
A catalogue quality audit tool

Ann Chapman and Owen Massey

The authors

Ann Chapman is Research Officer for Bibliographic Management at UKOLN, which is based at the University of Bath. Owen Massey has recently finished a recataloguing project at the Markland Library, Royal College of Obstetricians and Gynaecologists.

The Occasion

Owen Massey was the winner of the LIRG Postgraduate Prize 2001 for his Masters Dissertation at Loughborough University. The dissertation was co-supervised by Ann Chapman of UKOLN. This paper is parallel published in LIRN by kind permission of the Editor of Library Management.

Abstract

The current need for performance measurement and quality targets for services to users requires suitable performance indicators for libraries to use. This paper looks at the self-assessment audit tool for catalogue quality developed by UKOLN in collaboration with Essex libraries. For the tool a checklist of errors was drawn up, which can then be used to assess the quality of records within a catalogue using a sample of library stock. The tool can be used to assess the quality of catalogue records for monographs and non-book materials (but not serials), for complete collections or parts of collections and for records created at different periods. This paper describes the tool and the process of making the assessment and reports on the results of the pilot study carried out at the University of Bath Library in 2000.

Introduction

In recent years, library services have increasingly been encouraged to be more accountable to those who use and fund them, leading to the introduction of performance measurement and quality targets for services to users. One such service is the provision of a catalogue – but just how can you measure the quality of your catalogue?

The overall quality of a catalogue depends on a combination of factors, defined in 1980 in initial UKOLN research proposals as accuracy, consistency, timeliness and functionality. Each of these factors needs an objectively assessed performance indicator, chosen to conform to some standard of quality such as the six criteria proposed by Orr (1973): it must have informative content, allow comparison, be reliable, valid, accurate and practical.

This study concentrates on the accuracy of bibliographic records and the presence of 'dirty data'. In all performance measurement, there is a risk of measuring only what can be easily measured rather than what is most important. Nonetheless, mistakes are inevitable in databases the size of library catalogues, and without an idea of their

prevalence it is impossible to gauge the success of efforts to reduce errors, let alone answer more complex questions about the catalogue's effectiveness.

Background

Essex County Libraries approached UKOLN early in 1997 for information on methodologies for measuring catalogue quality. Initial work indicated that nothing suitable existed for use 'off the shelf' and UKOLN and Essex decided to collaborate in developing a self-assessment tool for libraries. However, despite early progress in designing the audit, Essex was unable to undertake the pilot study as planned.

Approaches to other library services suggested much interest in the tool but none had sufficient staffing resources to commit to performing the pilot study. The project was offered as a dissertation topic at Loughborough University library school resulting in its being piloted at the University of Bath library (Massey, 2000). The aim of the pilot was less to discover the incidence of errors in the catalogue than to investigate whether the right errors are being measured, while evaluating the practicality of the proposed technique.

Development of the audit tool

There is a profusion of literature on what makes good catalogues, or at least good catalogue records, but fewer studies have attempted to measure quality. These often concentrate on before-and-after studies of changing cataloguing systems or cleaning up a database without necessarily providing a method that can be generalised.

However, the stock and catalogue audit at the University of Tennessee (Kiger & Wise, 1996) provided a starting point for the new quality assessment tool. The audit used two samples: one sample checked items against the catalogue, measuring the percentage of missing records, and the other sample checked catalogue records against items, measuring the percentage of missing items. The use of attribute sampling seemed over-complex for simple self-assessment and although it was stated that the audit 'necessarily entailed a verification of the accuracy of the catalog record', this aspect was not reported in any detail.

The new tool (UKOLN, 2002) was designed with the same two tests, catalogue-to-collection and collection-to-catalogue. To measure the accuracy of bibliographic records, UKOLN had developed a checklist divided by type of error and bibliographic field. The choice of errors and fields to inspect was difficult. There was little consensus in the literature, and the variety of methodologies used reduced the validity of generalisations. Errors are particularly significant in controlled fields used as access points, while readers' consistent preferences for keyword searching require accuracy throughout all fields, quite apart from the need for confident identification from a retrieved record. Hence all bibliographic areas were treated uniformly in the pilot.

The desire for an objective performance indicator led to greater emphasis being placed on mechanical accuracy in transcription than on subjective intellectual accuracy. Errors have been divided into (a) fields containing incorrect information and (b) omitted fields, rather

than distinguishing misspellings, MARC coding errors and deviations from cataloguing rules.

A collection of juvenile fiction and a research-level university collection are likely to be catalogued intentionally to very different standards. Therefore, in this tool, catalogue quality evaluates deviations from the library's own standard, which may not coincide with an external standard such as AACR2. This requires a profile of the library's catalogue, describing the nature of the collection, past cataloguing regimes and reclassifications and salient local policies in order to account for known idiosyncrasies. The intended use of the results should be considered before auditing because this may affect the detail and quantity of data needed and the method of collection. The audit is intended to survey the whole catalogue but it can easily be adapted to concentrate on part of the collection, provided a suitable sample can be identified.

The tolerable error rate should be also decided before auditing. If an error rate over (say) 6% triggers a recataloguing exercise, sequential analysis can save time (DiCarlo & Maxfield, 1988) by halting the audit when there is enough evidence for a decision. If the concern is to correct errors without measuring them, this can be done more efficiently by inspection, browsing indexes and searching for likely errors such as those collected by Ballard (2001).

The audit technique

Having drawn up the profile and decided on error limitations, the next step is to print a random sample of catalogue records, if possible in shelf order to save time locating items. Errors are recorded either on the printout or on a worksheet for each item (reproduced as Appendix 1). The error coding given in the worksheet can be used simply for a tally chart, but the value of the audit is increased by recording errors for later correction. While minimal information identifies items, the full bibliographic record, in MARC format if this is used, is more useful for annotation.

The two tests are interlinked. For each record in the sample, find the corresponding item A on the shelves and also the item B five places to its left. Using the worksheet, check the cataloguing of item A. If it cannot be found and is not on loan, search for it on a second occasion; if it is on loan then recall it for checking. Next, search the catalogue for item B by control number and other access points; note if no record is found or if more than one relevant record is found.

Totals can be recorded in the worksheet table or the data input into a spreadsheet or database. The number of uncatalogued items found can simply be totalled and given with estimates of the number of duplicate records and missing items. The principal figure to calculate is the proportion of records with at least one error, either a keyboarding error (whether spelling or typing), incomplete data, or an omitted field. Other statistics such as the mean number of errors per record, or per incorrect record, may be derived as desired. All statistics should be quoted with their margin of error.

Not all keyboarding errors or missing fields have the same effect on the utility of the catalogue. The initial design of the checklist ignored this, but for the pilot, errors were divided into major (those that affect access) and minor (access not affected), with major errors taken as mis-keyings in any field (other than capitals, punctuation or diacritics), omitted words in titles and series, missing author headings and series, and incorrect classmarks. This level of detail requires the full bibliographic records to be printed.

Practical issues

Any member of staff can locate and recall items but comparison of items with records and confirmation of duplication requires professional expertise, even though value judgements have been kept to a minimum to ensure consistency. If suitable staff are not available at all sites, decide whether it is easier to move staff or stock. Allow time for the printing and distribution of the sample, the tests themselves, and the return of recalled items. It may be more efficient to locate several items at once and send them for verification elsewhere rather than checking them individually.

Academic libraries can perform the test in the summer vacation but other libraries have fewer quiet periods and might need more recalls. The sample will not be significantly affected by items being borrowed or returned so there is no need to close for the audit. Even additions and alterations to the catalogue are unlikely to make a major difference, although strictly speaking the results will be valid only for the date on which the sample was drawn.

The catalogue-to-collection test for accuracy

Using the catalogue profile, the checklist can be fine-tuned by adding and omitting fields. Miscellaneous issues such as absence of a uniform title or inappropriate subject headings should not be recorded as errors unless specifically requested, but should be noted for later correction. Multiple errors in one field should be recorded as a single error so that the incidence of errors in that field is not exaggerated. Decide at the outset on the status of aberrant spacing, punctuation, capitalisation, diacritics and the spelling out of numbers and symbols such as ampersands.

The worksheet contains the checklist in table format with rows listing the areas of the record to be checked and columns for the error types. A control number should be recorded for each item, being a unique identifier such as an accession number (ISBNs are unsuitable as there may be multiple copies of an item).

The areas of the record to be checked are: title, material description, statement of responsibility, author heading(s), edition, physical description, imprint, series, classmark/shelfmark, subject headings, genre/category, and location (or branch). This checklist was based on ISBD areas so should be straightforward to adapt for non-book materials.

