

## Perspective

# Subject Indexing in Context

JACKIE M. DOOLEY

**Abstract:** Integration of archival materials cataloged in the USMARC AMC format into online catalogs has given a new urgency to the need for direct subject access. The author offers a broad definition of the concepts to be considered under the subject access umbrella, including not only topical subjects but also proper names, forms of material, time periods, geographic places, occupations, and functions. It is argued that it is both necessary and possible to provide more consistent subject access to archives and manuscripts than currently is being achieved. The author describes current efforts that are under way in the profession to address this need.

*About the author:* Jackie M. Dooley is head of collections cataloging at the Getty Center for the History of Art and the Humanities in Santa Monica, California. She is a member of the Committee on Archival Information Exchange, serves as the Society of American Archivists' liaison to the American Library Association's Committee on Cataloging: Description and Access, and is vice-chair/chair-elect of ALA's Rare Books and Manuscripts Section. Her paper is based on a presentation to the Society of American Archivists in St. Louis in October 1989.

HISTORICALLY, ARCHIVISTS HAVE TENDED toward the view that access to archival materials by provenance renders access by subject unnecessary. This attitude may stem at least in part from the impossibility of analyzing the detailed subject content of millions of items in any single repository, and it has been rationalized by the claim that access by provenance leads, albeit indirectly, to the subject content of collections. Perhaps as a consequence, much less attention has been given to careful analysis of the merits of other attributes of archival materials for retrieval purposes.

Now that it has become routine to include descriptions of archival materials in integrated online catalogs, however, there is a growing sense among archivists that direct subject access to archives and manuscripts is not only desirable but necessary. Ample evidence of this need is found in the extent to which subject terms are found appended to USMARC Archives and Manuscript Control (AMC) descriptions in shared databases such as the Research Libraries Information Network (RLIN), as well as in hundreds of local online catalogs. Widespread acceptance of MARC AMC has provided both the opportunity and the urgency behind this change of attitude.

In recent years, subject access has been recognized as a significant concern by both the American and Canadian working groups on descriptive standards. In its final report, the Working Group on Standards for Archival Description, funded by the National Historical Publications and Records Commission (NHPRC), recommended that the profession address the need for guidelines for archival subject access.<sup>1</sup> Several years earlier, the Canadian Working Group on Archival Description Standards declared that

"access to the content of archival holdings is dependent upon detailed subject indexing."<sup>2</sup> However, both working groups have cautioned against blind adherence to the subject indexing practices followed by libraries, thereby establishing the need to examine subject indexing within an archival context.

This paper assumes that it is necessary to provide direct subject access to archival and manuscript materials (hereafter referred to jointly as "archival materials") and that it is possible to do so in a much more consistent fashion than currently is being achieved. The need for subject access is reviewed and the nature of archival "subjects" is briefly described, followed by an inventory of the ways in which the MARC AMC format accommodates subject data and some comments on the need to understand users of archival materials. A description of current efforts that are under way to address these needs then follows, along with some recommendations for action.

### The Need for Consistent Subject Access

In her 1986 study of archival indexing practice, Avra Michelson concluded that archivists are inconsistent in both what they describe and how they describe it.<sup>3</sup> She asked the forty major repositories then contributing to RLIN's AMC file to assign index terms to three collections, based on brief descriptions of their contents. The results demonstrated virtually no consensus on the concepts that should be indexed and the indexing terms that should be assigned. For example, no repository assigned personal

<sup>1</sup>Working Group on Standards for Archival Description, "Recommendations of the Working Group on Standards for Archival Description," *American Archivist* 52 (Fall 1989): 462-77.

<sup>2</sup>Bureau of Canadian Archivists, Working Group on Archival Descriptive Standards, *Toward Descriptive Standards* (Ottawa: Bureau of Canadian Archivists, 1985), 14.

<sup>3</sup>Avra Michelson, "Description and Reference in the Age of Automation," *American Archivist* 50 (Spring 1987): 192-208.

or family name access points (even for a collection of family papers), very few assigned form of material headings (such as *diaries* or *family papers*), and only about half provided direct access by geographical place. Topical subject headings ranged from the very broad to the very specific; this is hardly surprising, since no guidelines exist to advise archivists on indexing specificity.

