**Abstract:**
Cataloging is compared to descriptive bibliography, to enumerative bibliography and abstracting and indexing services as well as to metadata created by Web search engines or by nonprofessionals at sites such as Amazon.com. These four types of metadata are compared with regard to object of the description, functions, scope, number of copies examined, collective vs. individual creation, standardization, authority control, evidence, amount of descriptive detail, degression, time span the data is intended to last, and degree of evaluation.

# CATALOGING COMPARED TO DESCRIPTIVE BIBLIOGRAPHY, ABSTRACTING AND INDEXING SERVICES, AND METADATA

Martha M. Yee

**SUMMARY.** Cataloging is compared to descriptive bibliography, to enumerative bibliography and abstracting and indexing services as well as to metadata created by Web search engines or by nonprofessionals at sites such as Amazon.com. These four types of metadata are compared with regard to object of the description, functions, scope, number of copies examined, collective vs. individual creation, standardization, authority control, evidence, amount of descriptive detail, degression, time span the data is intended to last, and degree of evaluation.

**KEYWORDS.** Cataloging, metadata, descriptive bibliography, enumerative bibliography, abstracting and indexing services, Web search engines, Google, Amazon.com, authority control, standardization, degression.

Martha M. Yee, MLS, Ph.D., is Cataloging Supervisor at the UCLA Film and Television Archive, Hollywood Offices, 1015 N. Cahuenga Boulevard, Hollywood, CA 90038 (E-mail: myee@ucla.edu).

## *INTRODUCTION*

Powerful library administrators have been making negative comparisons between library catalogs and bookseller catalogs, such as Amazon.com, and between library catalogs and Web search engines, such as Google (Marcum, 2005; Flecker, 2005) Commercial abstracting and indexing tools are also sometimes used as examples of access tools that are created more quickly and therefore provide more timely access to current literature than do library catalogs. Such comparisons are like comparing apples and oranges. These are all very different tools with very different goals and objectives. One purpose of this article is to point out these very different goals and objectives to provide a little more clarity to these discussions.

The term metadata is now being used in a very loose fashion by both librarians and computer scientists. While some argue that cataloging is a kind of metadata, it is clear from actual metadata projects that much metadata is very far from being cataloging. An important distinction between cataloging and other kinds of metadata is in danger of being lost. Another purpose of this article is to make some needed distinctions in this area.

The final purpose of this article is related to the fact that new cataloging rules are on the way; it is planned for *Resource description and access* to replace the *Anglo-American cataloguing rules*, 2nd ed. rev. (RDA). The designers of these new cataloging rules seem to have a desire to create rules that can be used by a much wider community than that of libraries in their creation of item-level cataloging records. It appears that an attempt will be made to design rules that can be used to create metadata for electronic documents, descriptions of rare books and museum objects, and collection-level records for unpublished materials in collections based on provenance. *Functional requirements*

*for bibliographic records* (FRBR), a model that underpins the new cataloging rules, is aiming at an even wider audience, including scholars of folklore, descriptive bibliographers, and computer scientist designers of projects that analyze raw data using computers with no human intervention whatsoever. The assumption seems to be that all these diverse areas have more in common than they have differences. There is a danger that some important distinctions will be lost in the rush to be so all-inclusive. For all these reasons, this paper was written to try to make these distinctions and raise some questions about where we are headed.

For the purposes of this paper, "cataloging" is defined as "the creation of catalogs"; and a "catalog" is defined as: a guide to a particular collection or aggregate of collections created using standards that govern both the choice and the labeling of data in such a way as to result in the choice of preferred names for authors, works, subjects, and disciplines, with provision for access under variant forms, such that a user who searches under any variant is led to everything of interest (all works on the subject sought, all works by the author sought, or all editions or expressions of the work sought). In any given catalog record, sufficient data is recorded to allow a user to identify and distinguish one edition or expression from another and to select a desired work or expression of a work.

## *I. CATALOGING COMPARED TO DESCRIPTIVE BIBLIOGRAPHY*

### Object of the description

Tanselle argues that the essential difference is that descriptive bibliography describes the ideal copy, while cataloging describes a particular copy. (Tanselle, 1977, p.

4) This is not completely accurate; the ordinary cataloger, in describing a particular copy, aims at describing both the edition or expression to which it belongs, as issued, and the work of which it is an edition or expression; a conscious effort is made not to include copy-specific details in the bibliographic record or at least to differentiate clearly which details are copy-specific. Tanselle himself recognizes this when he characterizes AACR collation as aiming at "extent of work" (Tanselle, 1977, p. 52), although "extent of expression" would be more accurate within the FRBR model.

**Functions**

Lubetzky quotes Pierce Butler--"A catalog is a bibliography of the books in a particular collection"-- and goes on to elaborate the function of the catalog as being that of serving as a guide to a particular collection, indicating relationships between items in the collection. (Lubetzky, p. 269) Note that modern cataloging is also called on to serve the functions of a national bibliography; this can create tensions between local and national/international needs and make the demonstration of relationships quite a complex task. The function of descriptive bibliography is essentially to record physical evidence that bears on printing history of particular texts and on printing and publishing history in general.

