

Standards for Archival Description

The “Other” USMARC Formats: Authorities and Holdings. Do We Care to be Partners in this Dance, too?

LISA B. WEBER

Abstract: The archival profession’s attention has been focused on the USMARC Format for Archival and Manuscripts Control (AMC), which is part of the USMARC formats for bibliographic data, to the exclusion of the other two USMARC formats: authorities and holdings. The author examines the implications of adopting the USMARC formats for authorities and holdings by looking at the purposes and structures of the two formats, analyzing how they relate to archival description practice, discussing the changes required for the two formats to meet archival descriptive needs, and exploring strategies for the adoption of the USMARC formats to be used in archival information exchange.

About the author: Lisa B. Weber is assistant director for technological evaluation at the Records Program of the National Historical Publications and Records Commission. She prepared this background paper for the June 1989 meeting of the Working Group on Standards for Archival Description in response to issues raised at the group’s December 1988 meeting.

FOR NEARLY A DECADE, the archival profession's strategy for attaining the goal of archival information exchange has been embodied in the *USMARC Format for Bibliographic Data*.¹ Attention has focused almost exclusively on the bibliographic format, to the exclusion of the two other USMARC formats—authorities and holdings. The purpose of this article is to explore the implications of adopting these two USMARC formats for archival information systems.

What are the USMARC Authorities and Holdings Formats?

Any discussion of whether or not the archival profession should use the USMARC formats for authorities and holdings must be based on an understanding of what they are, why they were developed, and how they are used.

In the library community, the USMARC authority and holdings formats are adjuncts to the more important USMARC bibliographic format that contains generic information about mass-produced, commercially distributed library material. The authority format supports headings management for better retrieval of the mass-produced materials, whereas the holdings format enables libraries to record copy-specific or local information about these items.

Authorities format. The *USMARC Format for Authority Data* (USMARC authority format) is designed to carry authoritative information concerning the standard form of names and subjects to be used as access points, references to the standard forms, and the interrelationships among these forms.² Librarians have maintained manual

authority files for decades and could see many advantages to exchanging authority data, so the development of the USMARC authority format in 1976 was a logical step. In 1981, the Library of Congress revised the format due to the Linked Authority Systems Project, a part of the larger Linked Systems Project that aims to establish computer-to-computer linkages among bibliographic systems.

In libraries, the purpose of authority control is "headings management," i.e., to provide standard terminology for improved retrieval of the information in the bibliographic record. Jackie Dooley has explained that the process of authority control is to (1) establish the form of heading, (2) determine cross-references, (3) determine the relationships of the heading to other headings in the file, and (4) document the decisions.³ The USMARC authority format accommodates this process, permitting seven specific kinds of authority records including established headings, references traced and untraced, subdivisions, and node records.⁴ Though not always grouped consistently, the fields within the USMARC authority format generally fall into the following functional blocks: Headings (1XX); Cross-references (*see from* tracings, 4XX); Linking references (*see also* tracings, 5XX); and Notes (6XX).

The USMARC authority format functions most visibly in the control it exerts over the form of name and subject headings. Currently, the Library of Congress uses the USMARC authority format to ex-

¹*USMARC Format for Bibliographic Data* (Washington, DC: Library of Congress, 1988). The bibliographic format carries information about a variety of library materials: books, serials, maps, music (scores and sound recordings), visual materials, computer files, and archives and manuscripts.

²*Format for Authority Data* (Washington, DC: Library of Congress, 1987).

³Jackie M. Dooley, "An Introduction to Authority Control for Archivists," in *Archives and Authority Control*, ed. Avra Michelson, published as part 2 of *Archival Informatics Newsletter and Technical Report* 2:2 (Summer 1988): 5-18. The term *heading* refers to names (personal, corporate, meeting, jurisdiction, name/title combination, and uniform title) and subjects (topical and geographic).

