Minimal Level Cataloging: What Does It Mean for Maps in the Contexts of Card Catalogs, Online Catalogs, and Digital Libraries?

Zorana Ercegovac
Department of Library and Information Science, Graduate School of Education & Information Studies, 300 Circle Drive North, Mailbox 951521, University of California, Los Angeles, Los Angeles, CA 90095-1521.
E-mail: zercegov@ucla.edu

In this article, we examine some of the proposals which have dealt with the problems in cataloging in two different technological contexts: Printed-card catalogs and online catalogs. We first look at some of the measures which attempted to deal with the "crisis in cataloging" at the Library of Congress in the 1940s. Then we address some of the current problems in cataloging in the era of online public access catalogs (OPACs). In particular, we discuss the extent to which minimal-level cataloging, as defined by the Anglo-American Cataloguing Rules and implemented in the Online Computer Library Center (OCLC) database, works for maps. Our discussion is organized around two main groups of access points: Controlled vocabulary data elements, including name data elements, and free-text data elements. In closing, we discuss prospects that the next generation of online catalogs using the Z39.50 protocol and SGML format might offer to minimal-level cataloging for maps.

Introduction

One of today's solutions to deal with many of the current cataloging problems, similar to those facing the Library of Congress in the 1940s, is found in the proposal commonly referred to as "minimal-level cataloging." MLC is also known under the headings of short cataloging, brief record cataloging, simplified cataloging or K-level input, and less-than-full cataloging. In 1978, Library of Congress introduced MLC as a means to combat crisis in cataloging, described shortly, stating that some access is better than none. Ten years later, it was reiterated that while we have not learned as yet whether the provision for briefer records (of possibly more materials) is better than none, we read that MLC, in general, "has been highly effective in providing such access through the online retrieval system by means of such elements of bibliographic description as author, title, and series" (Library of Congress, 1990, p. 48). Has the power of online catalogs, to some extent, compensated for minimal records of maps? In 1994, a Task Group appointed by the Cooperative Cataloging Council, now known as the Program for Cooperative Cataloging (PCC), was charged to specify a minimum set of data elements for the core record bibliographic standards for several formats. The core level cataloging would facilitate a "national cooperative cataloging program that can help provide 'faster, better, cheaper' cataloging" (http://lcweb.loc.gov/catdir/pcc/).

In this article, we examine proposals which attempted to deal with the problems in cataloging in two different technological contexts, printed-card catalogs and online catalogs. We first look at some of the measures which dealt with the crisis in cataloging at the Library of Congress in the 1940s. Then, we review some of the problems in cataloging in our own era of online public access catalogs (OPACs). In particular, we discuss the extent to which minimal-level cataloging works for maps, as defined by the Anglo-American Cataloguing Rules (AACR2, 1988) and implemented in the Online Computer Library Center (OCLC) database. Our discussion is organized around two main groups of access points: Controlled vocabulary data elements including name data elements, and free-text data elements.

MLC was designed as a cost effective means of providing access 1) to items worth retaining in the Library's collections but not worth the expense of full cataloging and 2) to a very large number of items in the arrearages that were unavailable to users and for which the completion of full cataloging was highly unlikely. (Library of Congress, 1987, pp. 40–41)

Received May 9, 1997; revised July 16, 1997; accepted July 16, 1997.

© 1998 John Wiley & Sons, Inc.
It proposes to build graphical user interface (GUI)-based online catalogs that typically use the Z39.50 Information Retrieval Protocol (ANSI, 1995) and the Standard Generalized Markup Language (Goldfarb, 1990). If we wrapped MLC records around the SGML format, could some of the shortcomings of the minimal description input be, to some extent, alleviated? What access points are desirable and necessary for effective retrieval of maps in the networked online library catalogs of large collections?

We seem to have come full cycle, or at least back and forth on a swinging pendulum, from a “legalistic” approach to cataloging in the 1940s to a “pragmatic” approach to cataloging in the 1960s, and back to the “legalistic” approach. The Anglo-American Cataloguing Rules (AACR2, 1988) as well as additional conventions in OCLC-MARC formats suggest that we are reaching another crisis in cataloging. While we have no answers to the question of how to use the power of next-generation library catalogs in order to reduce the costs of cataloging and enhance the usefulness of the catalog to the users, in closing, we suggest a number of research questions to be considered for further examination of minimal-level cataloging for maps.

The “Crisis in Cataloging” at the Library of Congress and Means to Deal with it

In 1939, Archibald MacLeish, the Librarian of Congress at the time, convened two committees of experts to investigate the situation of processing operations.

One of the two committees, the Coordinating Committee of the Processing Divisions, reported that an unprocessed “arrearage” existed of 1,670,161 volumes, out of a total estimated collection of 5,800,000 volumes, and that the backlog was increasing at a rate of 30,000 volumes a year (MacLeish, 1944). The other committee of experts, the Librarian’s Committee, found a rather low throughput per technical assistant in the Catalog Division with corresponding high costs to catalog each item (Library of Congress, 1940). To deal with some of these problems, the Librarian’s Committee made the following four recommendations: Administrative reorganization, simplified cataloging, cooperative cataloging, and a change in management methods to control operational costs. While each of these four measures were addressed earlier in the general context (Yee, 1987), this article focuses on “simplified cataloging,” and particularly, in the context of maps.

Simplified Cataloging

The Librarian’s Committee (Library of Congress, 1940), chaired by Carleton Joeckel, recommended that the following three measures be taken: First, that the cataloging rules be made less “detailed.” Second, that three levels of detail be established, that is, “detailed,” “standard,” and “simplified” levels of cataloging. Finally, the Committee recommended that rules concerning corporate names and series be re-examined.