Notes fields have been excluded despite the substantial number of errors they may contain in case their diversity, multiplicity and length introduce enough errors to distort

overall results. However, individual libraries may wish to check specific notes fields, such as contents notes in music.

The checklist gives textual descriptions of fields rather than their MARC codings that may be unfamiliar to some staff or indeed not used. If MARC coding is used then additional checks of tags, indicators and subfields may be desirable, especially if data has been converted between different MARC flavours. There is no attempt to verify the information in fixed fields, as manual checking is laborious and the contents are often validated automatically. Where it is of interest to readers it is normally duplicated elsewhere in the record.

The collection-to-catalogue test for completeness

The percentage of items with (or without) catalogue records is measured by checking a random sample of items from the shelves – five to left of each item in the catalogue-to-collection sample. If auditing a sub-population interfiled with other items, such as videocassettes with books, the rule becomes ‘first appropriate item to the left’. Inaccuracies in the records found can be noted for correction but are irrelevant to this test.

Using items on the shelves excludes those on loan, in processing, in use by readers or awaiting shelving. Consequently completeness could be somewhat underestimated (material on loan may be more likely to have records) or marginally overestimated (for multiple copies, items five positions apart may be the same work, already known to be catalogued).

This test incidentally provides a rudimentary search for duplicate records, though confirmation of duplication requires care, especially when multiple copies are held at different sites.

Sampling

Sampling minimises the amount of checking and allows decisions to be made on incomplete information while quantifying and justifying the inaccuracy of any statements. The set of records to be evaluated is called the population, usually including all catalogue records.

Two common techniques of random sampling are systematic samples and stratified samples. In a systematic sample, records are chosen from the population at a fixed interval. This is effective unless some repeated structure in the population leads to distortion by omitting or disproportionately selecting certain materials. A systematic sample can be generated from sequential control numbering, from the shelves or by classmark.

A stratified sample divides the population into sub-populations each of which is separately sampled. As long as there is no overlap, it is valid to pool the results to produce overall statistics. Possible divisions include audience (adult/children), format

(print/audiovisual), genre (fiction, grey literature), subject (local history) or cataloguing history (retrospective conversion, or created before or after a certain date).

Random sampling may be impractical, necessitating a convenience sample (such as a predetermined number of books catalogued since a certain date) or a purposive sample (examining only books known or suspected to be inaccurately catalogued); inferences about the whole population from such samples are invalid.

Multi-site libraries must decide whether to check an item at the site chosen at random or to check a copy held at a convenient location. Records for items on order should be removed from the population, but not by ignoring records with no copies attached, as such 'ghost' records (other than orders) frustrate readers and must be considered errors. It is convenient to remove records for items away for binding or repair or marked as missing, although this information itself may be inaccurate; this will reduce the sample size but should not skew results unless such items typically have different characteristics from items on the shelves.

A sample may include items from a sub-population that cannot be screened out beforehand. In order to obtain a sufficient sample it may be necessary to use substitutes when items outside the sub-population are drawn. The ideal procedure is to select another at random; it is not acceptable to substitute the next appropriate item on the shelf, because this may introduce a systematic bias.

Sample size

The number of records to audit depends on the acceptable margin of error in the result, but it does not depend on the size of the population (provided fewer than 10% of records are to be sampled, which is almost always the case). Ready-reckoner tables are available (Kiger & Wise, 1993); for a margin of error of 4.9%, 3.5% or 2.5%, choose a sample of 400, 800 or 1500 items, respectively. If the error rate in the catalogue can be estimated, however roughly, then it is often possible to justify smaller samples using a formula such as that provided by Hernon (1994).

Increasing the sample size improves the estimate, but it cannot validate a biased sampling procedure. Also, attempting to halve the sample size by combining the results of the audit's two tests would invalidate the results by doubling the effect of any bias in the original sample.

Constraints of time or staffing or other practical considerations may limit the number of sample items to around 300 or fewer. In this case, the sample size should be fixed in advance and the margin of error calculated subsequent to the audit.

The pilot study

The Library & Learning Centre at the University of Bath has approximately 400,000 books and 2,000 serial subscriptions, supporting study in a range of disciplines with a

bias towards the sciences. Universal Decimal Classification is used with some local modifications. There are small but growing non-book collections.