**Scope**

Most writers, including Lubetzky and Tanselle, emphasize that the scope of a catalog is a particular collection. (Lubetzky, p. 269; Tanselle, 1977, p. 10) This is an important distinction in that it points up the fact that, in contrast to catalogers,

bibliographers are free to define their scope as necessary to serve their particular purposes and the fact that catalogers can usually examine only those copies that are in their particular collections. However, in the last hundred years, it has become increasingly the case that catalogers catalog not just for their local catalog but also to contribute to the national bibliography.

**Number of copies examined**

Tanselle rejects this criterion in favor of the object of description as the essential difference. This seems to be a very important difference, however. It is only after examining *all* copies of a text that one can be sure of what the discriminatory details are in any particular case, i.e., the details that discriminate between particular states, issues, impressions, and editions of a text. Since this extent of research cannot be performed by catalogers, it is impossible for them not just to describe an ideal copy (in Tanselle's sense) but even to be sure that they are including in their descriptions the pieces of evidence that a bibliographer would need to tell exactly what is held in that particular collection. The cataloger can examine copies outside his or her collection only in surrogate form (i.e., in the form of cataloging records created by other catalogers in other collections) and only if surrogates for them are readily available. These surrogates will have been created under the same handicap; that is, their creators will not have been able to examine all extant copies in order to identify variations and determine the important discriminatory detail to describe them.

**Collective vs. individual creation**

It is extremely rare for even a single local catalog to be the product of the work of a single person. A descriptive bibliography is usually the product of a single person, who plans the scope, looks at all of the items described, decides on what kinds of evidence are relevant, and records the relevant evidence according to his or her own rules that are devised to fit the circumstances of that particular project. Surely, a descriptive bibliography is likely to be carried out in a more uniform and predictable fashion than a catalog, which is gradually built up over many years due to the efforts of countless catalogers. It should be noted, of course, that non-institutional catalogs, such as bookseller's or collector's catalogs, may be the work of a single person. Union catalogs, such as OCLC or the ESTC, are bound to have even more inconsistencies due to differences in cataloging policy from one institution to another.

**Standardization**

Because catalogs are created collectively, rules and other standards are written to try to attain some measure of uniformity from one cataloger to another and from one institution to another (to allow construction of union catalogs or national bibliographies). Because the data in a catalog is both gathered and presented in a predictable fashion, it is possible to design complex indexes and displays of records from many different institutions; for example, it is possible to display all the editions of a particular work, along with all of the works related to it and all of the works about it in response to a user's search that uses a variant of the author's name and a variant of the title of the work.

A descriptive bibliographer is much freer than a cataloger is to tailor-make descriptive rules and formats to fit the materials being described although the necessity to

communicate findings does impose some requirements for use of standard terminology, standard collation formulae (or at least comprehensible ones), etc.

**Authority control**

Authority control of names, work identifiers, and subjects is a special kind of standardization that is much more likely to be necessary in large cataloging and union cataloging projects than in descriptive bibliography projects. When authority control is applied effectively, users are enabled to find the works they seek under any variant of author name (Mark Twain vs. Samuel Clemens), variant of title (Tom Sawyer vs. Adventures of Tom Sawyer), or subject synonym they might use in their search (e.g. hypnosis vs. hypnotism). In addition, homonyms are differentiated (e.g. cold the disease vs. cold the temperature), people with the same name are differentiated (e.g. George Bush), and works with the same title are differentiated (e.g., 5th symphony). And finally, a structure is created that allows a user to broaden and narrow a search as desired.

**Evidence**

In ordinary cataloging, the title page and preliminaries are the primary evidence; and they are taken at face value, unless there is clear evidence of falsehood or error. The title page is only roughly transcribed. The descriptive bibliographer may use a much vaster array of evidence: signatures, running titles, paper (format, watermarks, chain lines, etc.), typography (quasi-facsimile transcription, identification of type faces, etc.), binding, ownership, textual variation (catchwords, fingerprints), etc. Tanselle is right to reject this as an essential criterion in differentiating between cataloging and descriptive

bibliography (Tanselle, 1977, p. 5) since rare book catalogers, to one degree or another, often record much of the same evidence that descriptive bibliographers do. However, in this criterion lies much of the distinction between ordinary cataloging and descriptive bibliography. Although edition is defined in AACR2 as all copies produced from substantially the same type image, the evidence employed to determine whether two items are the same edition are in practice essentially the title page transcription and the paging (at least in ordinary cataloging).