David Bearman has concluded from Michelson's results that consistent subject indexing of archival materials is unattainable.<sup>4</sup> This paper proposes instead that there is an urgent need for archivists to learn more about subject indexing and to develop thesaurus application guidelines tailored to the archival context. Bearman is correct when he observes that library and information sciences have shown that perfect consistency among indexers is nearly impossible. Within the archival context, however, archivists certainly can devise guidelines to improve on the current situation.

It should be noted that the approach taken in this paper focuses on broad subject indexing at the collection level, not detailed content analysis at the folder or item level as has been described by Richard Lytle<sup>5</sup> and others. This focus is in recognition of the fact that most repositories create MARC AMC descriptions at the collection or series level and that they cannot afford to undertake more detailed description or subject analysis for formal data exchange. In this collection-level descriptive context, standard catalog descriptions are created, based on the rules in *Archives, Personal Papers, and Manuscripts* (APPM),<sup>6</sup> using

the scope and content descriptions from detailed finding aids as the basis for summary notes in MARC AMC records. Within the context of a MARC AMC database, it is principally from such summary descriptions that subject descriptors can most usefully be derived, as Richard Smiraglia has so clearly described.<sup>7</sup>

Widespread application of this MARC AMC-based collection-level approach gradually has served to enhance, and at times replace, several traditional means of subject access to archival materials. Mary Jo Pugh, one of many authors who have noted this trend states, "The researcher . . . usually presents a subject request."<sup>8</sup> Archivists and researchers already depend on an array of published subject guides, collection surveys, repository guides, and local card catalogs for subject access to broad topics such as the Civil War, black women, railroads, or education in Pennsylvania. Although such guides can be effective, they exist for only a limited number of subjects, and they become dated the moment they are published.

Shared databases provide archivists the opportunity to build a universal "subject guide" covering all repositories and all subject areas—a guide that would be increasingly up to date, not immediately obsolete. If such a guide is to be effective, however, the data must be consistent, or retrieval will be haphazard. Many archivists have experienced the frustration of searching a large database containing records that were created without a standardized approach to either authority control or subject indexing, or of finding that entirely different terminologies have been used in

<sup>4</sup>David Bearman, "Authority Control Issues and Prospects," *American Archivist* 52 (Summer 1989): 288.

<sup>5</sup>Richard Lytle, "Intellectual Access to Archives. I. Provenance and Content Indexing Methods of Subject Retrieval," *American Archivist* 43 (Winter 1980): 73.

<sup>6</sup>Steven L. Hensen, *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Li-*

*braries* (Chicago: Society of American Archivists, 1989).

<sup>7</sup>Richard Smiraglia, "Subject Access to Archival Materials Using LCSH," *Cataloging and Classification Quarterly* 11, nos. 3,4 (1990): 63-90.

<sup>8</sup>Mary Jo Pugh, "The Illusion of Omniscience: Subject Access and the Reference Archivist," *American Archivist* 45 (Winter 1982): 35-36.

the separate indexes to specific collections, even within a single repository. In response to such problems, the Canadian working group recommended that "institutions which index their holdings [should] prepare cumulative indexes based on consistent procedures and existing subject headings rather than separate indexes for each fonds."<sup>9</sup>

Increasingly, archival descriptions are found in the same databases as books, periodicals, visual materials, museum objects, and other media. Lytle has pointed out that "many users . . . ask questions which cross traditionally defined boundaries."<sup>10</sup> He refers to the traditional boundaries between archives and manuscripts, but his argument can easily be extended to other media as well. In a similar vein, Ann Gordon's report on the NHPRC-sponsored Historical Documents Study states that "researchers must cross the boundaries between types of sources when they select their evidence. . . . Their work is made more efficient and comprehensive by greater integration of the systems."<sup>11</sup> And Lawrence Dowler says, "Just as the boundaries between academic disciplines are weakening, so too may the boundaries between libraries, archives, museums, etc., be diminishing."<sup>12</sup>

As scholars begin to realize the potential power of databases that integrate access to published and unpublished texts, images, and objects, they quickly will come to expect integrated access, regardless of the types of repositories in which particular materials are housed. The door is suddenly open to

much greater use of archival materials than ever before—if users can locate them in databases.

Studies of searches in library online catalogs show that a preponderance of queries for published materials are requests for information by subject; the critical place of subject access also has been reported in the context of visual archival materials, such as documentary photographs<sup>13</sup> and newsfilm.<sup>14</sup> Researchers who pose subject questions in those contexts may also wish to retrieve textual archival records and manuscripts, and emerging realities such as the widespread availability of research library catalogs via Internet mean that searching will be increasingly unmediated by reference archivists.

Many authors, subjects, and forms of material are relevant across media, but users will not discover this if archivists choose indexing practices and vocabulary that differ radically from those used by librarians and other database builders. This discovery will also be impeded if the archival approach continues to require that questions be mediated by reference staff. We must find ways to preserve access to the special characteristics of archival materials, while carefully considering the ramifications of being different.

### The Nature of Archival "Subjects"

Do archives and manuscripts *have* subjects? Bearman has stated that subject content analysis of archival materials is impossible because, unlike the way in which books or journal articles are written, most original source materials are written with

<sup>9</sup>Bureau of Canadian Archivists, Working Group on Archival Descriptive Standards, *Toward Descriptive Standards*, 71.

<sup>10</sup>Lytle, "Intellectual Access," 66.

<sup>11</sup>Ann D Gordon, *Using the Nation's Documentary Heritage: The Report of the Historical Documents Study* (Washington, D.C.: National Historical Publications and Records Commission in Cooperation with the American Council of Learned Societies, 1992), 57.

<sup>12</sup>Lawrence Dowler, "Conference on Research Trends and Library Resources," *Harvard Library Bulletin*, New Series 1, no. 2 (Summer 1990): 7.

<sup>13</sup>Barbara Orbach, "So That Others May See: Tools for Cataloging Still Images," *Cataloging and Classification Quarterly* 11, nos. 3,4 (1990): 163–91.

<sup>14</sup>Martha Yee, "Subject Access to Moving Image Materials in a MARC-based Online Environment," in *Beyond the Book: Extending MARC for Subject Access*, edited by Toni Petersen and Pat Molholt (Boston: G.K. Hall, 1990), 97–115.

no conscious subject or thesis in mind.<sup>15</sup> Even though this essential difference between archival and published materials is valid, it hardly implies that subject terminology cannot usefully be applied to archival descriptions.

Consider, for example, a collection of Civil War diaries. Despite the fact that the soldiers who penned these journals did not set forth formal theses, draw particular reasoned conclusions, or neatly package their work with a table of contents and index for ease of consumption, their writings are *about* certain things: the soldiers themselves, life in their regiments, specific places and times, particular Civil War battles, their hometowns, and their thoughts on life and death, to mention only a few obvious possibilities.

Similarly, the records of the chancellor's office at the University of California at Berkeley during the late 1960s contain materials about the particular person, institution, place, and time, but they might also document less obvious subjects that have no explicit link to the provenance of the records or the functions of the office. Some possible subjects include persons with whom the chancellor corresponded extensively; significant events, such as the People's Park crisis and the free speech movement; and other aspects of sociological change, such as debates on changes to the undergraduate curriculum. The difficulty lies not in the fact that such materials do not *have* subjects but in the process of determining which are the most important subjects to index.

What is the scope of "subject access" in the archival context? One author has stated that "requests for records by proper name, geographical area, date, or form may conceal a subject request."<sup>16</sup> Such requests do not conceal subject requests, they *are* sub-

ject requests. The archival literature often gives the impression that subjects are strictly generic topics—rain forests, football games, railroads, generals, or skyscrapers. In reality, specific named entities, including particular people, organizations, government agencies, geographic places, and events are no less subjects than are generic topics.

Other aspects of archival materials that traditionally have not been considered subjects also provide important clues to subject content and form strong links between provenance and subject content without being purely one or the other. For example, given the evidentiary nature of archives and manuscripts, the time and place in which they were created are often, in a very real way, their subjects. Time and place are particularly potent access points for the papers of little-known or anonymous persons. Form of material also can be a powerful indicator of subject content, particularly when combined with time or place. Consider, for example, account books, overland journals, genealogies, field notes, death certificates, business correspondence, ships' logs, cartoons, and voting registers: All these forms suggest a great deal about the content of records.

### Subjects in MARC AMC Records

The MARC AMC format provides twelve fields for describing and indexing the subject content of archival and manuscript materials.<sup>17</sup>

Generic topical subjects are entered in field 650 (for subject heading lists such as

<sup>15</sup>David Bearman, "Authority Control: Issues and Prospects," unpublished paper presented to the Society of American Archivists, Atlanta, Georgia, October 1988, 12.

<sup>16</sup>Lytle, "Intellectual Access," 68.

<sup>17</sup>These fields and their archival implementations are described in more detail in Harriet Ostroff, "Subject Access to Archival and Manuscript Material," *American Archivist* 53 (Winter 1990): 100–05. It should be noted that all of the fields mentioned, except fields 656 and 657, are available in all MARC formats, not just AMC. Following integration of the MARC formats, all fields will be available in all formats.

*Library of Congress Subject Headings*,<sup>18</sup> or LCSH) or 654 (for faceted thesauri such as the *Art and Architecture Thesaurus*,<sup>19</sup> or AAT). Fields 600, 610, 611, and 651 are for proper name headings for persons, corporate bodies, conferences, and geographic places, respectively. All these fields are meant to be used with controlled vocabularies or authority files, and they are in widespread use in library and archival cataloging.

Field 655 is used for many special-format materials, including rare books, two-dimensional graphics, three-dimensional artifacts, and moving-image films, in addition to archival materials. Archivists have yet to identify a vocabulary that fully meets their needs, but *Form Terms for Archival and Manuscript Collections* (FTAMC)<sup>20</sup> has been used extensively. The AAT, which has added numerous archival terms to its document types hierarchy, is also increasingly gaining an archival audience.

Field 656 is for occupations reflected in the contents of the materials, and field 657 contains terms for the activity or function that generated archival records. Consensus does not yet exist on which controlled vocabularies should be used for occupations, but some archivists use LCSH and others are beginning to use AAT. As for functions, the RLIN Governmental Records Project produced an extensive vocabulary that was incorporated into the AAT functions hierarchy. Occupations and functions

are inextricably linked with provenance, and it is conceivable that these data could better be stored in authority records than in MARC AMC descriptive records.<sup>21</sup>

Fields 043, 045, and 072 contain coded data for access by geographic place, time period, and broad subject area. Standard MARC code lists exist for place and time but not for broad subjects; archivists should feel some urgency about publishing an appropriate list of codes for field 072 if exhaustive retrieval is considered important.<sup>22</sup> These three fields could be of great benefit for archival retrieval if they were widely used and were searchable online in a majority of retrieval systems, conditions that unfortunately do not yet exist. Nevertheless, if archivists worked together with other interested constituencies to make a strong case to vendors, this could change.

Clearly, the MARC format presents many choices, perhaps dauntingly so, for subject indexing. In the context of an integrated online catalog containing a variety of archival and nonarchival primary and secondary sources, it is critical that archival catalogers use the various fields in generally the same way that catalogers of other media use them. Once subject terminology or other access points have led users to archival descriptions of possible relevance, the full descriptive record that elaborates on provenance will provide the necessary context.

Subject analysis is an extremely complex business, and as Michelson's results re-

<sup>18</sup>Library of Congress, Office of Subject Cataloging Policy, *Library of Congress Subject Headings*, 14th ed. (Washington, D.C.: Cataloging Distribution Service, 1991).

<sup>19</sup>*Art and Architecture Thesaurus*, Toni Petersen, Director (New York: Oxford University Press, "Published on Behalf of The Getty Art History Information Program," 1990).

<sup>20</sup>Thomas Hickerson and Elaine Engst, comps., *Form Terms for Archival and Manuscripts Control* (Stanford, Calif.: Research Libraries Group, 1985). This list of terms is not actively maintained and is problematic due to lack of cross-reference structure, scope notes, or other elements of a formal thesaurus.

<sup>21</sup>SAA's Committee on Archival Information Exchange plans to work with the Philadelphia Area Consortium of Special Collections Libraries (PACSCL) on a proposal to enhance the USMARC authorities format to incorporate data such as agency history and personal biographical information, which are presently embedded in AMC descriptive records. Written communication from Kathleen Roc, CAIE chair, to committee members, 10 March 1992.

<sup>22</sup>Ostroff notes that users of the *National Union Catalog of Manuscript Collections* have frequently requested access to broad subjects; Ostroff, "Subject Access to Archival and Manuscript Material," 101.



veal, current archival practice is virtually random. In the same way that archivists have come to realize the importance of APPM for establishing consistent data-content guidelines for MARC AMC descriptive records, they must now recognize that the mere existence of appropriate MARC fields does not guarantee consistent data content. If databases such as RLIN AMC are to serve as effective reference tools, archivists must agree on what to index and how to index it.

The issues to be resolved are numerous, and they include the following:

- What level of indexing depth is appropriate for a large collection?<sup>23</sup>
- Is access by broad subject area desirable? If so, can a list of broad subject codes be agreed upon for use in field 072?
- Should forms of material be indexed exhaustively or selectively?
- Under what circumstances should names of people be indexed as subjects in field 600 and as authors in field 700?<sup>24</sup>
- Is direct access by time period important?
- Do existing LCSH application guidelines, as described by Smiraglia,<sup>25</sup> provide acceptable access points for archival descriptions?<sup>26</sup> If not, can better guidelines be devised for indexing archival materials using LCSH?
- Within the context of useful application guidelines, would the terminology contained in LCSH suffice or not?

### The Need for Archival User Studies

Who uses archives? How do they use them? And are they satisfied with the results? These questions are frequently in the archival news, and the need for answers is painfully evident in the context of archival subject analysis.<sup>27</sup> If the answers to these questions were known, meaningful answers to the questions posed in the previous paragraph would also be possible.

Archivists often assume that their users are interested principally in high recall and are willing to endure low precision to achieve it.<sup>28</sup> In other words, they believe users want above all to retrieve *everything* of interest and will suffer through mounds of irrelevant material to achieve this end. Janice

<sup>23</sup>Smiraglia very usefully describes "depth level" indexing in Smiraglia, "Subject Access to Archival Materials Using LCSH," 65-74.

<sup>24</sup>Although this issue may seem niggling to some archivists, the retrieval implications of inconsistent practice are extremely significant, given that most online catalogs require searchers to select an author *or* subject index for any given search. In a discussion on the ARCHIVES electronic bulletin board (listserv address archives@indycms) between 3 and 16 March 1992, Steve Hensen expressed his view (on 5 March) that 7xx fields "should be reserved for a true 'joint-authorship' relationship to the described materials (cf., APPM 2.2, 2.3)," whereas 6xx fields are for "name indexing" on the theory that, from an archival point of view, a file of correspondence is more important for what it reveals *about* the person or agency that wrote it than it is for the fact of its authorship." Unfortunately, this distinction is not clearly expressed in APPM or any other authoritative source, and there appears to be little consensus in the field, judging from records found in RLIN and other databases and from the various comments submitted during the listserv discussion.

<sup>25</sup>Smiraglia, "Subject Access to Archival Materials Using LCSH."

<sup>26</sup>The application guidelines followed by the Library of Congress and by libraries that use LCSH are published as Library of Congress, Office for Subject Cataloging Policy, *Subject Cataloging Manual: Subject Headings*, 4th ed. (Washington, D.C.: Cataloging Distribution Service, 1991). For a discussion of some particular problems encountered when following these guidelines for indexing forms of material, see Jackie M. Dooley and Helena Zinkham, "The Object as 'Subject': Providing Access to Genres, Forms of Material, and Physical Characteristics," in *Beyond the Book: Extending MARC for Subject Access*, edited by Toni Petersen and Pat Molholt (Boston: G.K. Hall, 1990), 43-80.

<sup>27</sup>For an excellent overview of the need for archival user studies, see Lawrence Dowler, "The Role of Use in Defining Archival Practice and Principles: A Research Agenda for the Availability and Use of Records," *American Archivist* 51 (Winter-Spring 1988): 74-86.

<sup>28</sup>Michelson, "Description and Reference in the Age of Automation," 199.

Ruth recently has questioned this assumption;<sup>29</sup> in the context of subject access, it is critical that an answer be found. If high recall is paramount, archivists should focus on providing broad subject access to all collections. If precision is also required, they must learn to assign specific subject descriptors in a consistent manner.

Another typical assertion is the notion that "known-item" requests (the simplest example of perfect precision), such as "I need the first edition of James Joyce's *Finnegan's Wake*," are nonexistent in the archival world. But might not a researcher wish to locate the first *manuscript* version of this work? And might not a request for Joyce's correspondence or financial records be construed as a known-item search? We should consider defining "known item" in the context of manuscripts and archives, discovering how often researchers do in fact need "known items," and determining how archival access systems can best respond to the need for precision in searching.

One of the great strengths of archivists is their familiarity with the scholarly research process, and this knowledge can be put to good use in evaluating use and users. It is critical, however, that unsubstantiated biases and impressions now skew this evaluation. By way of comparison, it was once routine for rare book libraries to forgo subject access, based on the snobbish premise that scholars who deserve to use rare materials are intimate with the source materials in their subject areas and therefore will require only a catalog with author entries. This approach has been discredited, in part because it has become clear that even experts sometimes need subject access, particularly as research becomes increasingly interdisciplinary. Studies of online catalog use in libraries have demonstrated that even the most sophisticated users find subject

searching critical under certain circumstances.

Archivists also are aware that scholarly interests and methodologies change radically over time. In recent years, for example, many historians have shifted their interests from elites to masses; in subject access terms, this means a change from people of known name and title to those of anonymous name and no title at all. A writer's occupation, birthplace, and approximate life dates have assumed new importance.

Who are the users? For whom are these elaborate systems of access being constructed? Archivists often want to believe that most users of archives and manuscript repositories are sophisticated historians and other academics doing "serious" research, but in some environments such users are far outnumbered by the genealogists.<sup>30</sup> What ramifications might this have for the type of subject access archivists should provide? In spite of anyone's value judgments about the importance of various types of research, users are users, and one must consciously evaluate the effect of various indexing methodologies on different types of users, be they professionals, students, or amateurs.

Why are certain collections in MARC AMC databases heavily used, whereas others, ostensibly of equal research value, are not used at all? Is it possible that materials currently accessible only through provenance might find their way to additional users if a variety of subject access points were added? This question merits unbiased investigation, particularly now that many records are available to researchers via unmediated online searching.

The existence of large MARC AMC databases makes possible a new type of research methodology: the study of online

<sup>29</sup>Janice Ruth, "Educating the Reference Archivist," *American Archivist* 51 (Summer 1988): 273.

<sup>30</sup>See, for example, Dowler, "The Role of Use," 76, and Lytle, "Intellectual Access," 66.



searching transaction logs. Although such studies potentially have great merit, researchers must carefully consider the complex context of online searching. For example, if few searches on personal names are performed in an online catalog's *subject* index, the researcher must ask whether users know such a search is possible before concluding that users feel it is unnecessary. If users ignore a particular index entirely, is it because they cannot easily determine the type of data the index contains? If users frequently are frustrated by "zero hits" subject search results, the researcher should ask whether this might in part reflect inconsistent indexing and should consider whether cross references from a controlled vocabulary would help, rather than just concluding that subject indexing is useless for access to archival materials.

Designing and implementing meaningful user studies is no small feat, but considerable help can be found in the literature of librarianship and information science, where researchers have worked in pursuit of effective methodologies for decades.<sup>31</sup>

### Recent Progress

The footnotes appearing throughout this paper serve notice of the burst of recent activity on issues related to subject access. As mentioned earlier, it is notable that the national working groups on descriptive standards in the United States and Canada have commented on the importance of subject access and called for the development of archival guidelines for application of ex-

isting thesauri. Harriet Ostroff's paper on subject access prepared for the American working group suggested needs related to specific MARC AMC data elements,<sup>32</sup> and the Canadian working group is in the process of drafting a report on subject access.<sup>33</sup>

In a 1990 issue of *Cataloging & Classification Quarterly* entirely devoted to use of the MARC AMC format, Smiraglia described how to index archival materials with LCSH, and others discussed subject access to archival graphics, sound recordings, and maps.<sup>34</sup> Authors of essays in the 1990 compilation titled *Beyond the Book* described subject indexing of art objects, three-dimensional artifacts, moving-image films, and other primary resources, as well as access to archival materials by form of material and by function.<sup>35</sup> Various reports, such as those previously cited from Harvard and the Historical Documents Study, have described scholars' desire for integrated searching of primary and secondary resources regardless of form of material or type of repository.

Other recent activities also indicate the breadth of archival subject access activities. For example, at the invitation of Toni Petersen, editor of the *Art and Architecture Thesaurus*, and Kathleen Roe, chair of SAA's Committee on Archival Exchange (CAIE), seven librarians and archivists met in March 1991 to begin drafting an archival application guideline for the AAT. It quickly became clear that extensive groundwork would have to be laid before such a guideline could be written, so the group instead

<sup>31</sup>See, for example, Miranda Lee Pao, "Uses and Users," in *Concepts of Information Retrieval* (Englewood, Colo.: Libraries Unlimited, 1989) 40-53; Tefko Saracevic and Paul Kantor, "A Study of Information Seeking and Retrieving. II. Users, Questions, and Effectiveness," *Journal of the American Society for Information Science* 39, no. 3 (1988): 177-96; and Ronald R. Powell, *The Relationship of Library User Studies to Performance Measures: A Review of the Literature* (Urbana, Ill.: University of Illinois, GSLIS Occasional Paper, no. 181, 1988).

<sup>32</sup>Ostroff, "Subject Access to Archival and Manuscript Material."

<sup>33</sup>The Canadian report, which was not finished as of May 1992, will be issued as Bureau of Canadian Archivists, Planning Committee on Descriptive Standards, *Report of the Subject Indexing Working Group*.

<sup>34</sup>Richard P. Smiraglia, ed., *Cataloging and Classification Quarterly* 11, nos. 3,4 (1990).

<sup>35</sup>Toni Petersen and Pat Molholt, eds., *Beyond the Book: Extending MARC for Subject Access* (Boston: G.K. Hall, 1990).

worked to define the range of specific issues that would have to be addressed in such a guideline. In essence, the group recommended that consistent approaches must be developed for use of the array of subject-oriented data elements available in the MARC AMC format which are discussed in this paper: topic (generic and specific), time, place, form, occupation, and function.

In response to the agenda mapped out at the AAT/CAIE meeting, CAIE plans to prepare a series of grant proposals designed to assess researcher needs for archival subject access, describe current problems, identify potential options for archival indexing, conduct a series of online indexing experiments, and, ultimately, develop a set of recommendations for providing subject access to archival materials.<sup>36</sup> Since many subject-oriented facets overlap with aspects of provenance (these include function, occupation, time, place, and form of material), CAIE's concurrent efforts to evaluate and enhance the USMARC authorities format will be closely linked.

At the 1991 SAA conference in Philadelphia, a call for expressions of interest in a Description Section discussion group on subject access elicited widespread interest. Also, the *SAA Art and Architecture Thesaurus* roundtable appears to be thriving.

At the invitation of the AAT and the Bibliographic Standards Committee of the ALA Rare Books and Manuscripts Section, and with funding from the Council on Library Resources, archivists and librarians responsible for editing a variety of thesauruses for indexing genres and forms of material met in August 1992 in an attempt to reconcile differences in terminology among thesauri. Such differences complicate both the selection of a thesaurus and retrieval across databases.

Clearly, interest in subject access is widespread in the archival community. Active participation must be equally broad if solutions that meet the needs of as many archivists and users as possible are to be found.

### Recommendations

Each of the authors and working groups mentioned above has called out aspects of an agenda for improvement of subject retrieval of archival materials. The following list represents some of the seemingly more pressing needs.

- Archivists should confirm their professional consensus that it is necessary to provide subject access to archival materials in order to supplement existing access by provenance, recognizing that subject access includes not only generic topics but also specific named persons, organizations, places, and events, as well as time, place, form of material, occupation, and function.
- The archival community should support and participate in CAIE efforts to study researcher needs for subject access, to assess the value of existing subject vocabularies such as *Library of Congress Subject Headings* and the *Art and Architecture Thesaurus*, and to develop application guidelines oriented toward archival needs.
- If there is a consensus regarding the need for a thesaurus of broad subject areas for use in MARC AMC field 072, archivists should develop such a thesaurus and encourage its widespread use.
- Archivists should identify the subject-oriented fields in MARC AMC which have not been implemented for display and/or indexing purposes by vendors of online systems and should present a united professional front to change this.

<sup>36</sup>Written communication from Kathleen Roe, CAIE chair, to committee members, 10 March 1992.

- Archivists should support and participate in CAIE efforts to enhance the USMARC authorities format in order to add contextual information to name authority records.
- Archivists must investigate user needs and behavior in the specific context of subject-oriented queries.
- SAA should continue to develop and implement a strong education program to train both new and experienced archivists in subject indexing and authority control techniques.

Michelson has stated that "extreme inconsistency in describing materials presents the key problem facing archival reference."<sup>37</sup> There will be no quick fix,

given the complexities of designing effective user studies, developing subject access guidelines, and learning to use controlled vocabularies, let alone creating basic MARC AMC descriptions of the tens of thousands of collections that remain inaccessible beyond their repository doors. Many of the specific problems to be solved require archivists to tackle issues that may be likened to the proverbial question of the number of angels that can dance on the head of a pin. Nevertheless it is from just such minutiae that meaningful standards are forged. Archivists have traveled an extraordinary distance in the last decade in their development of descriptive standards, but the journey is long and far from over. All aboard?

---

<sup>37</sup>Michelson, "Description and Reference," 194.