Joeckel asked Andrew Osborne to report findings of the Committee’s investigation, which he did in his landmark article, “The Crisis in Cataloging” (Osborne, 1941). Osborne strongly criticized what he titled the “legalistic” approach to cataloging which attempts to create the added bulk of rules making cataloging “. . . elaborate, highly technical, a skill too often existing in and for itself” (Osborne 1941, p. 393). By contrast, Osborne recommended a few simple rules for catalogers trained to use judgment, an approach he termed “pragmatic” cataloging. For Cutter (1904), cataloging was an art, not a science; no rules could take the place of experience and good judgment, but some of the results of experience might be best indicated by rules. For Osborne, cataloging is an art, a technical art: Its basic rules are few and simple, and as such, it is a delightful art to practice. By criticizing “legalists,” Osborne prepared a groundwork for Lubetzky’s work (1960, 1969). By rigorous and systematic investigation of each rule, one by one, Lubetzky’s inquiry resulted in a design of the “pragmatic” code which laid down guiding principles upon which all rules were based.

Three Levels of Cataloging

Harking back to Cutter’s “general remarks” in his Rules for a Dictionary Catalog (1904), the pattern of “short,” “medium,” and “full” cataloging was revived by the Committee and, as noted earlier, restated as one of the measures to deal with cataloging problems. With regard to the second recommendation by the Librarian’s Committee to allow for three levels of bibliographic detail, the Committee’s Report (Library of Congress, 1940) outlined a program for developing a standard for short cataloging.

The Committee recommended a number of tentative rules including:

i) Research time, restricting it to 15 minutes, required in establishing authors’ names:
ii) treatment of some very extensive titles requiring that judgment be exercised in curtailing them;
iii) treatment of collation (i.e., now known as “physical description” statement of an item) and suggesting that collation statement be abbreviated to essential items only;
iv) omission of series entries, content notes, repeated author’s statement in the title, and added entries such as those for titles, editors, translators, and such;
v) one subject heading be given with preference to specific headings.

These tentative rules were to be applied to mimeographed and much other near-print material; textbooks below the college level; juvenile literature and picture books, except where literary or bibliographical reasons call for full treat-
ment; books, and especially pamphlets of slight interest and importance (Library of Congress, 1940).

The phrase “standard” cataloging requires clarification. At the turn of the century, an Advisory Committee on Cataloging Rules appointed by the American Library Association Publishing Board (Henderson, 1976), set the pattern for future ALA codes by deciding that the plan for the code should be “carried out for the large library of scholarly character” (Kroeger, 1901, pp. 211–212). In this view then, the following interpretation of “standard” cataloging is quoted from a response to an inquiry from abroad regarding the North American system:

By the word “cataloging” we mean the preparation of printed or other duplicated entries according to standard rules whereby these entries are rendered usable for all library and bibliographic purposes. This involves the authoritative investigation of author headings by research and correspondence; transcription of title, collation, contents, etc., according to standard rules; duplicate entry under joint authors, editors, translators, subjects, title, catchword title, etc.; and the integration of all these entries by a network of cross references (“see” and “see also”), tracers, authority cards, guide cards, etc., which converts the catalog from a mere assemblage of cards into an organic unit designed to meet all needs of all classes of searchers—in other words, a universal apparatus as distinguished from mere check lists, want lists, “location” lists, inventories, and similar partial or temporary lists designed to meet only occasional or limited needs. (Library of Congress, 1935, p. 241)

About the same time, in the provocative article titled “The Forrest of Pencils,” Pierson wrote:

Catalogers are not only transcribers. They are seekers after truth. It is argued by many . . . that all that is needed for a catalog entry is the author’s name as it appears on the title-page, a brief title, synthetic collation, and size . . . a mechanized process for quantity production. Experience has shown that such cataloging is the most expensive . . . for it leads to endless confusion. It is actually cheaper . . . to take time to make a full and complete entry than to dispatch a hasty, ill-considered one . . . . The few (catalogers and bibliographers) that are painstakingly accurate remain and are treasured . . . . Printed catalog cards made for distribution to other libraries to save duplication of work become an economy only when they are complete and correct in every detail; time spent in making simple, unverified entries . . . is time and money thrown away. (Pierson, 1934, p. 313)

Obviously, Cronin, who headed the Coordinating Committee, had in mind the pattern for future ALA codes as set by the Advisory Committee (Kroeger, 1901) and Pierson’s remarks (1934) when he wrote:

The high cost of cataloging at Library of Congress is necessarily so because of the high scholarly standards of cataloguing practice of the Catalogue Division . . . . The present high standard of cataloguing practice should be continued and there should be no lowering of standards. (Library of Congress, 1940)

In 1947, the Rules for Descriptive Cataloging (RDC, preliminary edition; Library of Congress, 1947) were drafted but did not codify the levels of cataloging. The same year, Luther Evans, the Librarian of Congress, issued General Order No. 1340 on “cataloging categories.” These were:

1 (a) Material of primary importance to be cataloged individually and fully;
(b) Material of secondary importance to be cataloged individually but briefly;
2 (a) Groups of material of primary significance as groups to be cataloged by individual entries describing the particular collections; and
(b) Material of minor importance by a given author or on a given subject to be cataloged collectively by form cards. (Library of Congress, 1952, p. 11)

The lack of codification of what actually was meant by “individually but briefly” and “individually but fully” remained undefined with the exception of individual memos published in a slim “Supplement” in the RDC (1952, pp. 12–13, 16–19). These memos instructed catalogers as to what bibliographic data were to be treated, the amount of research to be performed, the types of materials to be identified and described, and the “no conflict” policy pertaining to establishing new personal name entries “in the form given in the work being cataloged without further search.” The 1949 Rules would apply in full only to basic reference and research tools, scholarly works, and rare books. Most other material, including maps, was to receive “limited cataloging” with simplified collation and notes.

So far, we directed our attention to the historical perspective of MLC in general and in the context of printed card catalogs. We now look more closely at the prospects MLC has in the online environment, and, specifically its implication to cartographic materials.

Minimal Level Cataloging in the Era of Online Library Catalogs

The recent growth of online bibliographic networks (e.g., OCLC, WLN, RLIN, UTLAS), the escalating cost of cataloging entries, the use of minimal-level cataloging in Machine Readable Cataloging (MARC) records, the Cataloging-In-Publication (CIP) Program, have all highlighted the need for empirical research on the usefulness of MLC and on the relative merits of varying levels of detail for describing and identifying materials. As stated earlier, the Library of Congress introduced the standard for MLC (Library of Congress, 1978, p. 695) “to provide a standardized brief record for materials that might otherwise not be cataloged at all” with a possibility for the records to be completed subsequently.
The Library of Congress issued “projected uses of minimal records” as follows:

Minimal records may be used for the cataloging of ephemeral or marginal materials to put them under bibliographic control to the extent of providing the descriptive block and certain access points. The minimal record is not intended to take place of a full cataloging record, but to be used to provide a standardized brief record for materials that might otherwise not be cataloged at all, and therefore could not be searched for retrieval. Furthermore, minimal records might be useful for the control of books in process of being cataloged, as a means of making these books accessible, either for technical processing or reference use. In addition, minimal records could be input to a network and could be contributed to LC’s database and redistributed by LC via its MARC distribution services. (Library of Congress, 1978, p. 695).

In 1983, Library of Congress began offering a subscription to its own MLC through MARC Distribution Service. Typically, the records lacked subject headings, notes, assignment of access points especially important to the retrieval of some non-book materials, classification numbers, and were not subject to the usual editorial procedures.

There is evidence attesting that most academic library users at the University of Bath are not interested in more than five bibliographic elements (e.g., Seal 1983; Stamm 1996). However, there is little empirical evidence to study the long-term savings in moving from full to minimal description. When Weilh (1979) discussed some of the problems and prospects of MLC for non-book cataloging, she stated that administrators should question the belief that certain types of categories, such as slides and other non-book items, are used only by few of the patrons. Actually, she found that the slides were uncataloged and stored in the workroom; thus, people could not use what they did not know existed. Research is required to establish needs of various classes of library users; how often they would make use of these data elements, and the purposes for which they would be using them. Lambrecht asks:

If 10 percent of users and staff of the most sophisticated research libraries are unable to locate an item, distinguish it from a very similar item, or determine needed details about it, has the catalog served them well enough? Cutter . . . would not believe it had. (Lambrecht, 1992, p. 20)

Lambrecht reminds us that the library literature of the period 1982–1991 has seen numerous proposals which call for the reexamination of the costs and benefits of less-than-full cataloging; little progress, however, has been made in this direction to this day.

Gorman (1984) summarized findings of a survey of the ALA’s Technical Services Directors of Large Research Libraries on the use of “official” and “unofficial” types of MLC, and the methods of distributing those minimal records. While Gorman admits that the survey was in “no sense scientific,” it does convey information worth pondering and does suggest further lines of inquiry. For instance, the questionnaire found that eight institutions out of 25 large research libraries used National Level Minimal Bibliographic Record (NLMBR) as defined in documents issued by LC; and that the categories of materials cataloged according to the standard were diverse. Some of the types of materials cataloged at the minimal level were pamphlets, theses, monographs older than 4 years, movie, radio, and TV scripts, art exhibition catalogs, audiocassettes, foreign monographs, and older materials. Furthermore, some of these types of materials selected for MLC treatment are ones which are often poorly represented in printed bibliographies or not represented at all. Intner elaborates:

We are especially prone to give MLC to nonbook media items kept in closed stacks that also won’t be found in network databases and require more specialized and time consuming original work involving viewing, listening, or running on computers to be done properly. (Intner, 1994, p. 5)

In 1991, Lambrecht surveyed 22 national cataloging agencies to assess their practices and attitudes toward MLC (Lambrecht, 1992). The survey indicated that only seven of the 22 national agencies included all mandatory ISBD data elements in all of their cataloged records and that only four of the 20 elements designated as mandatory in the ISBDs were considered of vital importance. Furthermore, the statement of responsibility, place of publication, series title proper, and edition statement of responsibility were considered the bare bones of description for new ISBD(MLC) practices. Going back to the main purpose of MLC, one wonders if the economic benefits of MLC might be more than offset by the increased costs of providing local reference services to collections with inadequate bibliographic control (Rhee, 1986; Swank, 1956).

In some of the most advanced online catalogs of large collections, users will typically retrieve some useful items from the collections in response to a given query, but they have no way of knowing what has not been retrieved. Handman persuasively argues that for media collections that are becoming “bibliographically accessible from remote locations, the lack of effective subject access means . . . valuable resources potentially lost on the shelves forever” (Handman, 1991, p. 45). He shares anecdotal information that MLC records when searched on the University of California systemwide online catalog, MELVYL®, are virtually useless to the majority of users. This is in sharp contrast to the belief which was expressed 10 years ago: “Although represented by a less-than-full catalog record, it was judged that the power of the online retrieval system would, to some extent, compensate” (Library of Congress, 1987, pp. 40–41). Handman again:

Even under the best of circumstances provided by the MELVYL® catalog, including to-the-hilt standard cataloging; keyword and Boolean searching; and format, date,
language, and location limitation capabilities, our patrons frequently find themselves rudderless and listing in stormy bibliographic seas. (Handman, 1991)

Other writers have also discussed limits with traditional cataloging of cartographic materials, especially in the area of providing access to geographic data by geographic name subject headings (Holmes, 1990). Holmes identifies numerous reasons why map-based computerized geographic systems are necessary. He makes the point that such a system is a superset of the Geographic Information Systems (GIS); while the former provides access to geographic information available in all types of media, GIS software displays and manipulates digital geographic information for a specific site (Holmes, 1990, p. 39).

While some of the writers have offered different proposals to combat the crisis in cataloging in the 1970s and 1980s suggesting extension of the Linked System Project to the online exchange of bibliographic records among networks, and establishing shared cataloging to cover special categories inadequately cataloged, others have experimented with their own in-house standards of what minimal-level cataloging should be (e.g., Stanford’s Freund, 1990; Northwestern’s Horny, 1991; Stamm, 1996; Michigan’s Marko & von Wahlde, 1986).

