Title

Permalink
https://escholarship.org/uc/item/9fw5b54w

Journal

Author
Yee, Martha M

Publication Date
1995

Peer reviewed
We all of us, grave or light, get our thoughts tangled in metaphors, and act fatally on the strength of them.

—George Eliot, *Middlemarch*

Things should be made as simple as possible, but no simpler.

—Alfred Einstein

In an attempt to divine the cataloging future with regard to editions, I will first establish some rough definitions. Then I will try to outline the goals and objectives to be met in describing and providing access to editions. Finally, I will ask how we can best meet those goals and objectives.

Please note that I have put the term *brainstorming* into my title in order to emphasize that the following is an essay about what the ideal approach might be for creating a national catalog built by cooperative cataloging. I have ignored all the political and economic obstacles that would lie in the path to creating such a catalog, if it were ever decided to do so. In other words, I have taken *brainstorming* to mean that one can temporarily shed one’s obligation to be practical and dream a little.

Before beginning, perhaps I ought to discuss which parts of AACR2 might be affected by changes discussed here. In the narrowest approach to change, only the glossary definition of the term *edition* might need tinkering with. In the broadest approach, the entire code could be affected, since one of the main purposes of the descriptive part of a cataloging record is to identify a particular edition of a particular work and distinguish it from other editions of that work. Each element of the description has a role to play in that identification and distinction process. The overall organization of chapters 1 through 12 of AACR2 undoubtedly needs work, now that any format can be combined with any other(s). In addition, one of the main purposes of headings, uniform titles, and references (chapters 21 through 26) is to demonstrate relationships among the various editions of a work. How much of the code needs change depends on the degree to which the creation, distribution, access, and use of the documents we are trying to describe have changed and will continue to change as we undergo a tremendous technological revolution.

**Some Definitions**

To begin, in order to provide some quick context for the discussion that follows, I am going to define the following terms by way of example, rather than trying to develop full explanations, complete with reasoned and watertight arguments for defining these concepts this way. (Those interested in fuller explanations I refer to my series of articles on “work,” published in *Cataloging & Classification Quarterly*, as well as my series of articles on “manifestations” and “near-equivalents,” published in *Library Resources & Technical Services.*

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
<td>Shakespeare’s <em>Macbeth</em></td>
</tr>
<tr>
<td>Version</td>
<td>A translation of <em>Macbeth</em> into German</td>
</tr>
<tr>
<td>Edition</td>
<td>A second edition with corrections of the translation of <em>Macbeth</em> into German</td>
</tr>
<tr>
<td>Near-equivalent</td>
<td>A microform copy of the second edition with corrections of the translation of <em>Macbeth</em> into German</td>
</tr>
<tr>
<td>Copies</td>
<td>Two copies of the microform near-equivalent of the second edition with corrections of the translation of <em>Macbeth</em> into German</td>
</tr>
</tbody>
</table>

I had a great deal of help on this chapter from John Attig, Jo Crawford, Sara Shatford Layne, and Brian Schottlaender, and would like to thank them for taking time from their busy schedules to catch my errors, argue with me, and suggest ways to strengthen my arguments.—The Author

I should mention one other concept here, that of “super-work.” Some works spin off into other works: They are made into operas, plays, movies, and so on. When that happens, some would argue that the original
work becomes a “super-work” with other works derived from it. Two examples of works derived from a super-work would be the films that Roman Polanski and Akira Kurosawa made of Macbeth, Polanski’s Macbeth and Throne of Blood, respectively. Most would recognize that these films are not simply editions of Macbeth, but are new works. However, they certainly have been derived from Shakespeare’s play and should be brought to the attention of users interested in performances of Macbeth.

In this chapter, I want to focus on the concept of “edition.” I would like to propose the following working definition for the term edition: an item that is essentially the same work as another item, but with some differences significant to users. The first LRTS article mentioned earlier contains an extensive discussion of the types of differences that can occur and of the significance (or lack thereof) of these differences to users. Following are brief lists of the major types of differences, classified into two categories: (1) significant differences that might need a separate record to express them adequately to users, and (2) less-significant differences that can perhaps be recorded on subrecords. This classification, of course, is my own, and does not, unfortunately in my opinion, correspond to current practice. I jump right into it here in order to stimulate thought and discussion on potential directions our cataloging rules could take.