**Amount of descriptive detail**

Tanselle is right to reject this as a criterion as well (Tanselle, 1977, p. 5) as there is wide variation among catalogs and bibliographies alike. It is interesting, though, to contrast Bowers' assertion that a more fully detailed description should be given if fewer copies have been examined (Bowers, 1953, p. 5) with Schneider's assertion that entries in catalogs must be brief while entries in bibliographies must be full. (Schneider, p. 51) Schneider is talking primarily about enumerative bibliographies; the implication is that a fuller description may be needed to aid users in finding listed works whereas less detail may be needed in a catalog that can lead the user directly to the work itself (or at least an edition of it). Bowers seems to mean that the more investigation a bibliographer has done, the more assurance with which he or she can eliminate particular details as non-discriminatory. However, the message from these two bibliographers about the amount of detail appropriate to a catalog is contradictory!

**Degression**

Degression, in the modern sense (the use of only discriminatory detail for editions after the first), is used only in some book catalogs and descriptive bibliographies. Card and online catalogs historically have used unit records. One suspects that, in the case of the card catalog, this was due to a reluctance to modify a card once filed. In the online environment, it surely is connected with a shared cataloging environment in which records are copied into thousands of different catalogs.

Degression in Madan's sense (Madan) sounds like the modern cataloging concept of levels of description. For economic reasons, many catalogs contain minimal records for classes of materials identified by librarians as less important. The differences between rare book cataloging and ordinary cataloging might be considered to be due to degressive practice in Madan's sense of the term as well. Libraries identify particular materials as being rare or special; and these may be given descriptions using many of the techniques developed by descriptive bibliography, e.g., transcription of capitalization, punctuation, and/or line endings, transcription of full imprints without transposition, collation by signatures, notes on binding, typography, paper, etc.

It would seem that a certain amount of degression in this sense might be a wise thing. Elizabeth Tate has shown that the common 80:20 ratio operates on the number of works that ever appear in more than one edition. (Tate) Even though bibliographers protest that any book is interesting as a physical object regardless of its content, in fact, there has been a marked tendency in the past to invest bibliographical effort in investigating classic and influential literary texts that have gone into many editions and for which scholars and students require authoritative texts. Greg actually defined

bibliography as "the science of the material transmission of literary texts." (Greg, 1914, p. 83)

**Time span**

Descriptive bibliographies, like catalogs, are meant to be used in perpetuity by future generations and are not considered by their makers to be of transient or ephemeral interest. A catalog or a descriptive bibliography is designed to be part of the permanent documentation of humanity's cultural record.

**Evaluation**

It could be argued that works described using the techniques of descriptive bibliography are inherently valuable to scholarship, or the considerable investment in time and travel involved in creating the descriptive bibliography would not have been made by members of the scholarly community.

Not all works collected by libraries are valued that highly by the scholarly community, but works collected by a library still represent a considerable investment in processing time and in storage costs and have not been acquired lightly. There is a selection process involved--first in the publishing process (reputable publishers make value judgments in deciding which works to publish) and secondly in the library selection process. No library has an unlimited budget, and works are acquired only when they are considered useful to the community served by the library. Thus, only the more reliable and reputable works on a subject even appear in the library catalog. In addition, a library catalog allows a user to determine which writers have published the most works on a

subject of interest, a fact that should also bear some weight in the process of evaluating the best works on a subject. However, once works are included in a library (and its catalog), librarians are careful not to include evaluative comments in their actual descriptions. When librarians compose summary notes, they are scrupulously careful to avoid reviewing or critiquing works; instead the aim is to provide an objective description of the subject matter of the work.

**Section summary**

The essential criterion would seem to be that of function since from it all else follows. The function of a library catalog is to serve as the guide to a particular collection. The function of a descriptive bibliography is to record the findings of an investigation into printing and publishing history of a particular work or works. The latter is *not* a function of a library catalog. Aside from whether or not it should be the function of a library catalog, no particular collection is likely to have enough of the evidence to make such an investigation feasible. While other kinds of catalogs have different functions, e.g. the bookseller's catalog is designed to sell books, the private collector's catalog is designed to reflect glory on the collector, these catalogs, in listing the contents of particular collections, are as limited as library catalogs in how much evidence they can contain to support studies of printing and publishing history. All these catalogs can draw on the findings of descriptive bibliographers to identify and describe precisely the publications that they list, but the catalogers who construct them cannot carry out the investigations themselves.

Tanselle's criterion of ideal copy seems unsatisfactory for the following reason. One could argue that a library cataloger aims at describing the ideal copy as well in that he or she aims at excluding from the description (or at least clearly indicating) details which are specific to the copy cataloged and including those details which apply to the entire edition as issued. The cataloger differs from the descriptive bibliographer only in the methods of investigation and kinds of evidence that he or she brings to bear in this process.

The function of the library catalog is to serve as the guide to a particular collection. As such, it must be able to tell the users which subjects and which authors are represented in the collection, which works of those authors, and which editions of those works. It is in the last object that the work of the descriptive bibliographer can be tapped to allow us to more clearly identify and describe, e.g., identical editions hiding under differing title pages or different editions hiding under identical title pages.

