The Worst of Both Worlds: How Old Rules and New Interfaces Hinder Access to Music

By Alastair Boyd
Robarts Library, University of Toronto

Once upon a time, looking something up in a music library catalogue meant riffling through a drawer full of cards rather than typing on a computer. Antediluvian though the concept now seems to anyone under forty, in many ways it worked rather nicely. Granted, you actually had to visit the library in order to consult its catalogue, which was annoying if in the end you had to leave empty-handed. But once your fingers were doing the walking in search of an author, title, or subject, what you encountered in the catalogue looked familiar, since cards and card catalogues all followed essentially the same layout in any library. On each card a brief paragraph or two described in a conventional format all the salient details of a particular book in the collection, much like the citations in annotated bibliographies, with a call number telling you where to find each book or score on the shelf. There were multiple copies of each card filed under the name or names of whoever had written it, the title, and up to three subject headings. In addition, extra cards redirected you from any variant forms of these access points that cataloguers had anticipated you might use.

Not all searches worked equally well this way (subject searches could be particularly laborious). But for so-called "known item" searches, when you wanted to discover whether or not the library had a particular item (and if so, where to find it), then it was reasonably foolproof. If you were looking up an author or composer, all you needed to know were the first few letters of the surname. Everything the library owned written by this person was then at your fingertips, sub-arranged by title in a single alphabetical sequence. And in a music library that used uniform titles, then all variant titles, editions, translations and arrangements of a work were to be found in this sequence.

The greatest strength of the card catalogue—what made it yield such comprehensive search results under the right circumstances—was also its greatest drawback: the only access points were "controlled" access points. (Bibliographic control means adding consistent, standardized forms of access to the inconsistent world of publishing, so as to save library users time and effort.) In the card catalogue era the degree of control was of necessity very high, since the only access points were those that cataloguers had provided. Once you had found the right place, browsing through the catalogue let you cover a lot of ground very quickly. If you had started out by looking up Beethoven's *Moonlight Sonata* you would have been sent to look under *Beethoven, Ludwig van, 1770-1827. Sonatas, piano, no. 14, op. 27, no. 2, C# minor* instead. Once there, at your fingertips was everything the library possessed relating to his Fourteenth Piano Sonata, no matter what the publisher had called it: *Mondschein-Sonate, Sonata quasi una fantasia, Claire de lune, Соната no. 14 для фортепиано*, or whatever. Of

course you could only be redirected to the place where everything was collocated like this if a cataloguer had previously anticipated all the other places where you might look first.

The various incarnations to date of the *Anglo-American Cataloging Rules* (abbreviated as *AACR1* and *AACR2*) and the *Library of Congress Rule Interpretations* that gloss those rules, are teeming with instructions on the making of access points: main entry (so-called, for the author or composer), uniform titles, name added entries (for co-authors, editors, arrangers, etc.), title added entries (for when the titles on the title page and cover differ, or when a subtitle might be taken for the title, or when the title contains abbreviations that might be searched as if spelled in full)—and so on and on. The LCRIs covering title added entries alone fill almost a dozen pages. The very terminology (*main entry*, *added entry*) betrays its origins in the card catalogue, as do the arbitrary limits to adding access points prescribed by the rules. When each added entry required the typing (or printing) and then filing of an additional copy of the catalogue card under the added heading, it was expedient to limit the number of controlled access points per record, balancing the benefits of extra standardized headings against the costs of creating and maintaining them, and of increasing the size of the catalogue.

