# CASE STUDIES

# RLIN, AMC, AND RETROSPECTIVE CONVERSION: A CASE STUDY

PATRICIA CLOUD

*ABSTRACT:* For one year, beginning in October 1984, the Northwestern University Archives was involved with eleven other large research libraries in an archives and manuscripts retrospective conversion project sponsored by the Research Libraries Group (RLG) and funded in part by a grant from the National Endowment for the Humanities (NEH). This article is based on the experiences of one institution involved in the project, and considers some of the specific steps that RLG has taken to implement the new Archival and Manuscripts Control (AMC) format, assesses the ways in which the new format fits into the Research Libraries Information Network (RLIN) national bibliographic database, and recounts the experience of creating records in RLIN.

During 1984 and 1985, the Northwestern University Archives was one of twelve participants in a project sponsored by the Research Libraries Group (RLG). The project's goals were to expand the newly established national online database for archival and manuscripts material, and to involve a number of large research institutions in use of the newly available MARC (for MAchine-Readable Cataloging) format for control of such material.

USMARC Format for Archival and Manuscripts Control (AMC), published by the Library of Congress in 1985, is one of seven such formats devoted to recording bibliographic information in machine-readable form. MARC has been part of the Library of Congress's cataloging distribution service since 1968; it was not, however, until 1985 that the MARC format provided a satisfactory tool for control of archival and manuscript materials. The earlier MARC Manuscript Format, published in 1973, was not widely used, principally because it was designed chiefly to accommodate single-item manuscripts and proved inadequate to the task of collection level description.[1]

The revision of the MARC format for archives and manuscripts was a highly cooperative venture. The Society of American Archivists, through its National Information Systems Task Force (NISTF), contributed heavily to both

the development of a theoretical framework for the use of a standardized format for control of archival and manuscripts material and, in extensive consultation with the Library of Congress, to the format revision itself.[2] The SAA remains involved in the maintenance of standards for the new format.

Another important participant was the Research Libraries Group. RLG is a consortium of university libraries and research institutions whose aims include the development of shared resources among its member institutions. Among RLG's principal programs is Technical Systems and Bibliographic Control, under which falls RLG's bibliographic utility, the Research Libraries Information Network (RLIN). RLIN's online database functions chiefly as an aid to library acquisition and processing of research materials, and additionally as an increasingly valuable aid in research and reference work. The database contains records reflecting significant portions of the collections, in all media, of over sixty research institutions, and so serves as a powerful tool for bibliographers and others concerned with collection development. In 1982, RLG began developing a system to be established in RLIN which could automate accessioning, processing, and retrieval of information about archival and manuscript materials.[3] Throughout the life of the National Information Systems Task Force of the SAA, RLG contributed to its deliberations as well.[4]

Since the preliminary approval of the format in 1983, RLG has vigorously pursued its implementation of the AMC format and the development of the AMC database in RLIN. A pilot project involving Yale, Stanford, and Cornell University libraries and the Hoover Institution,[5] begun in January of 1984, had by October of that year produced approximately 3,500 AMC records and had helped to establish working standards for record creation. RLG also has sought actively the participation of numerous repositories, among which are the National Archives and the State Historical Society of Wisconsin.[6] In addition, RLG has sponsored a cooperative project, involving the retrospective conversion of approximately 21,000 records describing archival and manuscripts material held by twelve participating institutions, including Northwestern. The project was funded jointly by the National Endowment for the Humanities (NEH) and the Pew Memorial Trust. It began in October 1984, and continued through March of 1986. The institutions involved in the project were Brown University, Rutgers University, Columbia University, New York University, the New York Historical Society, the University of Michigan, Johns Hopkins University, the State University of New York at Buffalo, the University of California at Berkeley, Brigham Young University, Dartmouth College, and Northwestern University.

RLIN offers several major advantages to archivists. Many wonder whether automation will genuinely improve information retrieval. RLIN's immensely powerful searching capacity provides great flexibility in this regard. The database includes over thirty searchable indexes. The "general" (i.e., utility-wide) indexes include most significantly: personal name, title word or phrase, corporate name, and subject word or phrase. Searches are very easy to construct: a personal name search can be full or truncated; key words can be strung together with or without boolean operators (logical characters that expand or refine searches); title phrases can be used as they are encountered. Several indexes may be included in one search, and it is possible to search across RLIN files, to retrieve, for instance, the published and manuscript materials of an

individual or concerning a subject. Local indexes are available for searching the records of an institution. These include call number (or record group number), record status, and, unique to the AMC file in RLIN, donor phrase and donor word. The boolean capacities of RLIN encourage the greatest precision in constructing searches and permit refinements of terms while conducting searches.

The RLIN AMC file is designed to accommodate the archival practice of using personal names to enhance subject access. Where in other RLIN files, a personal name search would yield bibliographic records only of material authored *by* an individual, in the AMC file, such a search would call up records of material both by and about an individual. For example, searching the name James Joyce in the AMC file in RLIN would provide the searcher with records of material, such as literary manuscripts, created by Joyce; it would yield, in addition, records of material concerning, but not created by, Joyce, such as correspondence of associates in which Joyce is noted. RLIN accomplishes this by indexing personal names in the Personal Name Index (which in AMC includes the 100, 600, and 700 fields). The personal name may be coded as either principal creator (100), correspondent (700), or subject (600) and be found in one search of the Personal Name Index in RLIN.[7]

RLIN has taken another special step in implementing the AMC format, which allows for administrative control as well as descriptive control of collections. Two fields that have received special attention in the revision of the manuscripts format are the 541 "acquisitions" and 583 "action taken" fields. Both were "exploded" in MARC, that is, provided with many additional subfields, in recognition of the great significance attached to actions taken on archival collections, such as processing, and to the method of acquisition of collections. Within these fields, the MARC format allows specification of the source and means of acquisition, accession number, description of actions taken (such as processing, exhibiting, lending) on the collections, and so on.[8] RLIN chose not to include these fields in the bibliographic segment (that is, the primary descriptive section of the record, all of which is public in RLIN; see Figure 1), but rather to create a separate, associated segment in which the material of the 541 and 583 fields for each item can be handled. This "archival control segment" (see Figure 2) is represented by one or more screens which appear following the bibliographic segment wherein all the 541 and 583 subfields are presented with mnemonic tags. The display makes fairly intensive use of these fields quite easy: information can be repressed; some or all of it can be displayed in a split screen with the bibliographic segment (see Figure 3). Currently it is possible to record as many as forty administrative "actions" within the archival control segment for most records.[9] This means that it is possible to keep track of a wide range of actions associated with a collection or its parts, including accessioning, processing, exhibiting, lending, and preservation and conservation actions. The National Archives is currently testing the flexibility of the archival control segment in a project with RLG designed to track a sample of records through their life cycle.[10]

Northwestern University Archives decided to participate in RLG's AMC retrospective conversion project for several reasons, some general and some specific to the repository. As an RLG member, the Northwestern University Library supports the value of regional and national utilities as online union

```
AMC/PROD  Archival     FUL/BIB    ILNG85-A615              Catalog          ILNG-TJD
Record 1 of 1
+
  ID:ILNG85-A615       RTYP:d     ST:p            NLR:        MS:  EL:  AD:06-05-85
  CC:9554  BLT:bd       DCF:     CSC:d    MOD:    PROC:b                 UD:08-13-85
  PP:ilu     L:eng       PC:i     PD:1932/1972    REP:
  MMD:       OR:     POL:      DM:      RR:          COL:       EML:       GEN:    BSE:
  040      IEN #cIEN#eappm
  100 1    Wilde, Oscar,#d1854-1900.
  245 13   An ideal husband :#kproduction file,#f1932-1972.
  300      ca. 35 items.
  520      Scripts, production notebooks, lighting plot, press cuttings, programs,
           stage plan, set designs, cast, music, wardrobe, and property lists, and
           costume sketches from the October 18, 1932 production and the 1945 and
           1972 revivals by the Dublin Gate Theatre.
  555      Container list in Special Collections Department.
  600 10   Wilde, Oscar,#d1854-1900.#tAN IDEAL HUSBAND#xStage history#zIreland.
  655  7   Scripts.#2ftamc
  655  7   Clippings.#2ftamc
  655  7   Programs.#2ftamc
  655  7   Lists.#2ftamc
  655  7   Sketches.#2ftamc
  773 0    #7c2bc#aDublin Gate Theatre Company Papers, 1928-1979.#w(CstRLIN)ILNG85-
           A318


AMC/PROD  Archival     FUL/BIB    ILNG85-A615              Catalog          ILNG-TJD
Record 1 of 1
+
  851      Special Collections Department,#bNorthwestern University Library,#c1935
           Sheridan Rd., Evanston, IL  60201.
```

*Figure 1*
*RLIN Full Bibliographic Display*
*AMC Record*

```
AMC/PROD  Archival     FUL/ARC    ILNG85-A615:1            Catalog          ILNG-TJD
Record 1 of 1
UPD
   RGPN Spec Ms                          INS
   PCID 1        PCDP N                  PST P                        PSTD
   ACCN NA                               ACCD 1971
   MATL
   SRCE
   ADDR
   MTHD
   PPRI                  OWNR
   PLOC Spec

   ACID 1          ADP N        AINS
   ACT  Cataloged.
     AIDN
      TAC 04/15/85      TFAC              AIN          AIR
     CONT
     AUTH
      JUR
     METH
     SITE
     STAT
      AGT J. Carson
```

*Figure 2*
*RLIN Archival Control Segment Display*
*AMC Record*

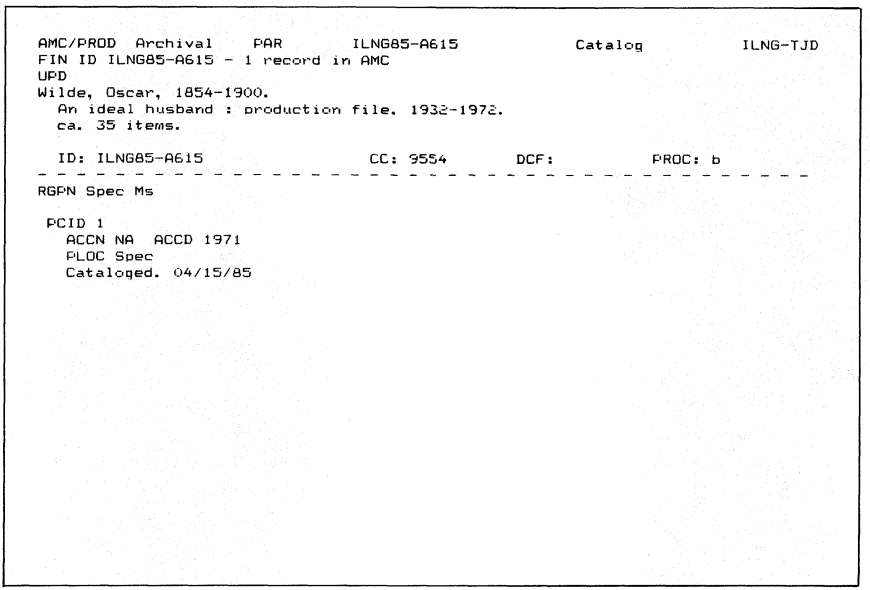


*Figure 3*
*RLIN "Partial" Display*
*AMC Record*

lists; the University Archives felt the significance of creating access to collections to be especially important where archival and manuscript materials are concerned. The new format offered a means for integrating archival and manuscript holdings into a national bibliographic network, and for gaining experience in the use of the new format.

The University Archives felt a particularly strong need for experience using MARC. The Northwestern University Library has been involved since 1970 in developing and refining NOTIS (Northwestern On-Line Total Integrated System), its own automated system for technical services, public access catalog, authority record management, and circulation. Anticipating the implementation of the AMC format in NOTIS, the University Archives felt that the knowledge to be gained by using AMC in RLIN would be of value. NOTIS implemented AMC in June 1986; the records created in RLIN by the University Archives are ultimately to be transferred to NOTIS. Future records created in NOTIS are themselves ultimately destined to be loaded into the RLIN database.

In archival automation, as elsewhere, it is very difficult to isolate one issue from another: getting involved with AMC means getting involved with *Anglo-American Cataloging Rules, 2nd. ed.* (AACR2),[11] with the Library of Congress Name Authority File (LCNA), and with all else that pertains to library standards of description. No review of the arguments for standardized description, or for the choice of one particular standard rather than another will be attempted here,[12] except to say that automation *per se* does not require

the kind of standardized description with which the Northwestern University Archives became involved. Where, however, the issues of integration of archival and manuscript materials in the larger universe of the bibliographic utility are concerned, standardization of some sort is inevitable. The University Archives accepted the standards developed by RLG for AMC.

What are these standards? In keeping with their wish to encourage the growth of the AMC database, RLG established a minimum standard for records which is quite basic:[13] the only required fields are main entry, title, dates, and extent (of material) or physical description. From this minimum, which serves in part to accommodate newly acquired, unprocessed collections, it is possible to elaborate records considerably. There are practical limitations on the length of records which can be successfully indexed in RLIN, but it is possible for a rather long record to accommodate well over 100 added entries (or additional access points). Shorter bibliographic text can be associated with more added entries, and vice versa. And, as noted before, up to forty "control structures" are permitted within the archival control segment, with bibliographic text of 1.5 pages and thirty-five added entries.[14]

Description is based on Steven Hensen's revision of chapter four of AACR 2, *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries.*[15] Topical subject headings are not required, but where any are used, one at least must be drawn from Library of Congress Subject Headings (LCSH). It is quite proper, in addition, to use locally established subject terms, which are specially tagged in the AMC record. Form of personal and corporate names must be established according to AACR 2; this involves searching of the LC Name Authority File and the use of AACR 2 principles to establish names not found there.

On the need and value of standardization of name form, there is not a true consensus in the archival community. Some feel that the imposition of bibliographic standards will result in the loss of important information, since many archivists have traditionally used the fullest approach to establishing names, including full middle names for individuals and place names for institutions. AACR 2, of course, dictates use of the most common form whatever it may be.[16] Henson recommends including fullest form in the note fields of a record.[17] New procedures will involve, among other things, the use of bibliographic authority files for the establishment of name form. Most institutions will establish their own authority routines, which will depend on the resources available to them. The LC Name Authority File is available online through RLIN,[18] and searching is charged according to the amount of central processing unit time which is used. In many places, particularly those associated with research libraries, the LC file is available directly either online or on microfiche. An authority routine can involve searching the local authority file of an institution, the LC file, and a third step may be searching of the *National Union Catalog of Manuscript Collections* (NUCMC) or the *National Union Catalog* (NUC). Clearly the most efficient searches will be those that can be ended with the first or second resource consulted. Unfortunately, archivists and manuscript curators will have a high percentage of authority searches involving several steps. These will be commensurately time and money-consuming, since many of the names from archival collections will not be those of authors,

and so will not appear in the typical bibliographic sources. Ideally, an authority routine would be completely exhaustive and include searching of local authority files, the LC Name Authority File, NUC, NUCMC, and other appropriate reference sources.

The routine followed by the University Archives was abbreviated and did little more than conform to the minimum standard. Nevertheless, even the most straightforward records took approximately one to three minutes per added entry for authority searching. Where a conflict was discovered, it became necessary to devote far more time to establishing forms. The Northwestern University Archives profits from the availability of the NOTIS online authority file, which has been in operation since 1982 and has currently over 120,000 records. Searches of both the NOTIS authority file and NOTIS bibliographic file online, and LC Name Authority File online through RLIN were made. