583 tag with only one subfield. For monographs, this would need to be unambiguously associated with a given institution within the shared system. This might be achieved by adding an institutional symbol in subfield \$5 at point of loading into Copac or OCLC if one is not already present. However, in the current environment, this minimal level of input is probably not required by any institution and should not be an option.

583 1 \$a Committed to retain

Minimal level of input

Level 2. A recommended set of subfields for monographs might consist of:

\$a Action e.g. Committed to retain

\$c Date e.g. 2012

\$f Authorization e.g. UKNMS [UK National Monograph Strategy]

\$2 Source of term e.g. pda [but see discussion in section 2, question 5 above – perhaps omit?]

\$5 Institution code e.g. UkleU

583 1 \$a Committed to retain \$c 2012 \$f UKNMS \$2 pda \$5 UkleU

This is compliant with relevant standards, albeit with a minimal level of information.

Level 3. Where a library desires to record additional information, the full range of MARC subfields and PDA terms can be used. This information might either relate to the intention to retain i.e. be expressed within the same 583; or alternatively relate to additional information about the item held i.e. be expressed in one or more subsequent 583 tags. These 2 possibilities are shown in the two following sets of examples.

583 1 \$3 v.1 only \$a Committed to retain \$c 2012 \$d 2022 \$f UKNMS \$f SCURL \$2 pda \$5 UkleU

In this case, the Library has committed to retain vol. 1 only, for a 10 year period, within 2 separate agreements: UKNMS and SCURL

583 1 \$a Committed to retain \$c 2012 \$f UKNMS \$2 pda \$5 UkleU

583 1 \$a Condition reviewed \$c 2012 \$l mutilated \$z pages 9-15 damaged \$2 pda \$5 UkleU

583 1 \$a Housed \$c 2012 \$i box \$2 pda \$5 UkleU

In this case, the Library has committed to retain the item. On reviewing the condition, it has noted that the item is mutilated and has also decided that it needs a storage enclosure (to be housed in a box).