In 1993, in an effort to expand the scope of national cooperative cataloging programs, the Cooperative Cataloging Council (made up of representatives from OCLC, Library of Congress, and Research Libraries Group) appointed a task force to recommend cost effective cataloging standards aimed to be acceptable to as large a segment of the cooperative community as possible. One recommendation, referred to a core level bibliographic standard, presents a flexible, less-than-full cataloging standard that emphasizes local needs and the exercise of judgment. This was developed in response to the perceived inadequacies of the standards for minimal level and full level cataloging. Minimal level cataloguing is widely perceived to be cost effective but of limited utility, whereas full level cataloging is frequently very expensive. The core level standard represents a third option that has some of the advantages of both levels of cataloging with fewer drawbacks (Cromwell, 1994). The effectiveness of the Core Record Pilot Project is being investigated by Kelley and Schottlaender (1996). While core standards have been proposed and developed for several formats (e.g., books, music, audiovisual materials), and more are under development, this author is not aware of the core record standard for cartographic materials (P. Rau of the Library of Congress, personal communication, May 1997). Instead, less-than-full level cataloging is used to catalog all new receipts except for certain categories (Library of Congress, 1991, p. B1).

Has MLC Worked for Maps?

It is commonly assumed that an integrated catalog containing records for different physical formats (e.g., books, serials, maps, scores, visual materials, sound recordings) might benefit by applying data elements uniformly, across all media. However, since maps are not books, the data elements describing maps will differ from those describing books. It follows that, at least in theory, minimal-level cataloging for describing maps will differ from minimal-level cataloging for describing books. In this section, we look at the MLC conventions as they apply to maps and discuss its implications in the context of map retrieval.

AACR2 provides three levels of detail in the description as follows: First-level of description, or “minimal” (AACR2, 1988; rev. rule 1.0D1); second-level or “core” (AACR2, 1988; rev. rule 1.0D2); and third-level of description or “full” (AACR2, 1988; rev. rule 1.0D3). However, the brevity of the first level of description is achieved by omitting certain data elements that are the most essential part of map description. When the Committee's Report recommended a number of tentative rules for developing a standard for short cataloging, these rules were to be applied to mimeographed and other near-print material; textbooks below the college level; juvenile literature and picture books, except where literary or bibliographical reasons call for full treatment; books and especially pamphlets of slight interest and importance (Library of Congress, 1940). The rules called for: Restricting research time to 15 minutes for establishing authors’ names; curtailing very extensive titles; abbreviating the physical description statement; omitting series entries, content notes, repeated author’s statement in the title, and added entries such as those for titles, editors, and translators. The rules also restricted subject analysis to only one subject heading per record. If these rules were automatically applied to cartographic materials, the brevity would be achieved at the significant cost of clarity and accuracy. While we have no empirical evidence to support Pierson’s statement cited earlier in this article, it is worth reminding us that

It is actually cheaper . . . to take time to make a full and complete entry than to dispatch a haste, ill-considered one . . . . Printed catalog cards made for distribution to other libraries to save duplication of work become an economy only when they are complete and correct in every detail; time spent in making simple, unverified entries . . . is time and money thrown away. (Pierson, 1934, p. 313)

Fifty years later, a typical MLC skeleton map entry lacks second and subsequent statements of responsibility, as well as names of second and subsequent publishers; added entries (7xx), except for series added entries (8xx), are rarely made even for items entered under title. Dimension statement in the physical description area (300 subfield c) is used sparingly. In reality, the size of a map is important because it may help distinguish one map from another. The statement of coordinates, which is per-
haps the most exciting contribution of the ISBD (CM), is given optionally (International Federation, 1977). Combined with other numerical data, those of scale, projection, and date, the statement of coordinates would allow the user to search the map by precisely specifying the latitude and longitude of the area covered by the map in the online environment. All notes (5xx fields) are used rarely. Finally, only limited subject analysis is applied (Library of Congress, 1987, p. 43). Given that most of these data elements define a map, we need to re-examine the ways maps are searched for and used in order to define levels of description which would be both cost effective and useful to our users.

To answer the questions: (i) Is the provision for briefer records of more maps better than none, and (ii) what are the economic implications of MLC in the context of cartographic material, requires research which goes beyond this article. Here, we give examples of the effect MLC has had on some of the data elements. We group data elements into two main groups of access points: 1) Controlled vocabulary data elements, including name data elements; and 2) uncontrolled (free-text) data elements.

**Controlled Vocabulary Data Elements**

In this group, we discuss mathematical data area, area related classification fields, as well as topical and geographic subject headings. In addition, name data elements will also be discussed.

*The mathematical data area* (a 255 MARC field) includes three subfields specific only to cartographic material: Statement of scale (255 subfield a), statement of projection (255 subfield b), and statement of coordinates (255 subfield c). In the minimal-level description, catalogers are instructed to include projection statement(s) if given, and not to include coordinates. The equivalent to the bibliographic 255 field is a computer searchable field (034 MARC field) which codes information about scale, coordinates, and equinox, but *no projection*. One might argue that giving a statement of coordinates would save time and resources, since catalogers would not have to conform to various forms of geographic place names which are treated differently in different editions of AACR2 codes. Additional savings could be also achieved, since catalogers would not have to verify name authority files under the name of a political area shown on a map (e.g., for the former USSR, Yugoslavia, Germany, Czechoslovakia) and then to check the latest *Library of Congress Subject Heading* changes to see if there is a difference. A recent addition of the *geospatial reference data* (342 MARC field), to be used in conjunction with the Federal Geographic Data Committee FGDC’s *Content Standards for Digital Geospatial Metadata* (1995), is a direct result of the converging communication standards for the representation of geospatial datasets. However, catalogers are required to enter this field, if applicable, in full-level cataloging only. Provided that the 342 field is entered and coded in the OCLC/MARC record, this field has the potential to provide access to a host of graphical datasets: Via geographical information systems (GIS) which use FGDC’s Standard, or via World Wide Web (WWW) pages which have adopted the Standard Generalized Markup Language (SGML). In this case, short-term “ savings” that are achieved locally through minimal-level cataloging using a single bibliographic network might be a shortsighted decision, especially if it hampers access to a variety of remote multimedia repositories beyond a single database (e.g., OCLC/MARC bibliographic database).