Those users of the library who are not bibliographers are assumed to be more interested in the content of the works or editions they seek than in physical variation that does not affect content. The bibliographers themselves are interested in physical variation that does not affect content essentially because it is evidence that ultimately supports either the establishment of authoritative texts or the description of printing and publishing practice that can later support the establishment of authoritative texts. Thus, the bibliographer's work will ultimately serve the library's users; and librarians and bibliographers must cooperate to that end.

Unfortunately, before the bibliographer has done his or her work, the cataloger cannot be sure of what details are discriminatory in any particular case. In ordinary

cataloging, we rely on rough title page transcription and overall paging as the primary evidence for variation in edition. When this is not enough, we must wait until the bibliographer has done his or her work before we can do ours.

## *II. CATALOGING COMPARED TO ENUMERATIVE BIBLIOGRAPHY AND ABSTRACTING AND INDEXING SERVICES*

### Object of the description

The object of a description is a published item as in cataloging. Frequently, however, more analysis will be practiced. In other words, articles in journals, papers, short stories, poems, and the like in compilations, pamphlets, and other materials not generally cataloged in libraries will be included in an enumerative bibliography; abstracting and indexing services specialize in the journal articles that are not cataloged by libraries.

### Functions

Instead of a guide to a particular collection, an enumerative bibliography (sometimes called systematic or reference bibliography) or an abstracting and indexing service serves to aid the user interested in a particular subject to discover the existence of works on that subject; but generally no help is provided in locating an actual copy of a given publication. In fact, the American Film Institute catalogs (actually enumerative filmographies) are an example of an enumerative bibliography that lists titles that probably no longer exist; it is known that many films listed in the AFI catalogs no longer

exist from having been lost to nitrate film deterioration. (*American Film Institute Catalog*)

Some attempts have been made to provide more linking from abstracting and indexing services to available electronic copies in the on-line environment; but, so far, the mechanisms do not work very well. First, the mechanisms are very dependent on numerical identifiers such as ISSN's, which may be erroneous, treacherous, or non-existent. Second, these mechanisms must navigate an impenetrable tangle of rights management algorithms.

**Scope**

Scope is usually defined as limited to a particular subject or academic discipline as opposed to a catalog, the scope of which is defined by the collection to which it is a guide.

**Number of copies examined**

The creator of an abstract or an indexing entry or an entry in an enumerative bibliography is probably seeing only the copy in front of them and is unaware of other copies or versions although, depending on the subject area, some will list the various extant editions of a particular work of interest.

**Collective vs. individual creation**

An enumerative bibliography is likely to be created by a single person who will have defined a consistent scope and style. Abstracting and indexing services typically

employ staffs of abstracters and indexers, who will presumably be following some sort of in-house guidelines.

**Standardization**

Standardization will be employed within the enumerative bibliography project or in-house at the abstracting and indexing service, but it is rare for outside standards to be followed. Abstracting and indexing services generally see their function as being that of providing speedy access to current materials that will be of interest for only a short time. Their tools are generally not designed to be used in perpetuity by future generations and are considered by their makers to be of transient or ephemeral interest. Thus, for example, it is not uncommon for older records to be formatted or indexed differently than newer records with no attempt made to bring the records into sync. Because of the lack of standardization over time, it can be difficult to design anything other than very simple indexes and displays, especially if the index or display must cover records from more than one time period. Interoperable displays of records from more than one institution or indexing service would probably necessarily have to resemble a primitive Google-like display.

**Authority control**

As with standardization in general, there are few incentives for abstracting and indexing services to provide elaborate authority control. Speed of publication is valued over quality of output. They see their role as that of helping someone find out what a particular author wrote in this past year, not helping someone find out what he or she has

written over his or her entire career. Frequently, authors' forenames will be represented by initials with no attempt to distinguish one AB Smith from a different AB Smith. Users who search under variant author names or variant titles will be out of luck. There are the exceptional abstracting and indexing services that use, for example, Library of Congress subject headings (LCSH); however, they are unlikely to invest any resources in changing older records to match newer forms of heading when LCSH headings change over time. Often abstracting and indexing services do use a thesaurus since the emphasis is usually on subject access rather than on access to authors and works. However, the thesaurus usually bears no relationship to any other thesaurus or subject heading list; and, even internally, no effort will be made over time to keep older records in sync with newer ones when the thesaurus changes.

**Evidence**

Enumerative bibliographers are likely to use title pages and preliminaries in much the same way that ordinary library catalogers do. Often particular style manuals, such as the *Chicago Manual of Style*, will be followed. Abstracting and indexing services usually derive titles from the actual publications being indexed although liberties may be taken with authors' names as noted above.

**Amount of descriptive detail**

Probably only enough detail will be provided to allow the user to find the work described in a bookstore or a library; usually, little attention will be paid to

discriminatory details pertaining to expression or edition beyond explicit title page edition statements and publication dates.

**Degression**

Since there is usually little interest in providing a record of the various extant editions or versions of a listed work, degression is not a factor in either enumerative bibliography or abstracting or indexing services.