Notice that the focus in these lists is on differences that are observable by a working cataloger. One of the questions I attempted to address in my dissertation research on moving-image materials was that of how often differences occur that are not observable by working catalogers. This question could profitably be posed for other types of materials as well, and the answers could bear on future approaches to the problem in the cataloging rules.

1. Significant differences that might need a separate record to express them:
   a. Change in title;
   b. Change in series title;
   c. Change in statement of responsibility, including subsidiary responsibility, signaling change in actual intellectual and artistic content (commentators, editors, translators, illustrators, etc.);
   d. Explicit statement of change in intellectual and artistic content (e.g., revised edition, enlargements, abridgments);
   e. Change in extent (i.e., change in actual intellectual and artistic content, or potential change because of, for example, resetting of type);
   f. Explicit change in list of contents (e.g., different songs or different takes of songs listed on CD compared with sound recording disc or audiocassette); and
   g. Change in language.

2. Less-significant differences that can perhaps be recorded on subrecords:
   a. Change in edition statement without corresponding change in intellectual and artistic content (e.g., paperback edition);
   b. Change in distributor or distribution date or both without corresponding change in intellectual and artistic content; and
   c. Change in physical format without corresponding change in intellectual or artistic content (e.g., a microform copy of a text, or an audiocassette copy of a CD; this category is commonly known as a reproduction).

The Future

If every document is likely to become an electronic document in the future, one has to wonder if some of these types of differences might change. For example, if pathways to documents become relatively stable eventually, and invisible to most users, will the publication/distribution area of the cataloging record tend to wither away? Or does this area have other functions besides that of recording where one can obtain a copy? If so, will such other functions continue to operate? For example, some might argue that publication source can bestow prestige on a document (or conversely, cast suspicion on it). On the other hand, perhaps that is a type of responsibility that could be recorded with other statements of responsibility.

In the current state of electronic publication, the pathways to documents are far from stable; in addition, the same document can be available from a number of different sources and, therefore, have a number of different URLs. Different documents in different locations can have significant intellectual and artistic differences over time as one is revised but another is not. In fact, some types of electronic publications can be put up so differently in two different sites, with such different types of access to the document available, that it could be argued that each time the document is put up a new edition is created. Ownership—even authorship—of documents seems to be in a state of flux, leading one to wonder if we are headed for the kind of community ownership of a text that used to prevail.
in oral cultures. In this context, as John Attig points out, publication information may be an indication of authenticity of a version that has some provenance relationship to the original creator of the document. If these trends continue, they will have the effect of increasing the importance of the “distributor/publisher” (i.e., the place at which the document can be accessed), rather than diminishing it, and will simply exacerbate the current problems with editions and near-equivalents.

If all formats eventually can be run on any computer system, will parts of the physical description area of the cataloging record, as well as its early warning flag, the GMD, tend to wither away? Without the need to be so concerned about whether particular playback equipment will be available, will there be a tendency to generalize the physical description more, so as merely to indicate the presence or predominance of image (moving or still), sound, or text? What function will the physical description serve?

Certainly, at the current time it is not yet possible to run even all-text formats on any given computer system, thus the current proliferation of such near-equivalents as the document in ASCII, Acrobat or PostScript format, and so on.

One can’t imagine the extent statement losing its usefulness, can one? How will extent be communicated for electronic documents, when we no longer have physical volumes and pages? If character counts (for textual materials), frame counts (for visual materials), and playing time (for sound and moving-image visual materials) were to become the norm, it might become much easier to use extent to spot actual change in intellectual and artistic content. We could even envision computer programs that carry out comparisons between two items and highlight differences for catalogers to summarize for users, and for users to display on demand.