The computerised catalogue originated in an experiment on the effectiveness of short records. On the principle of 'minimal data, maximum access', records held only title proper, surnames and initials, date, edition, class and book numbers (Bryant, Venner & Line, 1972). Subtitles were included only when necessary to distinguish items, edition statements were abbreviated and all accents and apostrophes were omitted. Although not displayed, records included those few ISBNs borne at this time and the language of the item if not English. The only subject access was by classification number. While these records are of significantly lower quality than the rest of the catalogue, their accuracy must be evaluated in their own terms.

At the start of the 1980s, the library joined the SWALCAP (South West Academic Libraries Co-operative Automation Project) consortium, when these short-entry records were converted to UKMARC format. New records were created in-house or copied from SWALCAP records. In 1985 SWALCAP's system was replaced with URICA and all subsequent cataloguing has been performed in-house. During this migration all 248 fields were lost, although some have been reinstated, and in the 260 field, copyright dates 'c1980' were corrupted to 's1980'.

Changing to the Unicorn system in 1998, the creation date of existing records was reset to the date of the changeover. The University's catalogue is now shared with that of the University of the West of England (UWE), although difficulties in merging holdings records have prevented a union catalogue of serials.

Consequently there are three types of record in the catalogue. First, there are the original minimal records. Secondly, there are standard records created since 1982 to AACR description level 2 but without physical description or subject headings. Finally, records created by UWE include pagination and, for non-fiction, in-house subject headings; items held by both Bath and UWE have these enhanced records.

Records are upgraded and revised if and only if further copies of the item are bought or classmarks are changed, but the scale of acquisitions since 1982 means that minimal records constitute a minority of the catalogue. The proportion of each record type could not be calculated, as records hold no distinguishing data. Many personal name headings, and all those in minimal records, were generated directly from the statement of responsibility. Checks were only possible for presence or absence rather compliance to AACR2.

The omission of diacritics was not considered an error as they are usually neglected in transcription. The language field 041 has been included inconsistently so was not checked. Similarly, personal names occasionally appearing as subjects in the 600 field were ignored. The correct construction of classmarks was not evaluated. Each item in multivolume sets receives an individual record, risking repeating errors.

The first (convenience) sample

Generating a true random sample from the cataloguing system proved impossible for technical reasons. Instead, a convenience sample was taken consisting of 305 unique records for all books issued at Bath over one spring weekend. A dozen theses, short loan offprints and course packs were excluded from the sample. Each item's full MARC record was printed on a separate sheet, with room for annotation, and the records were arranged in classmark order.

The audit took place on three days during the university vacation to minimise inconvenience to readers. It was assumed that the twelve weeks since the sample was created would mean the majority of books would have been returned, although some might have been borrowed again. It was decided not to recall books as Bath estimated the proportion of items on loan was as low as 2%.

It took 17 hours to locate and check 305 records and the associated items five positions to the left: a rate of just under 18 records per hour. There were 13 items for which all copies were on loan (over 4% of the sample, but sufficiently few to justify not recalling them) and 3 items could not be found after two searches. One item in Cyrillic script could not be checked and records for four items in Chinese were checked only against cataloguer-supplied transcription. Thus 288 records were found and checked.

The potential for bias in this convenience sample illustrates the problems of designing a fair test and the invalidity of generalising from such a sample. The method of constructing the sample altered the population, excluding all items unavailable for loan, with items held in multiple copies more likely to be included. A solution is to remove all but one of repeated records (in fact there were none).

Inaccurate records may reduce the likelihood of items being found and subsequently borrowed; therefore a book which has been borrowed is likely to have relatively better cataloguing. This factor leads to an underestimation of the error rate in the catalogue.

It is also plausible that borrowing frequency is inversely related to age of item with newer books over-represented in the sample. Furthermore, readers will frequently borrow several books on related topics creating clusters of books in the loan sample with similar classmarks. While true random samples contain unpredictable clusters, this clustering by classmark is predictable and potentially introduces bias.

The 288 records consisted of 30 minimal records, 243 standard records and 15 enhanced records. Of these, 189 were free from error in the fields checked, leaving 99 records (34.4%) with at least one error. Assuming first that each of the 17 unchecked records contains no error, and then that they contain one error apiece, the figure for the sample of 305 records lies between 32.4% and 38.0%. It is not appropriate to calculate a margin of error for this estimate because the sample was not random.