Nowadays we enjoy the miracle of online computer catalogues with keyword access to almost every word in each catalogue record. But this is a mixed blessing, as blessings often are. The obvious drawback is that keywords in parts of the record transcribed from the published item—titles and table of contents notes, for example—are uncontrolled, i.e. they use the language and spelling of the publisher, and are therefore liable to give incomplete or misleading search results. If you were to try a title keyword search for *Rite of Spring* in most current online catalogues, you would not retrieve records for scores or recordings published under the titles *Sacre du printemps* or Весна священная. Worse, you would receive no indication that your results had been incomplete. "If I had to blame the stunted development of online catalogs on one thing," writes Andrew K. Pace in *American Libraries*, "it would be the introduction of keyword searching….Our profession's utter fascination with the ability to keyword search a catalog has kept both librarians and vendors myopic for nearly 25 years."[1]

Some think this problem can best be solved by applying more computing horsepower at the search end of the equation, rather than by the traditional cataloguing method of adding controlled headings to the bibliographic record and cross-references to the catalogue (what was described as "hand-crafted metadata" by Dale Flecker from Harvard University at last winter's PCC Participants' Meeting at ALA Midwinter).[2] These visionaries seem to have in mind a sort of next-generation Google, where search terms can automatically be translated and expanded to include all conceivable synonyms and variations, in order to make the search truly comprehensive. At the same time, a magic algorithm will detect *only* the relevant results and discard the rest. In the context of a large academic research

[1] Andrew K. Pace, "My Kingdom for an OPAC," *American Libraries* 36, no. 2 (Feb. 2005): 4.
[2] Dale Flecker, "OPACs and Our Changing Environment," summarized on the Program for Cooperative Cataloging Web site (Sept. 30, 2005): http://www.loc.gov/catdir/pcc/archive/ pccpart05m.html .

collection in many languages, it is going to take a lot of time and money before an automated solution will produce results half as satisfactory as those produced by "hand-crafted" means.

Surely it will be more productive to harness the powers of automation to reduce the cost of the hand-crafting—that is, to assist rather than supplant the human skills that create a catalogue. I will come back to this point in a minute. First, having suggested that it is premature to discard authority control and human skills, I must also emphasize that the status quo is not working well. Even though computer online catalogues have removed the physical limits to the number and nature of additional access points that could be made, cataloguers are still following AACR2 rules about which controlled headings to add. Thus our online catalogues contain a bizarre mixture of controlled and uncontrolled headings: standardized browse index entries for some names or titles but only uncontrolled keyword access for others. For instance, a CD containing up to 15 works, roughly speaking, will have standardized name-title index entries for all of them; but add another work, and suddenly none of them receive this treatment; and printed music has a cut-off point of three such added entries. If the "Contents" note is in the keyword index (it usually is) then names and titles can be retrieved that way, but only in the uncontrolled form transcribed from the item. Because keyword searches are based on this mixture of (some) controlled and (many) uncontrolled headings, it makes it very difficult to assess how complete the results are, and to see how they relate to each other, however one chooses to sort them. AACR2's roots in the card catalogue era are also evident in the kinds of material

that it was created (in 1978) to describe. That is to say, primarily print. Now we have to deal with digital media, e-journals, and the World Wide Web. It is this, as much as the points already mentioned, that has motivated the Joint Steering Committee for Revision of AACR to embark on a radical overhaul of the rules. Even the proposed name of the new code—*RDA : Resource Description and Access*—shows the committee's determination to be inclusive. The JSC's current deliberations are being informed by a new bibliographic model published a couple of years ago by IFLA (The International Federation of Library Associations and Institutions) under the title "Functional requirements for bibliographic records" (or FRBR for short).[3]

To summarize much recent discussion: FRBR makes four subtly graded distinctions between a *work* considered as an abstract entity, and a specific copy of an *item* representing a tangible example of a work. These form a hierarchy of varying degrees of abstraction. Some confusion arises because FRBR divides the three things we used to call "work", "edition" and "item" into four "entities." Barbara Tillett[4] has provided the most succinct explanation of these, which Patrick Le Boeuf, in a whimsical essay,[5] has paraphrased in terms of music to show how