**Time span**

Tools are generally not designed to be used in perpetuity by future generations and are considered by their makers to be of transient or ephemeral interest.

**Evaluation**

See above for a disquisition on the degree to which library catalogs evaluate the works they contain. Enumerative bibliography exists primarily for the purpose of evaluation. An enumerative bibliography is usually a list of the best books on a subject created by a scholar or expert in the given subject. Abstracting and indexing services exercise evaluation at the point of determining which journals will be indexed. There is generally some effort to include only reputable journals and to exclude those that do not follow acceptable scholarly practices.

**Section summary**

The function of a catalog is to provide a permanent record of the works held in a particular collection with the underlying assumption that they will be held in perpetuity as part of the cultural record of humanity. Even if a title is withdrawn or discarded in a particular collection, this is done only when it is known to survive in another. The function of enumerative bibliography and of abstracting and indexing services is quite different; it is to provide temporary and timely access to citations to works on a subject without guaranteeing that the user will actually be able to obtain those works. The user must take a second step of visiting a library and searching the catalog (or, nowadays, searching the catalog over the Internet and authenticating themselves as users who are entitled to access electronic documents that are licensed by a library) in order to obtain the works themselves. Abstracting and indexing services especially are in the business of analyzing what is held by libraries; they do not collect the resources themselves.

## *III. CATALOGING COMPARED TO METADATA*

Many consider cataloging to be a type of metadata. In fact, by that broad definition of metadata (data about data), descriptive bibliography is metadata; abstracts and indexes are metadata; and, for that matter, encyclopedias, dictionaries and telephone books are metadata. Here, I would like to use a somewhat narrower definition of metadata, that of metadata that is neither cataloging, nor descriptive or enumerative bibliography, nor abstracts and indexes. I would like to consider here both metadata that is automatically generated by Web search engines and other computer software for which the data is the metadata (e.g. Google) and metadata that is created by humans who are not trained and educated in cataloging and authority control or any other kinds of standards

(e.g., Amazon.com or the metadata created by an aerospace engineer and attached to his paper on the Internet). Here I am using a definition of metadata similar to Campbell's: "Metadata is not designed or created by a specially-trained cohort of professionals who have a specific skill set and a common slate of objectives." (Campbell, p. 59) In other words, we define metadata in this paper in such a way as to exclude cataloging.

**Object of the description**

There can be quite a bit of fluidity in determination of the object of a description. One metadata creator will create a separate metadata record for each chapter of a given work or each page of a web site, and another will create a metadata record only for the work as a whole. Since many electronic documents are diverse conglomerations of image, text, and audio, different metadata creators will make different decisions about levels of analysis. When the computer is left to make the decision about the object of a description on its own, as in web search engines, the results can be inexplicable and bewildering. As Bernhard Eversberg points out, "Search engine operators cannot afford to disclose their methods of searching and indexing." (p. 11)

**Functions**

One of the main functions is often that of capturing the attention of potential users, much like the function of advertising. It has been said that the economy of the Internet is an economy of attention. Instead of vying for capital, competitors are vying for the scarce moments of attention that each Internet user has available on a given day.

**Scope**

For Amazon.com, the scope is current trade publications plus out-of-print titles available for purchase. It could be argued that the scope for metadata as narrowly defined above is the entire Internet. One could also argue that one of the reasons that the Internet is so appealing to people who have never been taught research skills is that it allows people to search for facts (Eversberg, p. 11) within disembodied texts, thereby losing all context. One of the main ways in which the metadata on the Internet differs from cataloging, descriptive bibliography, or abstracting and indexing service data is precisely this loss of context. Because of the loss of context, it provides information of often dubious authenticity and reliability and cannot provide the discourse that can lead to knowledge and finally wisdom in minds that are trained in research and critical thinking.

**Number of copies examined**

Only the copy in front of the metadata creator will be examined.

**Collective vs. individual creation**

This is the ultimate in collective creation.

**Standardization**

Little standardization is possible when metadata creators are not trained or educated in the principles of bibliographic organization, and none is possible when only computer manipulation of data is available as in the case of Google. As Bernhard Eversberg points out: "The database consists of nothing but large inverted files, derived

directly from the documents." (Eversburg, p. 12) Thus, complex indexes and displays become impossible as the data is completely unpredictable.

Remember that here we define metadata more narrowly than does, for example, Lynne Howarth, who vividly describes the cacophony that results when many different cataloging standards are followed in creating metadata, in her case broadly defined to include cataloging. (Howarth, p. 48-49) We have excluded metadata created according to standard rules from our definition of metadata in this paper.

**Authority control**

As with any kind of standardization, authority control is probably impossible. Users interested in precision and recall will succeed only if they have all synonyms or variant names and titles available for searching and if they are patient in wading through an ocean of homonyms. Since there are few such users in the world, precision and recall will be possible only for searches for authors, subjects, or works with unique and unchanging names. Thomas Mann writes vividly about how much users lose in the transition from library catalog to Web search engine.