This figure seems high, but many of these errors could be considered minor. The subjective division into major and minor errors proposed earlier labels 22 errors as major; assuming these arise from 22 distinct records, only 7.6% of records checked contain a major error. The non-random sampling means that these rates may not reflect the true incidence in the catalogue; they may, but it is impossible to say.

Although it was not possible to divide the sample by cataloguing date, looking at the error rates for minimal, standard and enhanced records separately gives some indication of variation in quality with time. Date corruption has been ignored here because it skewed the results for minimal-level records, all of which had some kind of error in the date.

With this proviso, 3 of 30 minimal records had errors (10%). There were 221 standard records without pagination, of which 54 had errors (24.4%). Of 22 standard records with pagination – the most recent records, 6 had errors (27.3%), while the 15 UWE-enhanced records included 3 with errors (20%).

The minimal-level records probably have so few errors only because of their simplicity. Although recent records created at Bath appear to have more errors, a more detailed breakdown of the ‘standard’ records is needed to be able to judge whether quality has changed over this period.

The collection-to-catalogue test for completeness was also tested, using items five to the left of the items listed in the printout. With 16 items not located, it was decided that the remaining 288 items would be sufficient. Each one was found on the catalogue and no duplicates were discovered. As the original ‘mini-catalogue’ included separate records for multiple copies of a book, in those cases it was necessary to use accession numbers for identification.

The second (systematic) sample

Although the first sample was not random and therefore its conclusions could not be extrapolated to the whole collection, it served as a useful pilot of the audit technique and of the checklist. To gauge the extent of the underestimation of errors, a second, smaller sample was taken. The 100% intactness rate from the collection-to-catalogue test validated systematic sampling of the shelves, starting at a column randomly selected from the first twenty and taking the rightmost book on the top shelf of every twentieth column thereafter. This excludes items on loan so, by a previous argument, should overestimate errors; the proportion of items on loan had been estimated as around 2%, so this bias would be small.

79 items were checked in two-and-a-half hours, just over 30 per hour; as the collection-to-catalogue test was not performed, this is consistent with the time taken previously. There were 5 minimal, 70 standard (3 with pagination) and 4 enhanced records. By chance, none had corrupted dates.

There were 30 errors from 27 records, of which 9 were major errors. The total error rate (27 out of 79) is 34.2% with a 10.5% margin of error at the 95% confidence level (the margin is so high because the sample was small). The rate for major errors (9 out of 79) is 11.4% with a 7.0% margin of error at the 95% confidence level. The incidence of errors was similar to that found by the first, non-random sample, but with a slightly greater proportion of major errors.

Discussion and refinement

The catalogue-to-collection test

The overall error rate of 34.4% from the first (convenience) sample was higher than expected, given the presumed tendency of records in that sample to have better cataloguing, but the second (systematic) sample confirmed this figure. The incidence of major errors (between 7.6% and 11.4%) was less alarming. That 5.9% of records had errors in more than one field (ignoring those with distinct errors in just one field) suggests some tendency for errors to cluster, so when an error is corrected it is worth verifying the whole record.

The field with the most errors was imprint (21.5%), followed by series (16.3%), edition (10.0%) and title (8.7%). The audit procedure allows for known substandard cataloguing, such as short forms of names in the minimal records; predictable, known errors (such as the corrupted dates in 260) should possibly be discounted in the same way. In both cases the reader will simply note a poor quality record. A possible solution is to present two sets of figures, with such errors included and excluded, although this increases the labour of collection and compilation.

Ignoring the 43 corrupted dates reduces the overall figure for errors to 19.4%, and 6.6% for imprints. To be consistent, the instance of an incomplete title due to a deleted 248 field should also be ignored, although this makes only a minor difference to the results. These manipulations demonstrate an advantage of printing and keeping full bibliographic records: if cumulative figures had been recorded instead then it would have been appreciably harder to extract corrupted dates.

The classification of errors as 'incomplete or incorrect' or 'missing' makes the audit slower while not directly benefiting analysis. Incomplete information is usually better than missing information, but missing information is arguably preferable to incorrect information. The only fields with large discrepancies between the two error types were title (which cannot be missing), statement of responsibility (seldom missing) and imprint (where missing place of publication was recorded as 'incomplete'). It is not obvious that useful conclusions can be drawn from this information.