---

[3] IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records: Final Report*. München: K.G. Saur, 1998. Available online (Oct. 1, 2005) at http://www.ifla.org/VII/s13/frbr/frbr.pdf and at http://www.ifla.org/VII/s13/frbr/frbr.htm.
[4] Barbara B Tillett, "FRBR (Functional Requirements for Bibliographic Records)," *Technicalities* 23, no. 5 (Sept./Oct. 2003): 1, 11-12.
[5] Patrick Le Boeuf, "Musical Works in the FRBR Model or 'Quasi la Stessa Cosa': Variations on a Theme by Umberto Eco," *Cataloging & Classification Quarterly* 39, no. 3/4 (2005): 103-124

the word "score" can mean several different things to a musician. In essence he says that: (1) a score is a physical object you take to a rehearsal—FRBR calls this an *Item*; (2) a score is also a collective term meaning "publication," describing a particular edition—FRBR calls this a *Manifestation*. The Dover full score of Holst's *The Planets* is a manifestation; the Dover miniature score is a different manifestation; (3) a score can mean something yet more collective, denoting (for example) the printed version of the music as opposed to a recording—the FRBR term for this is *Expression*. At this level we say "score" to distinguish, say, print from audio; (4) "score" can also be a fancy synonym for "work"—*The Planets* is Holst's most extravagant score.

Although the FRBR terminology may be new, the underlying concepts are already familiar to music librarians. The term *work* as a sort of platonic fundamental underlying all expressions, manifestations and items has produced a variety of definitions, but IFLA's is the shortest: "a distinct intellectual or artistic creation."[6] Music cataloguers have a head start here, since they have been heavy users of uniform titles since practically the dawn of time. And uniform titles not only establish standardized titles for music that may be published under various titles (as I described above), but in so doing establish headings for works, in order to make explicit the "derivative" and "whole/part" relationships so common in published music.[7] (A name-title heading like *Wagner, Richard, 1813-1883. Ring des Nibelungen. Walküre. Feuerzauber; arr.* shows both the whole/part relationship of the "magic fire" music to the parent work, and also indicates that this is an arrangement, i.e. a work derived from the original one.)

An overhaul of cataloguing principles and rules has been suggested based on the FRBR model. Current rules, it is argued, were shaped by the limitations of the card catalogue: "Since the card catalog could not support a hierarchical model, the selection of the basic entity for cataloging was an either/or decision. Most cataloging codes, including AACR, chose the manifestation as the basic bibliographic unit."[8] Current catalogues, therefore, describe a manifestation of an expression (a particular edition of a work) in the bibliographic record, to which is hitched information about the number and location of items (how many copies, locations, call numbers and barcodes by which they may be retrieved—i.e., the "item record" or "copy record" already familiar for more than a decade). Where music is concerned, the concept of the *work* is often already embedded in the record in the form of the main entry combined with the uniform title, or with name-title added entries. Distinct records at the level of the work already exist in the shape of name-title authority records. Therefore the thing that most needs beefing up to make the four FRBR entities explicit in catalogues is the principle of uniform titles. Since this device for identifying works already exists as part of current records and authority files, it seems sensible to build on this. As Tillett puts it: "The JSC is using FRBR … to re-examine and, hopefully, improve the traditional linking devices of uniform titles.… Perhaps an expression-level citation or work-level

---

[6] IFLA, p.16
[7] Tillett, p. 11-12

[8] Edward T. O'Neill, "FRBR: Functional Require-ments for Bibliographic Records Application of the Entity-Relationship Model to *Humphry Clinker*," *Library Resources & Technical Services* 46, no. 4 (Oct. 2002): 150-159

citation … would offer more collocation and differentiation capabilities than current uniform titles."[9]