Thus, in Amazon.com, a person who looks up Thomas Mann's work, *Library Research Models*, will be told that users who bought this work also bought *Death in Venice* and *The Magic Mountain*. A person interested in reading the great Chinese novel by Cao Xueqin will find some English language translations under *Dream of the Red Chamber*, some under *Dream of Red Mansions*, some under *The Story of the Stone*, and some under *Hung Lou Meng*, with nothing to guide the user back and forth among them.

**Evidence**

Since data is unlikely to be marked up in any standard way, it is hard for computer programs to find automatically the title or creator of a particular assemblage of bytes. Human mark-up may improve computer performance a bit in this regard; but, if, as humans often seem to do, they display one title prominently on their electronic document and then put a different title into their metadata, they are not doing much better.

**Amount of descriptive detail**

Most extant metadata systems seem to be trying to reduce the amount of descriptive detail collected compared to the amount collected in ordinary cataloging in libraries.

**Degression**

The problem of describing the differences among multiple editions or expressions of the same work has not been tackled yet in the metadata world although the librarian interested in science fiction might contemplate a future in which all documents have been digitized and are therefore available for electronic comparison. Such comparison might be able to signal to a descriptive bibliographer controlling the comparison process where there are significant variations in text, sound, or image sequencing. Only a human, however, would be able to identify which electronic documents represent the same work and should be compared in this fashion. With current Web search engines, it would be impossible to assemble all of the expressions and manifestations of a given work in the first place in order to conduct a subsequent comparison.

**Time span**

Currently, the major interest in Internet access seems to be focused on the extremely current and ephemeral. So far, because of its lack of organization and its bewildering magnitude, the Internet does not seem to be a good place to keep the permanent cultural record although bits of it are starting to show up there in the form of electronic texts.

**Evaluation**

Web search engines often provide what is called relevance ranking. This can mislead users into thinking that the most valuable and useful indexed Web sites are at the top of the tens of thousands of hits that result from just about any search done using a Web search engine. In fact, all that any relevance ranking algorithm can do is to weight certain words (e.g. those that occur least frequently) more heavily, add up the results, and put the sites with the highest score at the top. Because this process has nothing to do with the actual meaning of the terms being weighted and because language is one of the most complex of human capacities and one which is not available to computers, this can frequently lead to ludicrous results.

Google, perhaps the most popular of all search engines, is so successful because, instead of using the standard relevance ranking algorithms, it ranks sites based on their popularity. Google is capitalizing on an insight that lies behind citation indexing, a tool long familiar in the library world: one way to evaluate the quality of a document is to ask how many other people have cited it in their writings. Google ranks sites based on how

many other sites link to them. This can often give valuable results, but it is a method that should be used with caution. Popularity does not always equal excellence!

**Section summary**

These various kinds of metadata have an even more ephemeral function than does enumerative bibliography or an abstracting and indexing service; and, consequently, even less control is exerted to facilitate ready access. Users are left on their own with no human intervention to organize data for access. These systems profit from the ignorance of users who do not know about what they do not find and are so bedazzled by computers that they place extra value on whatever they did find because they did not have to travel anywhere to get it and did not have to humiliate themselves by asking for help from someone else to get it.

## *ARTICLE SUMMARY*

If librarianship is to continue to be a profession, it is going to have to find a way to provide the majority of our users with what they need (rather than what they think they want) and be content, for awhile perhaps, with pleasing only an educated minority that realizes what it needs. Otherwise, humanity is in danger of losing access to its cultural record.

There is a folk tale about a man who captures a leprechaun and is therefore entitled to the gold belonging to the leprechaun. Unfortunately for him, however, the gold is buried in the ground under a particular tree; and he has not brought a shovel with him. He does have a red ribbon with him, however; and he ties it around the tree and makes

the leprechaun promise not to remove it. When he returns with his shovel, of course, every tree in the forest has a red ribbon tied around it. Lest humanity is left to search through such a forest for its cultural treasure, we need to use our ingenuity to figure out how to defend the value of the type of human intervention for information organization that we carry out as a profession from the many advocates of the cheap fix that leaves human intervention out of the solution.

## ***WORKS CONSULTED***


1. Alden, John E. "Cataloging and Classification." In: *Rare Book Collections*. Ed. Richard Archer. ACRL Monograph no. 27. (Chicago: ALA, 1965), p. 65-73.
2. Alston, Robin and M.J. Jannetta. *Bibliography, Machine Readable Cataloging and the ESTC*. (London: British Library, 1978)
3. Alston, Robin. "Computers and Bibliography: the New Approach in ESTC." *Papers of the Bibliographical Society of America* 75:4 (1981), p. 371-389.
4. Alston, Robin. "The History and Description of Books." In: *Searching the Eighteenth Century*. (London: British Library, 1983), p. 15-27.
5. Alston, Robin and Richard Christophers. "Machine-readable Cataloguing and Early Printed Books." *International Cataloguing* 8:1 (Jan./Mar. 1979), p. 9-12; 8:2 (Apr./June 1979), p. 16-18.
6. *American Film Institute Catalog of Motion Pictures Produced in the United States*. (Berkeley, Calif.: University of California Press, 1971-1999)
7. *Anglo American Cataloguing Rules*. 2nd ed., 2002 rev. (Chicago: American Library Association, 2002)

8. Bowers, Fredson. "Bibliography, Pure Bibliography and Literary Studies." *Papers of the Bibliographical Society of America* 46 (1952), p. 186-208.