The collection-to-catalogue test

It quickly became clear when performing the collection-to-catalogue test that very few items were not represented in the catalogue. A far smaller sample is therefore possible: if 1 in 1,000 items is estimated to be uncatalogued then a sample of just 30 will give the true proportion to within 1.1% with a 95% confidence level. Indeed, it is perhaps

unnecessary to investigate uncatalogued items at all without some other motivation such as problems at the point of issue.

One justification for keeping this test is that the catalogue-to-collection test is unsuitable for verifying the classmark given in the record, as items may be supposed missing when in fact the classmark in the record is incorrect. Checking that the shelfmark on an item in hand matches its record's classmark is more sensible than the reverse.

The test for the presence of duplicate records was always extremely cursory. The failure to find even near-duplicates suggests that the more sophisticated approaches of most deduplication studies are essential and that this duplicate test can be abandoned.

How representative is Bath?

There are few reasons to assume that the library at the University of Bath is unrepresentative of medium-sized academic libraries. Most large libraries have had to confront online conversion of card catalogues and this process has inevitably affected catalogue quality – quite possibly positively, as a result of checks at the time of input and increased consistency.

Non-academic libraries may have substantially different collections: public libraries will hold more fiction and non-book material, and workplace libraries will often have more grey literature. In principle, the technique can be applied with the same success regardless of the collection; in practice, the results may be accurate without being useful, giving undue emphasis to accuracy in fields which are seldom searched or failing to detect inadequate authority work. The feasibility of the technique for multi-site libraries also needs to be investigated.

Simplifying the audit

The pilot vindicated the use of sampling in auditing the accuracy of a catalogue. It did demonstrate the problems encountered when a random sample cannot be generated automatically, but the systematic sample taken from the shelves was a successful substitute that most libraries should be able to implement.

Applying the tests in the audit to just 18 items per person per hour may be seen as a luxury by technical services departments with few staff to spare or with a large backlog. Abandoning the collection-to-catalogue test can, at best, halve the time needed for the audit, and this is recommended. It is possible to reduce the sample size and hence the length of the audit by accepting a less accurate estimate of the error rate.

Collecting less information would also save time. The division into incorrect or incomplete information and completely missing fields could be scrapped, with the seriousness of an error decided solely by the field in which it occurs. This simplification makes it practical to include a check on notes fields, which other studies have found to be prone to errors, and which can be important for keyword searches.

The fields are now classed in three groups (Appendix 2 shows the revised worksheet). Group 1 contains authority-controlled access points for comparison with an authority file, while Groups 2 and 3 cover descriptive cataloguing, with indexed fields in Group 2. There is limited authority control for titles in the form of uniform titles, and series headings should, strictly, be controlled, but for simplicity these are not included in Group 1. This grouping allows a direct interpretation of the results in terms of errors in access points and arguably less significant errors elsewhere.

It is not expected that checks on material description, language, genre/category and location/branch, which were not tested, will cause any particular difficulties. The appropriateness of a genre heading is a matter of opinion, but as with subject headings, it is possible to check for keyboarding errors that will prevent access as surely as inappropriate headings.

Limitations

Authority control is a valuable form of quality assurance which the audit technique is weak in evaluating, compared to checking descriptive cataloguing. The pilot was unable to confirm the feasibility of comparing headings to an authority file, which would inevitably increase the time required. An alternative would be to make exploratory searches for the most common types of errors such as qualifiers and additions to distinguish personal names, especially by comparing consecutive headings in an index.

The audit is unsuitable for serials. Many of the fields in the checklist are inapplicable and there is nowhere to note discrepancies in holdings records. Sampling serial titles is theoretically valid, but it is difficult to know whether all holdings must be checked or whether one volume can be taken as representative. A separate tool should be developed for serial records.

The audit considers items individually so it cannot check consistency across records, for example, consistent recording of series statements. Nor is currency of cataloguing addressed, though a sizeable backlog could be said to reduce recall; currency could be measured by the percentage of the collection yet to be processed or the median length of time taken to catalogue an item.

The audit's diagnostic role has been downplayed in favour of simply recording the overall incidence of errors. Not just the quality of the catalogue is of interest but the means of improving it, which requires knowledge of collection areas subject to particular cataloguing problems. Existing suspicion about poor quality can be confirmed or dispelled by separate sampling of relevant sub-populations, but a random sample of the whole collection is the only way to reveal unsuspected or unpredictable errors.