Other area related numeric fields are classification codes, including LC call number (050 MARC field), geographic classification code for geographic elements of subject headings—each links up with every 6xx field (052 MARC field), geographic area code (043 MARC field), and locally assigned call number (990 MARC field). However, all these codes are to be used optionally for the MLC map input to the OCLC/MARC database. Because of the importance of geographic classification...
codes, less-than-full level cataloging instructs map catalogers to provide full classification for less-than-full records: To apply geographic classification codes (tag 052) for geographic elements of subject headings traced on less-than-full records according to normal conventions and routines (Library of Congress, 1991, p. B.1).

Subject headings. This brings us to the question of what is meant by the term “subject” as applied to maps. Cutter (1904) spoke of subjects as the theme(s) of the book, whether stated in the title or not. Recently, the term “theme keyword” is used by the Federal Geographic Data Committee (FGDC) in their Content Standards for Spatial Metadata (1995) as “common-use word or phrase used to describe the thematic content of a data set.” Larsgaard (E-mail communication, January 24, 1994) explains:

“What the attempt is in having both Theme Key Word and Geographic Key Word is that the first is for overall topics, such as geology, vegetation, and so forth, that pertain to the spatial-data item as a whole; the latter is intended for getting at features shown in that item. For example, on the USGS quadrangle of Santa Barbara, features would include beach, ocean, streams, hills, and so on.

Boggs and Lewis (1945, p. 19) defined cartographic subjects as follows: “The term subject as used in map classification and cataloging relates to the content of a map (apart from the area covered by the map), or to the type of information which it portrays.” This definition makes the distinction between topical subjects and geographical subjects. First, there are maps that are useful mainly for their display of features such as landmarks, distribution of refiners, the utilization of land, or the direction of currents (a 650 MARC field). Second, there are maps primarily designed to show physical features of a geographical region (a 651 MARC field). If we define the subject of a map as that which a map is about, then we have basically two types of subjects represented in the majority of maps. For example, a map about earthquake intensity in California, or skiing in Montana is an example of the topical subject heading:

650 0 Earthquake intensity $z California $z Sausalito $x Maps.
650 0 Skis and skiing $z Montana $ Garnet Range $x Maps.

A map representing the Yosemite National Park is about a specific geographic region. The Yosemite National Park is the type of information it shows, thus the Park is the subject or the content of this map. The following examples show geographical subject headings:

651 0 Yosemite National Park (Calif.) $x Maps.
651 0 Nevada $x Maps. Topographic.

Making the distinction between these two types of subjects is important because different users at different times will desire access to one type of subject and/or another type of subject.

Studies have illustrated that these two subject fields are the most frequently used entry headings for maps. Of the two subject fields, more end-users are likely to want to see a geological map of California rather than all geological maps and then to try to find those that pertain to the area (e.g., American Geographical Society, 1969; Fink, 1962; Larsgaard, 1987; Special Libraries Association, 1956). All of these studies suggest the area as the organizing principle for maps; thus, geographers and cartographers have traditionally organized map collections by means of spatial data—the area main entry heading. In contrast, the Anglo-American cataloging tradition, as noted earlier, has advocated that maps be organized and entered under their authors. Specifically, AACR2 states that the main entry for maps is under the person or corporate body primarily responsible for the content and design of maps (Library of Congress, 1981, pp. 21–24, per Rule 21.1B2f). The question, what is the most significant and useful order to display data elements in map entries has remained open to this day. While the Final Report (Special Libraries Association, 1956) recommended area-date scale, and the American Geographical Society’s (1969) system favored the area-date-subject approach, arranged chronologically. We might argue that in the online environment with Boolean “and” searching, subject-word searching, and free-text searching, these different display approaches would not really matter that much. This too needs research to examine the extent to which online library catalogs collocate and display map entries on a screen. Carlyle (1996), for example, studied the extent to which online catalogs collocate records representing particular authors (e.g., Homer)
and textual works (e.g., Bible) associated with a large number of relevant records. Results of the study showed that string matching collocated relevant records more successfully than Boolean matching, and that author records were collocated more successfully than work records. Similar studies are needed in the area of cartographic materials.

In the contexts of MLC and less-than-full level cataloging as it pertains to subject added entries for maps, catalogers are advised to practice limited subject analysis only. No subject access is to be provided for insets or ancillary maps. Only one subject heading, reflecting the primary area and subject coverage, is required on less-than-full records. Under no circumstances should any subject heading element be established for use as a secondary tracing in a less-than-full record. To provide additional, but less time-consuming and expensive subject access to less-than-full records, catalogers may add additional tracings, including free-floating headings, format headings, or established headings for additional, co-equal, primary areas. As in full-level cataloging catalogers are to assign only established, fully verified AACR2 subject