If the new cataloguing code enhances the scope of name-title headings (uniform titles) so as to better differentiate between "derivative" forms of works, and abolishes the old limits on the number of added entries, then our catalogue records will have the potential to provide amazingly precise and well-ordered search results. But what about the added costs of establishing all these headings? asks the head of many a cash-strapped cataloguing department. Costs could be reduced by greater sharing of the "authority" MARC records in which name, name-title and series headings are established and supplied with cross-references. Just as the widespread sharing of bibliographic records saves cataloguing time and effort, so too does the sharing of authority records. The Library of Congress maintains a national authority file of such records which is used throughout the AACR world. Under the umbrella of the Program for Cooperative Cataloging (PCC) an increasing number of libraries are contributing to this file through NACO (the Name Authority Cooperative Program of the PCC). As of September 2004 the libraries in the NACO Music Project had contributed over 110,000 name and name-title authority records.[10] It would help if there were more contributors, of course. (Canadian music libraries that are not cataloguing in the OCLC bibliographic utility can still join via NACO Canada.)

The features of most MARC cataloguing systems could be supplemented with macros and plug-ins that generate any number of added entries based on names that occur in the descriptive parts of the record (Joel Hahn of Niles Public Library District has already published a collection for use with OCLC's *Connexion* client). It would also save time and reduce costs if LC's authority file could be more seamlessly integrated (via macros and scripts) with the cataloguing modules of the systems now in use at most large public and academic libraries. Since the Internet searching protocols known as Z39.50 work for authority MARC records just as they do for bibliographic records, a consortium of libraries could even share the cost of mounting a copy of the LC name authority file on a Z39.50 server. (Sirsi already provides such a database for its customers, as does OCLC for its members.) So, updating the rules for cataloguing, FRBR, expanding the shared use of authority records, and automating the creation of headings and authority work: these things will improve the quality of the library database. But this will be of little use until improvements are made to the OPACs which stand between catalogue users and catalogue records. An eloquent denunciation of the status quo comes from Tim Burke (of Swarthmore College). In the January 20, 2004, entry ("Burn the Catalog") from his blog, *Easily Distracted,* he expresses his frustration "at just how useless a typical electronic catalogue has become.... Wherever you go in the academic world [they] have become a horrible crazy-quilt assemblage of incompatible interfaces and vendor-constrained listings."[11] Martha M. Yee has a more systematic examination of the mess, and the need "to improve catalog

[9] Tillett, p. 12.
[10] See the NACO Music Project Web site: http://www.musicoclcusers.org/nmp.html (Oct. 1, 2005).

[11] Tim Burke, "Burn the Catalog", http://www.swarthmore.edu/SocSci/tburke1/perma12004.html (Oct. 1, 2005)

Heading	Titles
beethoven ludwig van 1770-1827 sonatas piano no 14 op 27 no 2 c minor	42
beethoven ludwig van 1770-1827 sonatas piano no 14 op 27 no 2 c minor adagio sostenuto	1
beethoven ludwig van 1770-1827 sonatas piano no 14 op 27 no 2 c minor adagio sostenuto arr	1

Figure 1

design based on a better understanding of [existing MARC] records" in a paper on FRBR and OPAC design.[12]

It's those keywords: first, the incomplete and imprecise retrieval from keywords, already mentioned; and second, the display of the search results in most current OPACs. If we compare keyword to browse search results, we find that the latter are far more useful (as long as the indexes are properly built). Browse lists are sorted alphabetically and hierarchically reminiscent of the card catalogue filing order. So if one takes the trouble to search a composer-title heading such as "*beethoven ludwig van 1770-1827 sonatas piano no 14*" then most web catalogues will return the closest matches from the browse index in a list like that shown in Figure 1. Such a list identifies the work by the uniform title, and shows the precise number of records associated with each, letting users choose which in particular they want to look more closely at. Browse searches also make use of the cross-reference structure built into the underlying authority records, so if one had searched "*beethoven ludwig van 1770-1827 moonlight sonata*" one would be redirected to the established heading, just like the old "see" entries in the card catalogue.