Goals and Objectives

What would we like our catalogs to do for us and for our users with regard to editions? In the past, we have had two sets of answers for this question. One set of answers is found in the objectives of descriptive cataloging, which were actually included in AACR1, but have been dropped from AACR2. “Descriptive cataloging” here refers only to the preparation of the body of the record, as opposed to choice and form of access points. According to AACR1, the objectives of descriptive cataloging are:

1) to state the significant features of an item with the purpose of distinguishing it from other items and describing its scope, contents, and bibliographic relation to other items

2) to present these data in an entry which can be integrated with the entries for other items in the catalog, and which will respond best to the interests of most users of the catalog. . . . Extent of description: The item is described as fully as necessary to achieve the objectives stated above, but with economy of data and expression.

In other words, we would like the description to distinguish this edition from other editions of the same work. We would also like the description to identify it as being the same work as other editions of the work it represents. And finally, we would like the distinction and identification to be carried out as efficiently as possible, “with economy of data and expression.”

The other set of answers is found in the “objects,” or functions of the catalog, first recorded by Cutter, further developed by Lubetzky, and adopted as an international standard at the International Conference on Cataloguing Principles in Paris in 1961:

2. Functions of the catalogue
   The catalogue should be an efficient instrument for ascertaining
   2.1 whether the library contains a particular book specified by
      (a) its author and title, or
      (b) if the author is not named in the book, its title alone, or
      (c) if the author and title are inappropriate or insufficient for
         identification, a suitable substitute for the title; and
   2.2 (a) which works by a particular author and
      (b) which editions of a particular work are in the library.

Function 2.2b is that of being an efficient instrument for ascertaining which editions of a particular work are in the library. In other words, we would like our catalogs to enable users to see all the editions of a particular work so that those users can make up their own minds about which edition might best meet their needs. To do this, we need some sort of linking mechanism to tie together all the editions of a work in a particular catalog.

Lest we think of abandoning that goal (the linking goal, which is admittedly a very difficult one to carry out), Ross Atkinson reminds us, in his recent article, that linking—or what he refers to as “reference to the citation”—is the fundamental goal of all libraries and of the humanities in general. To paraphrase his arguments (I hope he will forgive me), humanists ask the question, “How did we get here (the present) from there (the past)?” In contrast, scientists are more apt to ask, “How can we get there (the future) from here (the present)?” Our job as librarians is to keep and organize for access the cultural record of our society, so that (1) humanists can try to answer their question, and (2) new scientists can be
relationships as part of our shared cataloging programs. Our inability to effectively and consistently share identified bibliographic information needs to be kept distinct. (See my article "Manifestations and Near-Equivalents" for a more extended discussion of this problem.)

As I have attempted to argue here, the problem is actually much bigger than just that of reproductions, although the speed with which records for reproductions are cluttering up our national databases makes that part of the problem particularly vivid for most of us. The real problem lies with conflating holdings information (which library holds what volumes of what titles) with near-equivalent information. For clear displays, such information needs to be kept distinct. (See my article "Manifestations and Near-Equivalents" for a more extended discussion of this problem.)

What might the future hold in the way of a solution to this problem? As I have attempted to argue here, the problem is actually much bigger than just that of reproductions, although the speed with which records for reproductions are cluttering up our national databases makes that part of the problem particularly vivid for most of us. The real problem lies with our inability to effectively and consistently share identified bibliographic relationships as part of our shared cataloging programs.
Martha M. Yee

Editions: Brainstorming for AACR2000

Figure 3. A VHS videocassette of Patton released in 1978 by Magnetic Video Corp.

Figure 4. A VHS videocassette of Patton released in 1984 by CBS/Fox Video. Playing time is still 171 min., so it is probably identical to the previous one.
Figure 5. Thesis—the original, cataloged at UCLA

Figure 6. Thesis—the photocopy made by UMI

Figure 7. Thesis—the microfiche copy made by UMI
How Can We Best Link the Editions of a Work?

Now

Currently, for works of personal authorship that are not subject to revision, for translations, for works about which other works have been written, and for anonymous works, we can follow Library of Congress practice and create a uniform title that stands for the work (but many libraries do not, and those that do often have systems that won’t display uniform titles). This uniform title can be added to all editions of the work and have the effect of collocating all the editions at one place in the catalog under the name of the principal author of the work or under the uniform title. In effect, the many editions link to the one main entry. A major problem with this approach is that the user, to see all the editions, has to be familiar with this structure. If a search matches on the main entry to be certain all editions, as well as all related works, were included. What works are linked (if at all) with chained added entries? Each new edition is given an added entry for the main entry of the last edition, if it is different. This is the same approach that is used for serials that change titles. This one-to-one link does not have the effect of gathering all records representing the work together at one point in the catalog. Also, if any particular catalog lacks one link in the chain, the connection is broken.