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Annotation</th>
<th>AACR2</th>
<th>FGDC (1994)</th>
<th>Emphasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>034 (C)</td>
<td>coded cartographic material</td>
<td>Y O</td>
<td>Spatial reference, chapter 4</td>
<td>F</td>
</tr>
<tr>
<td>043 (C)</td>
<td>geographic area code</td>
<td>Y O</td>
<td>Spatial reference, chapter 4</td>
<td>F</td>
</tr>
<tr>
<td>050 (C)</td>
<td>Library of Congress call #</td>
<td>Y O</td>
<td>NA</td>
<td>A</td>
</tr>
<tr>
<td>052 (C)</td>
<td>geographic classification code</td>
<td>Y O</td>
<td>Identification information, ch. 1</td>
<td>F</td>
</tr>
<tr>
<td>086 (C)</td>
<td>government document class #</td>
<td>Y O</td>
<td>Distribution, chapter 6; citation, 8</td>
<td>F</td>
</tr>
<tr>
<td>1xx (C)</td>
<td>name main entry heading</td>
<td>Y Y</td>
<td>Citation, chapter 8</td>
<td>A</td>
</tr>
<tr>
<td>7xx (C)</td>
<td>added entry heading</td>
<td>Y Y</td>
<td>NA</td>
<td>A</td>
</tr>
<tr>
<td>4xx (C)</td>
<td>series statement</td>
<td>Y Y</td>
<td>Identification, ch 1; citation, ch. 8</td>
<td>A</td>
</tr>
<tr>
<td>8xx (C)</td>
<td>series statement added entry</td>
<td>Y O</td>
<td>NA</td>
<td>A</td>
</tr>
<tr>
<td>245 (T)</td>
<td>title statement</td>
<td>Y Y</td>
<td>Identification, ch. 1; citation, ch. 8</td>
<td>A,F</td>
</tr>
<tr>
<td>255 (C)</td>
<td>mathematical data area</td>
<td>Y O</td>
<td>Spatial reference, chapter 4</td>
<td>F</td>
</tr>
<tr>
<td>260 (C)</td>
<td>publication data</td>
<td>Y Y</td>
<td>Citation, chapter 8</td>
<td>A,F</td>
</tr>
<tr>
<td>265 (C)</td>
<td>the acquisition address</td>
<td>Y O</td>
<td>Distribution, ch. 6; contact, ch 10</td>
<td>F</td>
</tr>
<tr>
<td>300 (T)</td>
<td>physical description</td>
<td>Y O</td>
<td>Entity and attribute, chapter 5</td>
<td>F</td>
</tr>
<tr>
<td>342 (C)</td>
<td>geospatial reference data</td>
<td>Y O</td>
<td>Spatial reference, chapter 4</td>
<td>F</td>
</tr>
<tr>
<td>510 (T)</td>
<td>citation/reference note</td>
<td>Y O</td>
<td>Citation, chapter 8</td>
<td>F</td>
</tr>
<tr>
<td>514 (T)</td>
<td>data quality note</td>
<td>Y O</td>
<td>Data quality, chapter 2</td>
<td>F</td>
</tr>
<tr>
<td>518 (T)</td>
<td>date/time and place of an event</td>
<td>Y O</td>
<td>Time period, chapter 9</td>
<td>F</td>
</tr>
<tr>
<td>520 (T)</td>
<td>summary, etc. note</td>
<td>Y O</td>
<td>Identification, chapter 1</td>
<td>F</td>
</tr>
<tr>
<td>521 (T)</td>
<td>target audience</td>
<td>Y O</td>
<td>NA</td>
<td>A</td>
</tr>
<tr>
<td>522 (T)</td>
<td>geographic coverage note</td>
<td>Y O</td>
<td>Spatial reference, chapter 4</td>
<td>F</td>
</tr>
<tr>
<td>524 (T)</td>
<td>preferred citation note</td>
<td>Y O</td>
<td>Citation, chapter 8</td>
<td>F</td>
</tr>
<tr>
<td>530 (T)</td>
<td>physical form availability note</td>
<td>Y O</td>
<td>Distribution, chapter 6</td>
<td>F</td>
</tr>
<tr>
<td>650 (C)</td>
<td>subject added entry, topical</td>
<td>Y O</td>
<td>Identification, ch. 1</td>
<td>F</td>
</tr>
<tr>
<td>651 (C)</td>
<td>subject added entry, geographic</td>
<td>Y O</td>
<td>Spatial ref., ch. 4; identific, ch. 1</td>
<td>F</td>
</tr>
<tr>
<td>653 (C)</td>
<td>index term, uncontrolled</td>
<td>Y O</td>
<td>Spatial ref., ch. 4; identific, ch. 1</td>
<td>F</td>
</tr>
</tbody>
</table>

C - control; T - free-text; I - full level cat; K - MLC; O - optional; A - emphasized in AACR; F - emphasized in FGDC.

FIG. 3. Map representation in two standards: AACR2 (full level, MLC) and FGDC.
Name authority data elements. The power of the main entry heading is given to the chiefly responsible map maker (110 and 100 MARC fields for corporate and personal main entry headings and related 7xx fields). Precise instructions are given to catalogers with regard to the time spent on consulting reference sources and on the procedures of establishing the name in name authority record.

Series statements (4xx) for less-than-full records are searched in a manner similar to the main entry headings. If the series is identified as a traced series on the series authority record (SAR) and the form of the series statement matches the established form, catalogers are instructed to content designate the series as traced the same (a 440 field). Only established AACR2 series traced differently are transcribed as series added entries (8xx) (Library of Congress, 1991, p. B.3, B.7). Special consideration should be given to series statements for large collections of remote sensing images, aerial photographs, landsats, as well as nondigital cartographic material traditionally collected by map libraries.

Instructions are given to catalogers to optionally add the full address of a publisher, distributor, etc., to the name of the place, and not to add the full address for major trade publishers (per Rule 4C7 in Cartographic Materials, 1982, p. 70). The acquisition address (a 265 MARC field) is omitted.

Uncontrolled (Free-Text) Data Elements

In this group, we discuss the notes area and title added entries.

With regard to the nature and scope of the note fields (5xx MARC fields), catalogers are instructed to be restrictive, and use such notes only to reflect the nature and scope information not apparent from the rest of the description and subject headings (Library of Congress, 1991, p. B.4). Some examples are: Omit all “source of title proper” notes; omit all “parallel titles and other title information” notes; generally omit base map notes. However, as illustrated with two records in the Appendix (Fig. A1), the two MARC fields are identical in all respects except for the note (500) and the government document classification number (086) fields, and could falsely merge into one record. Also, to be omitted or used sparingly, are dissertations and contents notes. Any notes pertaining to variations in title, mathematical and other cartographic data, which used to be the original version note (tag 534), and the notes relating to physical description and dissertations must also be omitted. It would seem that the notes area would be of a particular significance for aerial photographs in view of the fact that the 245 and 255 MARC fields are often not supplied on these maps. In this case, giving the source of title, situation date, and flight lines would seem well justified and worthy of the effort even in the minimal-level description. Other note fields that are treated as optional include data quality, target audience note, geographic coverage note (514, 521, 522 MARC fields), and an additional physical form available note (e.g., 530. Also available on the Internet on abagOnline at http://www.abag.ca.gov/bayarea/eqmaps/. Any future modifications or updates are likely to be available on the Internet prior to “hard copy” publication.) The physical description area (300), in spite of its significance to the identification and description of certain cartographic items, is to be omitted or used rarely.