But who is going to type in great long strings like these when they search? Keyword searches, particularly combined author-title searches, are a much more natural strategy. What is needed, then, is a system that will send such searches through the authority record reference structure (as browse searches do) in addition to the keyword indexes. The user, in addition to getting the actual keyword results, could then be offered some suggestions to guide further searches. Thus a keyword search on "*beethoven*" + "*moonlight*" would not leave the user to puzzle over a suspiciously short result list representing only editions published under the English title, but would provide links for follow-up browse searches based on matching uniform titles, which would then lead to the precise, complete, and ordered display shown above. If this sounds unrealistic, then look at Laurentian University Library's implementation of Sirsi's Unicorn system. This is currently configured so that keyword searches are sent in parallel to the subject authority file, offering users suggestions ("*Try these too…*") for further searches based on controlled subject headings, in addition to the keyword search results. It doesn't seem unreasonable to imagine that a similar mechanism could be applied to the name authority file for combined keyword searches involving names and titles.

---

[12] Martha M. Yee, "FRBRization: a Method for Turning Online Public Finding Lists into Online Public Catalogs," *Information Technology and Libraries* 24, no. 3 (2005): 77-95. Postprint available free at: http://repositories.cdlib.org/post prints/715.

The hope is that FRBR principles will lead to new web catalogues where the display of search results will be organized around the *work* and *expression*. Imagine users looking at the hitlist shown in Figure 1, and suppose that from this top-level view they identify the *work* they want and click on the first line, with 42 associated records. Current displays lack any intervening step between this overview (the *work* level) and a record-by-record list of all 42 results (the *manifestation* level). Shouldn't we have the option of seeing a summary of the various *expressions*? After all, users probably know if they want a score or a recording, and if the latter, which kind. The Research Libraries Group has a web search engine called RedLightGreen™ which summarizes results from participating libraries by expression.[13] And the Library of Congress MARC standards site contains documentation and a toolkit suggesting how we might begin to achieve similar clusters of expressions in a local catalogue.[14] Thus, clicking on the first result shown in Figure 1 could produce a list as shown in Figure 2.

Behind such a display there has to be a system-generated expression-level record, created and updated like a sort of super-index each time another manifestation (i.e. bibliographic record) for the same work is added to the database. If it seems starry-eyed to imagine Library System vendors taking on this challenge, then we need look no further than the VTLS *Virtua* system, which has gone a long way towards achieving this

Beethoven, Ludwig van, 1770-1827.
  [Sonatas, piano, no 14, op 27, no 2, c minor]
Scores:
    10 editions published between 1898 and 2004
Recordings (CD):
    14 editions published between 1986 and 2001
Recordings (LP):
    8 editions published between 1956 and 1983
Recordings (tape):
    4 editions published between 1979 and 1986
Related works:
    6 editions published between 1935 and 1995

Figure 2

FRBR-awareness already.[15] We should also remember that nowadays we are less helpless in system vendors' hands than we used to be. Open standards and scripting tools enable us to build additions to supplement the system vendor's options. Consider Kent State Library's ingenious added interface for searching by instrumentation,[16] or (to look beyond music for a moment) the University of Toronto's brand new "original script" viewer for records with Chinese or other non-roman characters. These things not only fill certain gaps in their underlying systems, but can even serve as models to help system designers better understand what we want, and how to achieve it.

These days we are asking far more of a library catalogue than could ever be attained through printed cards. But modern marvels of remote access, hyperlinks and colourful displays can surely supplement rather than ignore the orderly and comprehensive structure of a card catalogue. Let's aim for the best of both worlds.

---

[13]For details and examples, see the RLG Web site at http://www.rlg.org/en/page.php?Page_ID=433 (Oct. 1, 2005)
[14] See: http://www.loc.gov/marc/marc-functional-analysis/ (Oct. 1, 2005).

[15] VTLS has documentation on this (PowerPoint format) at: http://www.vtls.com/Corporate/ FRBR. shtml (Oct. 1, 2005).
[16]See: http://www.library.kent.edu/kentlink_ instrument _ search.php (Oct. 1, 2005).