Because neither of these approaches has been consistently applied, and because they are rather complex techniques, it would be very difficult to program a computer to find all editions of a particular work from the cataloging record for one particular edition.

In order to keep the costs of cataloging down, shared cataloging programs have been developed extensively in this country. However, shared cataloging can have the effect of working against the functions of the catalog. The products of shared cataloging are individual records in an atomized catalog, if you will. These atoms link to each other only when two records contain the same character strings in a normalized heading field. Certainly, we share the creation of authority records as well as bibliographic records. However, the creation of an authority record for a particular author or work does not automatically cause the form of that author’s name or the uniform title for that work to be updated in every bibliographic record. It could be argued that a cataloging department is taking on the responsibility for maintaining not just its local catalog, but a national utility database and the Library of Congress’s catalog as well. Maintaining three catalogs is more work than maintaining one and, even if those three are perfectly maintained, that does not take care of the problem of all the other local catalogs that are not updated when a heading is changed.

The Future

In the following discussion, I would like to make explicit four assumptions:

First, it is still necessary to keep the cultural record (but that record may not include absolutely everything on the Internet). Keeping the cultural record and making it accessible is the professional calling of the librarian. If someone does not take on this responsibility, we are likely to see a new Dark Age beginning in the next century.

Second, artificial intelligence is not the answer. As one computer scientist puts it, “After fifty years of effort... it is now clear to all but a few
diehards that [the] attempt to produce general intelligence [on the part of a computer] has failed. . . . The know-how that made up the background of common sense could not itself be represented by data structures made up of facts and rules." Machines have had a particularly hard time "learning" natural language and "learning" how to do recognition tasks, such as recognizing the nature of the relationship between two entities. Not unexpectedly, efforts in our field to build expert systems have not been very successful. Hjerpe and Olander report on a project that built two expert systems for cataloging; they note that "much of the present cataloging process consists of 'instinctive' interpretation, based essentially on experimental learning from examples in an apprenticeship manner." Among the number of interpretive acts they identify that are difficult for computers to carry out are "the recognition of an item as possibly being related to other item(s) and identifying such item(s)." Humans can perform such recognition tasks nearly effortlessly. For example,

This different name probably represents the same person.
This different title probably represents the same work.
This same name probably represents a different person.
This same title probably represents a different work.

Recognition of the likelihood of a relationship can then trigger research to confirm or deny the existence of one.

What we need is not artificial intelligence but, rather, human intelligence applied toward developing human–machine partnerships that maximize human intellectual input and minimize human drudgery. Just the reduction of the number of catalogs to maintain from three to one ought to increase the "more-better-faster-cheaper" numbers. If catalogers did nothing but identify relationships all day long, they could accomplish much more work in a day than they do now on largely antiquated editing software in many different systems, few of which have been effectively designed to support cataloging work per se.

Third, the term "Information Superhighway" can be translated to mean ubiquitous and cheap telecommunication that could enable us to create a single virtual catalog that would be more like a coral reef built up by catalogers over time, rather than the current catalog model that resembles a cloud of atoms buzzing about, sometimes linking up when they should and sometimes not.

Fourth, eventually we will be able to erase the distinction we make now between system design (indexing and display in OPACs) and record design (MARC and AACR2) and once again take a holistic view of catalog design. Because of this last assumption, this chapter will not suggest specific changes to the rules in AACR2. Instead it will suggest specifications for a holistic cataloging system that encompasses both record and system design. I have attempted to argue here that the problems described are rooted in our cataloging system as a whole, not just in AACR2. To solve them requires systemic change.

With these assumptions in place, I would like to suggest the following specifications for the ideal AACR2000 catalog system that would link editions for users, no matter what the initial search might be. Please remember that the real problem is not the need for mechanical linking devices per se. They are readily available now through hypertext linking. The problem is to devise a method for creating one-to-many links that are sharable, immediately ubiquitous, and permanent.