Title added entries are not normally included in less-than-full records, and secondary title tracings may be made on a restrictive basis only (Library of Congress, 1991, p. B.7).

Discussion

To reiterate:

MLC was designed as a cost effective means of providing access 1) to items worth retaining in the Library’s collections but not worth the expense of full cataloging and 2) to a very large number of items in the arrearages that were unavailable to users and for which the completion of full cataloging was highly unlikely (Library of Congress, 1987).

Now, 10 years later, is the time to reexamine the effect MLC has had in describing and providing access to cartographic items for the purposes of retrieving desired information. The question remains, whether the provision for briefer records (of more material) is better than none? How economical is MLC in the long-run? And which criteria are being used to demote maps to those categories which are not worth the expense of full cataloging? What have we learned about success rate of the “less-than-full” level of map cataloging in OPACs?

The process of full-level cataloging is seen as time consuming and costly mainly due to the authority work involved in consulting authority files for personal, corporate, and series names, as well as for subject headings (e.g., 1xx, 7xx, 4xx, 8xx, 6xx MARC fields). Another reason may be that many catalogers are not trained in map librarianship, non-book cataloging in general, and map cataloging, in particular. Many catalogers who catalog maps on a part-time basis may lack necessary cartographic expertise and experience in understanding certain geographic concepts and the related vocabulary (Ercegovac, 1990a, 1990b, 1991, 1992a, 1992b, 1992c). Some of the conceptual problems are related to the mathematical data such as statements of coordinates, projection, equinox, and scale (e.g., 255, 034 MARC fields). Other labor-intensive activities are measurement of a cartographic item (300 subfield c); articulation of the description that is given in the note area; variations in title; physical description; geographic coverage; accompanying material; contents; and copy-specific notes. Consider the amount of experience that is required to give the following notes:
We Need Geospatial Access to Geospatial Data

Since today’s operational OPACs support bibliographic description of mainly textual data types, items that are of non-textual nature, such as maps, are reduced to text and must be searched accordingly. Consider queries such as, “Show me the National Priority List (NPL) toxic dump sites that are scheduled for remediation in the Los Angeles County within 1 year from now”; “Find all marshes in the state of Louisiana within a specified timeframe”; “Show me a place where Mark Twain was born”; “Any pocket-size map(s) with the location of ski runs in Mono County?” In addition, the user needs maps with certain attributes (e.g., topographic elevation, length and vertical drop of ski slopes, average seasonal temperatures and wind factors). Each of these queries has strong spatial and oftentimes temporal features which must be reduced to textual data elements. The question remains, is it natural for map users to think in terms of authorial data elements? For example, to answer the inquiry on “toxic dump sites,” the user might enter “browse subject” search using the heading “hazardous waste sites” in UCLA’s online catalog ORION and select R12 from the alphabetical list of The Library of Congress Subject Headings (LCSH) (Fig. 1).

As a result, instead of a map, the user sees a bibliographic record containing data elements for title, edition statement, publication data, physical description of a subject heading. If we acknowledge that the matching process in the retrieval system is dependent on the accuracy of the indexing process, on the depth of indexing, and on the specificity (Farradane, 1979), retrieval of records representing the cartographic material is indeed wanting. To paraphrase Handman’s remark, cartographic items, along with other media, are unique resources in libraries and along with their bibliographic complexity and special access requirements, they are primary candidates for mega-cataloging rather than micro-cataloging.

Different Metadata Traditions Converge into a Single Representation Standard

While Anglo-American cataloging tradition has supported the author main entry catalogs whenever possible,
geographers traditionally preferred the area/subject primary access as the organizing principle for their cartographic collections (American Geographic Society, 1969; Special Libraries Association, 1956). It is worth noting that only recently have we seen a constructive cooperation between catalogers representing the library culture and geographers representing the geospatial culture. An attempt to mesh the Federal Geographic Data Committee (FGDC) Content Standard for Spatial Metadata (1995) with the OCLC/MARC Formats represents a new direction that accommodates multiple data types into a single underlying super-structure model. These two standards started off with diametrically opposed philosophical orientations (see Fig. 3). In the summer of 1993, when the preliminary revised draft of the Content Standards for Spatial Metadata of the Federal Geographic Data Committee was prototyped and tested, it consisted of elements that accommodated retrieval of geospatial datasets only; for example, it included spatial references (e.g., coordinates, projection), data structures (e.g., point, vector, raster, string, polygon), and other geographic features (e.g., themes, keywords), as well as elements for quality, accuracy, security, and contact information. However, there was no provision for bibliographic control of these datasets. In contrast, the library community (e.g., AACR2, ISBD(CM)) (1977) had initially reduced the description of maps to less-than-full cataloging. In particular, those data elements which define and distinguish maps from other materials have been treated only as optional. The extent to which the two standards (i.e., FGDC Content Standard vs. ISBD(CM)) (1977) emphasized different data elements, for identification and representation of cartographic items, was striking. While the geographers prefer the area related fields, the library community has paid only minimal attention to the 650 and 651 subject fields, classification codes such as the Library of Congress call number (050), the geographic classification code (052), as well as the mathematical data codes (255 and 034). Catalogers (ISBD(CM)) (1977) continue to be concerned with authorial statements (e.g., corporate body main entry, the 110 MARC field, statement of responsibility, 245 subfield c, name of publisher 260 subfield b, and the associated name added entries 7xx MARC fields). Areas of emphasis between the two national standards are illustrated in Figure 3. Of the 27 most important data elements for representing cartographic materials, only six, mainly name-related data are emphasized in AACR2; two data elements (i.e., title statement and publication data) are equally well represented in both AACR2 and FGDC; all other predominantly spatial and temporal data elements are significantly expanded in FGDC’s entire chapters. However, tighter communication between geographers and librarians has produced a working model which recognizes the importance of both geospatial and bibliographic metadata for representing all types of cartographic items (e.g., paper maps, remote sensing images, aerial photographs). Early synergy between the two standards has added new MARC data elements, including a 342 MARC field for geospatial reference data, a 514 MARC field for data quality note, and a 522 MARC field for geographic coverage note. Similarly, FGDC has extended its own standard both in structure and content. For example, with regard to its structure, the metadata standard, like MARC Formats and ISBD(CM), defines data elements as being mandatory (those that must be provided), mandatory if applicable, and optional (those that are provided at the discretion of the data set producers). In addition, FGDC’s standard is organized in seven main and three supporting sections. The main sections cover information pertaining to identification, data quality, spatial data organization, spatial reference, entity and attributes, distribution, and metadata reference. The supporting sections relate to citation information, time covered, and contact data (Federal Geographic Data Committee, 1995).