There are conflicting pressures for detail and simplicity, for standardisation and customisation. It may be that libraries have sufficiently different collections that no two will need identical audits; the question is whether common terminology for comparisons justifies a method that suits no library exactly. However, there is nothing to prevent

libraries collecting detailed data for their own purposes before condensing it for external or longitudinal comparisons.

Conclusion

The catalogue audit tool is successful as piloted for evaluating the accuracy of bibliographic records. It should be considered as just part of a toolkit for testing catalogue quality, along with such tools as the collection-to-catalogue test for completeness, focused checks of the validity and appropriateness of assigned name and subject headings, evaluations of serials cataloguing and duplicate detection.

Before starting the audit it is important to profile the collection and its cataloguing. The audit evaluates accuracy in the library's own terms, so variations resulting from known local practices are not penalised. At a higher level, knowledge of the sources of records and of the cataloguing rules and software used is invaluable when interpreting the results.

References

- Ballard, T.** (2001) Typographical errors in library databases. Available at <http://faculty.quinnipiac.edu/libraries/tballard/typoscomplete.html>
- Bryant, P., Venner, G.M. & Line, M.B.** (1972) The Bath mini-catalogue: a progress report, University Library, Bath.
- DiCarlo, M.A. & Maxfield, M.W.** (1988) 'Sequential analysis as a sampling test for inventory need', *The journal of academic librarianship*, vol. 13, no. 6, pp. 345–348.
- Hernon, P.** (1994) 'Determination of sample size and selection of the sample: concepts, general sources, and software', *College & research libraries*, vol. 55, no. 2, pp. 171–180.
- Kiger, J.E. & Wise, K.** (1993) 'Attribute sampling: a library management tool', *College & research libraries*, vol. 54, no. 6, pp. 537–549.
- Kiger, J.E. & Wise, K.** (1996) 'Auditing an academic library book collection', *The journal of academic librarianship*, vol. 22, no. 4, pp. 267–272.
- Massey, O.** (2000) Auditing catalogue quality by random sampling, M.A. dissertation, Loughborough University.
- Orr, R.H.** (1973) 'Measuring the goodness of library services: a general framework for considering quantitative measures', *Journal of documentation*, vol. 29, pp. 315–332.
- UKOLN.** (2002) Introduction to the CAT-ASSESS: the catalogue audit tool. URL: <http://www.ukoln.ac.uk/bib-man/projects/cat-assess/> (Accessed February 2002)

Figure 1. Pilot worksheet

Catalogue-to-collection test

Control number: _____

If the catalogue record lists no copies held for this item, record **code Z** or tick this box [] and go to the collection-to-catalogue test for the item five places left of its classmark. Please tick a box for each field in the catalogue record, or note the codes against each record (for example T2 M4 R1). For author and subject headings, consider only whether they match the forms given in the authority file.

	<i>acceptable</i>	<i>incorrect or incomplete</i>	<i>missing</i>	<i>not applicable</i>	code
Title					T
Material description					M
Statement of resp.					R
Author heading(s)					H
Edition					E
Physical description					P
Imprint					I
Series					S
Classmark					C
Subject heading(s)					D
Language					L
Genre / category					G
Location / branch					B
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	

If there are other errors in the record (for example absence of uniform title) then list them below but do NOT count them in the statistics.

Collection-to-catalogue test

Control number: _____

<i>present (P)</i>	<i>missing (M)</i>	<i>duplicated (D)</i>

Figure 2. Revised worksheet

Control number: _____

Date of cataloguing: _____

If the catalogue record lists no copies held for this item, record **code Z** or tick this box []
 Please tick a box for each field in the catalogue record, or note the codes against each record (for example T2 S1 M3). For author and subject headings, consider only whether they match the form given in the authority file.

	<i>acceptable</i>	<i>incorrect or missing</i>	<i>not applicable</i>	code
Group 1				
Personal authors				A
Corporate authors				C
Subject headings				D
Group 2				
Title				T
Series				S
Group 3				
Material description				M
Statement of responsibility				R
Edition				E
Imprint				I
Physical description				P
Language				L
Genre / category				G
Location / branch				B
Notes				N
	<i>1</i>	<i>2</i>	<i>3</i>	

If there are other errors in the record (for example absence of uniform title) then list them below but do NOT count them in the statistics.