1. The system would recognize the following six hierarchical levels: superwork, work, version, edition, near-equivalent, and copy. In a sense, this approach is a back-to-the-future approach (see Figure 9).
2. A human operator would be able to point to two records and click on a type of relationship (e.g., same work, same version, different edition; or same work, same version, same edition, different near-equivalent).
   a. This action of recording a relationship need be done only once (i.e., it need not be replicated in multiple databases).
   b. The recording of the relationship will be permanent (but editable).
   c. The recording of the relationship will be immediately ubiquitous (i.e., visible to all users; shared).
   d. At any level with levels below it, a textual label or citation form will be devisable to identify or name the one entity—for example, the superwork, work, version, edition, or near-equivalent—to which subrecords can be linked. This label can be derived from the description of the entity (e.g., main entry [author and title, or title] for the work).
3. As long as local physical collections exist, users should be allowed to limit or prioritize their searches to items that are either locally held or readily available online and that are in particular formats, and they should have ready access to any call number, location, holdings, and circulation information needed to obtain the item or a particular volume or part of it.
   Perhaps in the future we can build systems to maintain electronic documents online as works, with subsequent editions simply appended to the existing record.
AN ESSAY ON THE EXTERNAL USE OF WATER.

An essay on the external use of water. In a letter to Dr. **** with particular remarks upon the present method of using the mineral waters at Bath in Somersetshire, etc. London: printed for M. Cooper; sold by D. Wilson; Bath: sold by Leake & Frederick, 1792. C. 123. k. 3. pp. 48. 4°.


THE EXPEDITION OF HUMPHRY CLINKER.


Anonymous.


Anonymous.


Anonymous.

[Another edition.] Dublin: A. Leathley, etc., 1771. 1478. e. 41. 2 vol. 12°.

Anonymous. Vol. 2 is a duplicate of the preceding.


Anonymous.


Anonymous.

The expedition of Humphry Clinker, etc. 1775. See supra: [Collections.] The select works of T. Smollett, etc. vol. 7, 8. 1776. 12°.

Anonymous.


Anonymous.


Anonymous.


Anonymous.


Anonymous.

[Another edition.] Dublin: W. Sleater, etc., 1784, 85. 1471. de. 44. 2 vol. 12°.

Anonymous.
Figure 9.—continued

SMOLLETT (Tobias George)


   P.P. 5262. aa. [vol. 19.] 1785.


   12602. a. 5. 2 vol.: plates. 12°. (Cooke's pocket edition of select novels.) Anonymous.

   1608/5096. 2 vol. 12°. Anonymous. The booksellers' names in the imprint are probably fictitious.

   12612. de. 23. 2 vol.: plates. 12°. (Cooke's pocket edition of select novels.) Anonymous. [1800 1]


Some of the plates have been re-engraved.


   2 vol. 12°. (The British novelists. vol. 30, 31.) Anonymous. The booksellers' names in the imprint are probably fictitious.


   3 pt.: pp. 159; illus. 8°. (Limbird's novelists. pt. 32-34.) Anonymous. Published in parts.
SMOLLETT (TOBIAS GEORGE)


pp. xxxvi, 403: plates; port. 8°. (The novelist’s library. vol. 1.)

635. e. 16.


12267. a. 1/302.

pp. 387. 8°. (Collection of British authors. Tauchnitz edition. vol. 22.)


pp. 281: plates. 18 em.


12619. l. 31.

pp. iv, 125. 22 cm. (Routledge’s sixpenny novels.)


2502. c. 20.

pp. viii, 384: plates. 19 cm. (Bohn’s novelist’s library.)


pp. 419: plates; port. 18 cm. (Classic novels.)

012613. ee. 46/9.


012209. d. 186.

pp. xx, 440. 16 cm. (The world’s classics. no. 200.)


12613. b. 27.

pp. xxiii, 453: port. 17 cm. (New Nelson classics.)


012613. ee. 44.

pp. xvi, 433. 17 cm. (The modern library of the world’s best books.) The titlepage is a cancel.


pp. xi, 372. 18 cm. (Everyman’s library. no. 975.)