With regard to the content, the metadata standard has made a number significant improvements. The identification section, for example, makes a distinction between the “originator,” defined as person(s) and organization(s) that developed the data set, and the “point of contact,” defined as person(s) and organization(s) that can be contacted if questions arise about the data set (FGDC, 1–6). Other contributions pertain to subject access to geospatial data sets. Specifically, the standard provides four types of keywords: Theme (the subject of the data set, such as wetlands, vegetation, etc.), place (the geographic location of the data set, such as Montgomery County, Yellowstone National Park), stratum (the vertical location of the data set, such as seafloor, seabed, troposphere, stratosphere), and temporal (time references for a data set, such as pre-Columbian, World War II) (FGDC, 1–4). According to FGDC’s Metadata Coordinator, the Metadata Standard will continue to evaluate the existing versions, and work on the issues of compatibility with the emerging international metadata standard (B. Tolar, personal communication, March 1997).

Summary and Further Research

With regard to retrieval toolings, we need to examine information seeking patterns and retrieval success rates of the map user who uses OPACs where cataloging records for maps, books, serials, and other formats, are integrated according to the conventional Anglo-American map cataloging practice. What inquiries do users bring to map (imagery) collections most frequently? How do they actually search in OPACs to answer their needs? What are the most common types of errors which lead to aborted searches? How are broad searches modified? Is there an emerging pattern in the error behavior for certain types of searches? In the context of geo-referenced information systems and GISs, we need to study characteristics of GIS users: How do they differ from OPAC users? How
do we measure their success in retrieving relevant information? How do we isolate and analyze different variables which may influence the performance of GISs?

The ultimate goal is to design systems which would be compatible with users’ seeking behavior, level of search sophistication, and retrieval requirements. Whether the user is extracting geospatial meanings from cartographic collections, or wanting to locate a small-scale map of Northridge in order to reconstruct some of the objects after the 1994 earthquake, the system should be designed to accommodate various searching levels and collocate those entries that satisfy a given user’s goal.

Research is needed to determine which data elements are necessary, desirable, and unessential, to represent cartographic materials in OPACs in the contexts of the economic implications and users’ needs. We also need to understand which access points would be useful for different classes of users and how cost-effective different levels of description are in view of a growing number of virtual repositories that are searchable remotely. Is it natural for the user of maps to think in terms of data elements designed to represent non-graphical items? Making the distinction between different types of inquiries for different types of geospatial qualities is important because users at different times will, for varied purposes, desire various access points and retrieval mechanisms to browse, locate, and retrieve cartographic materials.

We envision that multiple user studies would be designed along two dimensions: The extent to which a given (i) technological delivery and (ii) level of detail influence users’ success to retrieve relevant information. Among information technologies, the proposed study might consider the technologies of printed-card catalogs, OPACs, Geographic Information Systems, and the World Wide Web (see Fig. 4). Among levels of detail, we need to study the extent to which minimal-level cataloging, full cataloging, and the extended SGML-based representations might be cost-effective in searching and retrieving spatial data. Such a four-by-three matrix would point to those areas which have been better understood, such as the information seeking of map users in the context of printed-card catalogs, than the other areas (e.g., success of map users in the environments of OPACs, GISs, and the Web). We view this discussion mainly as the framework for further study which would ideally give advantages and disadvantages for each of the data fields of cartographic items, and include the associated long-term cost savings and for each of the identified technologies.

**Appendix**

<table>
<thead>
<tr>
<th>1</th>
<th>010</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>040 GPO c GPO</td>
</tr>
<tr>
<td>3</td>
<td>034 1 a b 18/5000 d E0190000 e E0210000 f N0420000 g N0400000</td>
</tr>
<tr>
<td>4</td>
<td>052 6830</td>
</tr>
<tr>
<td>5</td>
<td>074 856-A-1</td>
</tr>
<tr>
<td>6</td>
<td>086 0 PrEx 3.1074:A1 1</td>
</tr>
<tr>
<td>7</td>
<td>090 b</td>
</tr>
<tr>
<td>8</td>
<td>049 CUYM</td>
</tr>
<tr>
<td>9</td>
<td>110 1 United States. b Central Intelligence Agency.</td>
</tr>
<tr>
<td>10</td>
<td>245 10 Albania.</td>
</tr>
<tr>
<td>11</td>
<td>255 Scale [ca. 1:1,875,000] c (E 19°00'-E 21°00'/N 42°00'-N 40°00').</td>
</tr>
<tr>
<td>12</td>
<td>260 0 (Washington, D.C. : b Central Intelligence Agency, c 1984.)</td>
</tr>
<tr>
<td>13</td>
<td>300 1 map : b col. : c 20 x 17 cm.</td>
</tr>
<tr>
<td>14</td>
<td>500 &quot;Base 800142 (A01/92) 5-84.&quot;</td>
</tr>
<tr>
<td>15</td>
<td>651 0 Albania x Maps.</td>
</tr>
</tbody>
</table>

**FIG. A1.** Two identical records except for the 500 and 086 fields.
Acknowledgments

I express my thanks to Professor Michael Carpenter for allowing me to use some of the material for this article. I also thank Gregory Leazer of UCLA and Pamela Rau of the Library of Congress for the reading of this manuscript, and their criticism.

References


Digital Library Initiative. (1996). The entire May issue of the Institute of electrical and electronics engineering (IEEE) is devoted to the various design and development topics of large-scale digital libraries: Six university sites, sponsored by the National Science Foundation, Advanced Research Projects Agency, and the National Aeronautics and Space Administration.


718 JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE—June 1998