⁴Node records enable the USMARC authority format to accommodate thesauri and other kinds of hierarchical and syndetic structures.

change name authority data in the Library of Congress Name Authority File and subject authority data in the Library of Congress Subject Headings file. The Art and Architecture Thesaurus Project is using the USMARC authority format for its implementation in the Research Libraries Information Network (RLIN). In conjunction with the Library of Congress, Name Authority Cooperative Organization libraries contribute name authority data directly to the LC Name Authority File through the Linked Systems Project. The Library of Congress is the only institution adding terms to the LC Subject Headings file, although other institutions can suggest terms to the Library of Congress. Both the LC Name Authority file and LC Subject Headings file are available through RLIN and OCLC, and are published on microfiche and CD-ROM. The LC Subject Headings file is also available in a paper format.

Holdings. The *USMARC Format for Holdings Data* (USMARC holdings format) is designed to be a carrier for holdings and locations data for all forms of material.⁵ It permits librarians to attach information about a specific, locally-held copy to the generic information about that work contained in the bibliographic record.

Most networks and system vendors defined their own local holdings fields for their specific implementations prior to the development of the USMARC holdings format. Serials librarians provided a strong catalyst for the development of a standardized holdings format. In a scenario that should sound familiar to archivists, they wanted to use automated techniques to track the multiple processes—such as check-in, binding, microfilming, and circulation—that occur to the individual items that comprise serial publications. Eight southeastern research libraries (the Southeastern Associa-

tion of Research Libraries Cooperative Serials Project) began to develop a format in 1982. The format remained in “final draft,” as published by the Library of Congress until 1990, when it was officially adopted as a USMARC format.

The complexity of the holdings format has limited its widespread use. The number of defined fields is relatively small, but there is a dauntingly complicated pattern of interdependent fields that define the enumeration and chronology for the bibliographic unit, supplement, and indexes, and the values contained in those categories. The format was defined to be “attached” to the bibliographic record through an explicit (machine) link, implicitly through common identifiers (i.e., ISBN, CODEN), or through being embedded within the bibliographic record itself. Vendors and users have not embraced the holdings format enthusiastically, although it has been implemented at Harvard University; the universities of Georgia, Kansas, and Florida; and by VTLS, Inc.

What do the USMARC Formats Have to Offer Archival Information Systems?

One of the difficulties in determining the value of the formats for archivists in creating authority and holdings data is the underlying lack of certainty about the role and purpose of archival information systems. Experience with the USMARC Format for Archival and Manuscripts Control (AMC) and efforts to define archival information systems as parts of, or separate from, the library bibliographic community provide some basis for assessing the potential value of the authority and holdings formats. By comparing our needs with those of the library community, we can see what the formats have to offer. The first point of comparison should be the bibliographic format, since it is the center of library automated information systems.

Bibliographic format. USMARC AMC,

⁵*USMARC Format for Holdings Data* (Washington, DC: Library of Congress, 1989).

although part of the bibliographic family of formats, is unlike any other kind of bibliographic record.⁶ The MARC Format for Manuscripts failed abysmally because it treated archival material like weird kinds of books. But the National Information Systems Task Force created USMARC AMC to accommodate the archival requirements of collective description, life-cycle management, access by provenance or context, intra- and inter-record links, and unique or local data. Although the library community shares several of these requirements, its vision has been restricted by the economic incentive of copy-cataloging that underlay the development of USMARC. For the most part, librarians see these requirements as solely those of the archival profession.

The library bibliographic record, the focus from which all else emanates, contains generic information about commercially distributed, mass-produced library material. If one expands the definition of authority control to include the concept of repeated use, then in a library setting bibliographic records function as authority records because the information contained in these generic records is used repeatedly for local cataloging.⁷ USMARC AMC records are not authority records in this sense.⁸

⁶This is true regardless of format integration, although the problems with leader 06 and leader 07 emphasize the point.

⁷For a detailed discussion of the expanded definition of authority control, see David Bearman and Richard Szary, "Beyond Authorized Headings: Authorities as Reference Files in a Multi-disciplinary Setting," in *Authority Control Symposium, ARLIS Occasional Papers #6*, ed. Karen Muller (Tucson, AZ: Art Libraries of North America, 1987). For evidence of the way in which the library community is coming to realize that bibliographic records have an "authority control" aspect, see *Multiple Versions Forum Report: Report from a Meeting Held December 6-8, 1989, Airlie, Virginia* (Washington: Library of Congress, Network Development and MARC Standards Office, 1990).

⁸I disagree with Szary's definition of authority records in his paper, "Design Requirements for Archival

Archival materials are unique—sometimes as individual items, nearly always as aggregations. The USMARC AMC records that describe them contain collection-specific information about intellectual content, physical aspects, provenance, and actions or processes that cannot be used for derivative cataloging by other archival repositories, except in the case of microform reproduction. So, this comparison exposes an important and fundamental divergence between library bibliographic records and USMARC AMC records.

Authority format. Authority data is profoundly important to archival information systems, but our needs both embrace and go beyond the "headings management" needs of librarians. Both Richard Szary and David Bearman have written extensively on the subject.⁹ Bearman has shown that the different kinds and levels of retrieval needed in archival information systems require different types of authority control.¹⁰ Archivists need relatively simple value tables for data elements such as action, language, and status, and more complex thesauri for subject, genre, and occupation. These needs parallel the headings management needs of the library community. The divergence comes when archivists think about provenance data. The characteristics and activities of the corporate and personal records creators are also authority data for archivists, but in a much more expanded concept of authority control

Authority Systems," presented at the 1988 annual meeting of the Society of American Archivists, Atlanta, Georgia, in which he states that archival bibliographic records are authority records.

⁹See the bibliography in Lisa B. Weber's "Development of Authority Control Systems Within the Archival Profession," in *Archives and Authority Control*, ed. Avra Michelson, published as Part 2 of *Archival Informatics Newsletter and Technical Report 2:2* (Summer 1988): 35-38.

¹⁰David Bearman, "Authority Control: Issues and Prospects," *American Archivist* 52 (Summer 1989): 286-99.

than that held by the library community. For archivists,

the expanded authority record would provide a structure and repository for information about the history and characteristics of cultural entities that could be used to determine which persons, geographic locations, concepts, or other entities (headings for which have been included in bibliographic records) would be useful concepts to incorporate as access points in a search strategy.¹¹

The expanded authority files (what Bearman calls *reference files*) would incorporate aspects of headings management, but would be reference sources in their own right. Provenance data, form-of-material data, general records schedules, and geographic data are examples of potential reference files.

One proof of how important expanded authority or reference files are to us is the fact that we are busy creating them. Agency history records, which contain provenance authority information, quickly became the backbone of the Research Libraries Group (RLG) Seven States Project and its sequel, the RLG Government Records Project. Numerous archival repositories are following suit in local implementations. The RLG Government Records Project is creating two more kinds of reference files: records schedules and form-of-material records.

The authority data that we are fashioning currently resides in the USMARC AMC format and not in the authority format, where it logically should be placed. At first glance, we should be trying to change the USMARC authority format to serve all of our different authority control needs. There is no technical obstruction in the USMARC authority format structure to prevent such expansion. It is important to note, however, that not all potentially valuable ref-

erence files would fit into the USMARC authority format with equal facility. For example, corporate and personal provenance data is an expansion of the concept of personal and corporate name headings. We could use the existing historical note field (655) and/or define new fields to carry this richer and more varied information. On the other hand, form-of-material and records-schedule data are unlike anything librarians now capture, and would require the definition of very different categories within the USMARC authority format.

According to Bearman, archival reference files should also be able to import authority information that is created by outside sources, such as biographical data from Marquis *Who's Who* publications.¹² Such importation presents a problem, not specifically with the USMARC format, but with the data from outside sources, which does not have the content designation needed to map specific elements into the USMARC authority format.

Archival implementation of the USMARC authority format would require a simultaneous revision of the USMARC AMC format, because it currently contains fields that more appropriately reside in an authority record. For example, the biographical or history note field (545), and the fields for the occupation (656) and function (657) access points are all authority-related data. The most appropriate kind of information to remain in a MARC AMC record is descriptive, both intellectual and physical.

Holdings format. The relationship of the USMARC holdings format to MARC AMC is distinctly different from that of the USMARC authority format. As already illustrated, USMARC AMC records are not authority records, but they are, concep-

¹¹Bearman and Szary, "Beyond Authorized Headings," 72.

¹²David Bearman, "Archives and Manuscript Control with Bibliographic Utilities: Challenges and Opportunities," *American Archivist* 52 (Winter 1989): 37.

tually, holdings records because they contain information about local or institution-specific data. Clearly, the thought of putting all USMARC AMC data in the USMARC holdings format is absurd because the holdings format was not designed to contain this kind of information.¹³ There is a temptation, however, to use the USMARC holdings format for box or container data because the need to maintain hierarchical links is similar for serials control in libraries and for container-level control in archives. Systems that have implemented the holdings format provide those links.

The lack of a place in USMARC for both the intellectual and physical data included in contents lists has been less obvious in the national arena because this level of detail is rarely included in RLIN or OCLC. A repository can choose to emphasize a specific part or item in a collection by describing it in a separate USMARC AMC record, linked to the collection-level description. The lack of a listing capability is far more apparent in a local system. Bearman maintains that NISTF designed the “contents note-formatted” field (505) together with the “location” field (851) to be used for this purpose. Unfortunately, no existing software implementations support the functions that archivists need for the data that should be carried in these fields. Currently, most repositories are keeping their container lists on separate systems. The Virginia State Archives, which uses VTLS, may be the only archives to apply the USMARC holdings format.

The simple fact that archivists do not have the kind of data for which the USMARC

holdings format was designed makes its use illogical. Nonetheless, it is a tempting option for archivists using software systems that provide some of the necessary capabilities. In fact, this resembles how the RLG repositories are using the bibliographic format for authority data in the guise of agency history records. Even though agency history data should logically appear in the authority format, the RLG implementation places it in the bibliographic format because of software implications. Making decisions about placement of data based on software implementation can cause more problems that it solves. This is likely to be the case in the use of the holdings format for management of information below the collection level. It is clear that the profession must pursue the development of system implementations that provide the functions we need to manage information below the collection level. Using the USMARC holdings format is not the answer.

How do We Get What We Want?

The first step in attaining what we want is to define precisely what it is that we need. This, of course, leads back to the original question, which the profession can no longer evade: “What is the purpose of an archival information system?” For authority records, we must define the data elements we want to include in our expanded concept, such as provenance, form-of-material records, and records schedules. In fact, the RLG Seven States and Government Records projects have done just that for their own use. The profession must develop a mechanism to adopt, enhance, or reject the RLIN-specific standards before there can be any thought about strategies for change. For data below the collection level (i.e., container lists), if we can be satisfied with the data elements that already exist in the USMARC AMC format, then we need to work toward appropriate implementations rather than change the format.

¹³This is the philosophical argument we employed to try to convince the Library of Congress and MARBI to keep the locations field 851 in the bibliographic format. MARBI is the American Library Association’s Machine-Readable Bibliographic Information Committee, which advises the Library of Congress about changes to the USMARC format.

With regard to authority data, there are four ways to achieve standardization. The first is to lobby the Library of Congress and MARBI (see footnote 13) to change the USMARC authority format. That would be the most logical, the most direct, and the most difficult. The library community is unlikely to welcome our changes, and our position on MARBI doesn't give us much clout—our MARBI representative isn't even a voting member. Unless we can convince the library community that they will benefit from the use of an expanded authority format, it will be an uphill battle. Our chances of convincing them are poor.

The second option is to continue what we are doing, but more openly. MARBI has approved a proposal to define leader 08 for "type of control" and code "a" in that leader as "archival control." In essence, this means that we could control our own records and proceed to define types of MARC AMC records such as provenance, records schedules, and form of material. Instead of trying to move MARBI to change the USMARC authority format, we would define what we need within AMC. This wouldn't be easy because we would still have to make changes to USMARC AMC via MARBI, and the committee is more sensitive to changing any bibliographic fields because of format integration. Additionally, it will be clear to MARBI that the kind of data we are defining is authority information.

The third option is to convince MARBI to define authority control fields within the USMARC authority format but label them as archival. Leader 06 in the format is coded "z" in authority records. We could persuade them to define leader 08 in the USMARC authority format as "type of control" and assign a code for archival. We could then define a set of fields that would accommodate our needs.

The fourth option is better labelled a possibility. In the paper Bearman prepared for MARBI to explain the distinctions between

bibliographic, serial, archival, and museum control, he concluded with a caution that if the Library of Congress and MARBI are not flexible enough to expand USMARC to accommodate the museum community's needs, that community will define an ISO 2709/ANSI Z39.2-compatible format, separate from USMARC.¹⁴ If that occurs, the archival profession may need to choose between the two communities. Because our records currently reside in bibliographic networks, we would be unlikely to throw our lot in with the museum community which, as yet, has no national networking structure. But the possibility may present itself and force a decision. The critical issue would be our ability to sustain a national network outside the bibliographic structure. Nonetheless, our best strategy may be to join forces with the museum community, which has similar authority control needs, convince MARBI of our requirements, and motivate vendors to develop viable implementations.

Implementation

Although only alluded to indirectly thus far, implementation—how the software handles the formats—is a key issue. The primary reason that we have used the bibliographic format for authority records is because RLIN allows archivists to make explicit links between the authority records and the descriptions of the archival materials. The structure is of no use if we don't have ways to implement what we require for our concept of authority files. It doesn't matter where the data elements reside.

Szary lists a number of design requirements for an archival authority control system.¹⁵ Three of those requirements are: (1) provision for multiple record types, each

¹⁴David Bearman, "Can MARC Accommodate Archives and Museums? Technical and Political Challenges," (unpublished paper, 1989).

¹⁵Szary, "Design Requirements," fn. 10.

structured to record the information characteristics of a particular type of entity, and restricted to that information; (2) no type of record must have a privileged status in the system; and (3) provision for links between any records regardless of type. These design requirements differ substantially from those perceived by the library community. There the bibliographic record has a privileged status and the linkages are not as important. The challenge is to convince both MARBI and vendors of our need.

Conclusion

Ultimately, the question of adopting the "other" USMARC formats comes down to the value to be derived from their implementation. How important to the profession are authority data and the ability to manage information below the collection level? Authority data is very important; the RLG projects are proving its significance in a shared environment. Developing standardized structures for archival reference files within the context of USMARC should take a high priority on our list of recommendations. The strategies and tactics we use to achieve this goal are unclear, but may become more apparent in the near future. At the very least, the museum community's interest in using USMARC for an exchange structure will make MARBI more aware of the needs and requirements of other communities.

Authority data is such an important part of archival control because some feel it is what we can afford to exchange. Thus far, the archival profession has been able to ride the economically profitable coattails of copy-cataloging, which has sustained the library networks for over two decades (and archivists for the last five years). But the moment of truth is near if we continue to lobby the bibliographic community to provide us with both the format and the software func-

tionality we need to achieve successful archival information systems. We will need to demonstrate economic reasons for exchanging information beyond the altruistic one of sharing information for broader access. According to some, sharing authority data or reference files is a major economic reason for the exchange of data.¹⁶

The fact that the profession has not been more vocal about the lack of management control below the collection level within implementations of USMARC AMC raises questions about the importance of this capability in a shared environment. From a management perspective, it is critically important to control information below the collection level. Many archivists are using commercial software packages (dBase, PC-File, Q&A, etc.) to manage their data. Is this capability important only in a local environment, or are archivists forced to use off-the-shelf software because the current USMARC AMC implementations do not meet our needs? Do archivists want to exchange container lists? The USMARC holdings format clearly is not the vehicle with which to accomplish this. What is confused is how we should proceed with issues of implementation and exchange for control below the collection level.

In the context of this Working Group, I believe authority control is a more important issue than the ability to manage information below the collection level, and thus should take precedence on our list of the Working Group's recommendations. The second issue also deserves serious attention. Just how high on our list of recommendations both issues are placed is a decision with which the entire group must grapple.

¹⁶David Bearman, *Towards National Information Systems for Archives and Manuscript Repositories: The National Information Systems Task Force (NSTIF) Papers, 1981-84* (Chicago: Society of American Archivists, 1987), 65-85.