W.P. 9139/66.

pp. xvi, 433. 19 cm. (The modern library of the world’s best books. no. 159.) Earlier publication in this series, [1937].


12653. aa. 45.

pp. 352: port. 19 cm. (Collins new classics. no. 632.)


12661. b. 26.

pp. x, 371: plates; port. 23 cm.


X. 0906/46. (8.)

pp. xxii, 375: map. 21 cm. (Oxford English novels.)


X. 0907/15. (21.)

pp. 414. 18 cm. (Penguin English library. no. BL 21.)


X. 908/20503.

pp. xxxvii, 346: plates; maps. 21 cm. (Riverside editions. no. B67.)
Conclusion

These are exciting times. The systems we design now will determine whether twentieth- and twenty-first-century culture will be preserved for the ages. Now more than ever, it is important to go back to first principles and remind ourselves of our goals and objectives. We must remember our responsibility to the many library users, who are mute because they know not how to speak, but who depend on us to make wise decisions so that they can find the works they seek.

Notes


3. My use of the term version here is closer to the traditional one than is the current use of the term in the phrase "multiple versions" (for which I prefer the term near-equivalent). Examples of traditional definitions of the term version follow. The ALA Glossary of Library and Information Science defines version as follows:
1. A particular translation of the Bible or any of its parts.
2. An adaptation, or modification of a work for a purpose, use, or medium other than that for which the original was intended.
3. One of the variant forms of a legend, fairy tale, or other work of unknown or doubtful authorship.

4. The sharp-eyed will have noticed that instead of using the term manifestation, which I used in my series of articles, I have used the term edition. I do so in the interest of better communication.


7. For an extensive discussion of the assumptions about user behavior that underlie these functions (e.g., that users more often seek a particular work than a particular edition of a work, and that users often are unaware of the existence of various editions of a sought work), please see Part 1 of my series of articles entitled “What Is a Work?” It is interesting to note, however, that at the OCLC/NCSA Metadata Workshop, the participants—many of whom were not librarians—when asked to identify 13 key data elements, chose two (relation and source) that indicate relationships between and among documents. (See Discussion Paper 86 in the USMARC Archive, available on LC Marvel (gopher or telnet: //marvel.loc.gov).)


12. Character strings that are identical but not normalized can also link. Problems arise, however, when the same character strings represent different entities: for example, two different works that happen to have the same title. Here normalization can have the potential to differentiate two entities with the same character string: for example, using qualifiers such as date of original publication or release, form (textbook vs. work), publisher, and so forth.


16. Ibid., 35.

17. In March 1995 a discussion of the implications of creating a single virtual catalog took place on the Intercat list maintained by OCLC (intercat@oclc.org); discussants included Hunter Monroe, Charles Floyd Bearden, Cynthia Watters, Charley Pennell, and George M. Sinkankas. Concern has been raised over the potential loss of local practice that can have the effect of customizing a catalog for special local needs. Certainly, it would be desirable for the “single virtual catalog” to allow for different views of the catalog; for example, using different subject thesauri and classifications, even different languages. It has also been pointed out that the idea of a single catalog is not a new one, that each bibliographic database such as OCLC and RLIN was meant to create a single catalog for many libraries. (For that matter, from the time of Jewett, various attempts have been made to create a single national catalog. It is certainly true that in some sense each national database was meant to form a single national catalog. The problem is that the emphasis was on creating atomized records, not on creating a catalog in which records were bound together by the demonstration of relationships between them. Also, system design assumed as the primary purpose the creation of a warehouse of records from which “stock” could be ordered up using LCCNs, ISBNs, and the like as “stock numbers.” The systems were never well designed to support direct user access. On OCLC, for example, it is still difficult to do searches that result in large retrievals; there are no effective displays of multiple headings (e.g., displays that link the editions of a particular work); and displays of multiple bibliographic records are cumbersome, badly arranged (editions don’t come together), and difficult to scan through.

18. Although the USMARC format is primarily a communications format, it contains record design rules. Whether or not a particular piece of data is tagged and coded determines whether or not it is readily available for indexing and display in a given system.