9. Bowers, Fredson. "The Function of Bibliography." *Library Trends* 7:4 (April 1959), p. 497-510.

10. Bowers, Fredson. "McKerrow, Greg and 'Substantive Edition.'" *The Library* 5th ser. 33:2 (June 1978), p. 83-107.

11. Bowers, Fredson. *Principles of Bibliographic Description*. (New York: Russell & Russell, 1949).

12. Bowers, Fredson. "Purposes of Descriptive Bibliography with Some Remarks on Methods." *The Library* 5th ser. 8:1 (March 1953), p. 1-22.

13. Butler, Pierce. "The Bibliographical Function of the Library." *Journal of Cataloging and Classification* 9:1 (March 1953), p. 3-11.

14. Campbell, D. Grant. "Metadata, Metaphor, and Metonymy." *Cataloging & Classification Quarterly* 40:3/4 (2005), p. 57-73.

15. "Critical Bibliography." *Encyclopedia of Library and Information Science.* (New York: Marcel Dekker, 1971), v. 6, p. 276-286.

16. Crump, Michael. "Introduction." In: *Searching the Eighteenth Century*. (London: British Library, 1983), p. 3-12.

17. Crump, Michael. "Stranger than Fiction: the Eighteenth Century True Story." In: *Searching the Eighteenth Century*. (London: British Library, 1983), p. 59-73.

18. Currier, Thomas Franklin. "What the Bibliographer Says to the Cataloger." *Catalogers' and Classifiers' Yearbook* 9 (1941), p. 21-37.

19. Davis, Stephen Paul. "Computer Technology as Applied to Rare Book Cataloging." *IFLA Journal* 10:2 (1984), p. 158-169.

20. Davis, Stephen Paul. "Night Thoughts on Automation and Rare Books." Typescript of lecture delivered July 30, 1984, Columbia School of Library Service Rare Book School.

21. "Descriptive Bibliography." *Encyclopedia of Library and Information Science.* (New York: Marcel Dekker, 1972), v. 7, p. 1-17.

22. Dunkin, Paul S. *Bibliography: Tiger or Fat Cat?* (Hamden, CT: Archon books, 1975).

23. Dunkin, Paul S. "On the Catalog Card for a Rare Book." *Library Quarterly* 16:1 (Jan. 1946), p. 50-56.

24. Eversberg, Bernhard. *On the Theory of Library Catalogs and Search Engines*. Supplement to a talk on Principles and Goals of Cataloging, given at the German Librarians' Annual Conference, Augsburg, 2002. Available on the Internet at: http://www.allegro_c.de/formate/tlcse.htm

25. "The First Phase: an Introduction to the Catalogue of the British Library Collections for ESTC." *Factotum* 4 (1983).

26. Flecker, Dale. *OPACs and Our Changing Environment*. (Washington, D.C.: Program for Cooperative Cataloging, January, 2005). Available on the Internet at: http://www.loc.gov/catdir/pcc/archive/opacfuture-flecker.ppt

27. Foxon, David F. *Thoughts on the History and Future of Bibliographical Description*. (University of California, 1970).

28. *Functional Requirements for Bibliographic Records*. (Munich: K.G. Saur, 1998)

29. *Functional Requirements for Bibliographic Records: Hype or Cure-all?* Ed. by Patrick le Boeuf. (New York: Haworth Information Press, 2005).

30. Greg, W. W. "Bibliography--an Apologia." (1932) In: *Collected Papers*. Ed. by J.C. Maxwell. (Oxford: Clarendon Press, 1966), p. 239-266.

31. Greg, W. W. "The Present Position of Bibliography." (1930) In: *Collected Papers*. Ed. by J.C. Maxwell. (Oxford: Clarendon Press, 1966), p. 207-225.

32. Greg, W. W. "What is Bibliography?" (1914) In: *Collected Papers*. Ed. by J.C. Maxwell. (Oxford: Clarendon Press, 1966), p. 75-88.

33. Hibberd, Lloyd. "Physical and Reference Bibliography." *Library* 5th ser. 20 (1965), p. 124-134.

34. Howarth, Lynne. "Metadata and Bibliographic Control: Soul-Mates or Two Solitudes?" *Cataloging & Classification Quarterly* 40:3/4 (2005), p. 37-56.

35. Laurence, Dan H. *A Portrait of the Author as a Bibliography*. (Washington, D.C.: Library of Congress, 1983).

36. Library of Congress. Processing Dept. *Studies of Descriptive Cataloging: a Report to the Librarian of Congress by the Director of the Processing Dept.* (Washington, D.C.P: Library of Congress, 1946).

37. Lubetzky, Seymour. *Writings on the Classical Art of Cataloging*. (Englewood, Colo.: Libraries Unlimited, 2001).

38. McKenzie, D. F. "Printers of the Mind: Some Notes on Bibliographical Theories and Printing-House Practices." *Studies in Bibliography* 22 (1969), p. 1-75.

39. Madan, Falconer. "Degressive Bibliography." *Transactions of the Bibliographical Society* 9 (1906-8), p. 53-65.

40. Mann, Thomas. *Survey of Library User Studies*. (Washington, D.C.: AFSCME 2910, July, 2005). Available on the Internet at: http://www.guild2910.org/google.htm

41. Mann, Thomas. *Will Google's Keyword Searching Eliminate the Need for LC Cataloging and Classification?* (Washington, D.C.: AFSCME 2910, 2005). Available on the Internet at: http://www.guild2910.org/searching.htm

42. Marcum, Deanna B. *The Future of Cataloging*. Address to Ebsco Leadership Seminar, Boston, Mass. (Washington, D.C.: Library of Congress, January 16, 2005). Available on the Internet at: http://www.loc.gov/library/reports/CatalogingSpeech.pdf

43. Mitchell, Jim. "Investigating False Imprints." In: *Searching the Eighteenth Century*. (London: British Library, 1983), p. 43-58.

44. Osborn, Andrew D. "The Crisis in Cataloging." In: *Readings in Library Cataloguing*. Ed. and introd. by R.K. Olding (London: Crosby Lockwood & Son, 1966), p. 225-41.

45. Osborn, Andrew D. "Relation Between Cataloguing Principles and Principles Applicable to Other Forms of Bibliographical Work." In: International Conference on Cataloguing Principles (1961 : Paris). *Working Papers*. Working Paper no. 1.

46. Pollard, Alfred W. "Relations of Bibliography and Cataloguing." In: International Library Conference (2nd : 1897 : London). *Transactions and Proceedings*, p. 63-66.

47. Pouncey, Lorene. "The Fallacy of the Ideal Copy." *The Library*. 5th ser. 33:2 (June, 1978), p. 108-118.

48. *RDA: Resource Description and Access: Prospectus*. Available on the Internet at: http://www.collectionscanada.ca/jsc/rdaprospectus.html

49. Rota, Anthony. *Points at Issue: A Bookseller Looks at Bibliography*. (Washington, D.C.: Library of Congress, 1984).

50. Schneider, Georg. *Theory and History of Bibliography*. Trans. Ralph R. Shaw. (New York: Columbia University Press, 1934).

51. Shatford, Sara. "Catalogue Access to Early Children's Books." *AB* (Nov. 12, 1984), p. 3411-3416.

52. Tanselle, G. Thomas. "Descriptive Bibliography and Library Cataloguing." *Studies in Bibliography* 30 (1977), p. 1-56.

53. Tanselle, G. Thomas. "The Editorial Problem of Final Authorial Intention." *Studies in Bibliography* 29 (1976), p. 167-211.

54. Tate, Elizabeth L. "Main Entries and Citations: One Test of the Revised Cataloging Code." *Library Quarterly* 33 (1963), p. 172-191.

55. Taylor, Archer J. *Book Catalogues: Their Varieties and Uses*. (Chicago: Newberry Library, 1957).

56. Taylor, Archer J. *Catalogues of Rare Books: a Chapter in Bibliographic History*. (University of Kanasas Publications Library Series no. 5) (University of Kansas Libraries, 1958).

57. Taylor, Arlene. *The Organization of Information*. 2nd ed. (Westport, Conn.: Libraries Unlimited, 2004), p. 29-48.

58. Taylor, Earl R. "Cataloguing and Computers: Librarians and Cyberphobia." *Papers of the Bibliographical Society of America* 75:4 (1981), p. 392-400.

59. *Times Literary Supplement*.

July 7, 1966. "Review of William B. Todd's *A Bibliography of Edmund Burke*."

Aug. 4, 1966, "Letter from John Carter re: review."

Sept. 1, 1966, "Letter from William B. Todd."

Sept. 22, 1966. "Letter from Tanselle."

60. Todd, William B. "The ESTC as Viewed by Administrators and Scholars." *Papers of the Bibliographical Society of America* 75:4 (1981), p. 389-392.

61. Tolman, Frank L. "Bibliography and Cataloging: Some Affinities and Contrasts." *Public Libraries* 10:3 (March, 1905), p. 119-122.

62. Wright, Wyllis E. "Some Fundamental Principles in Cataloging." *Catalogers' and Classifiers' Yearbook* 7 (1938), p. 26-39.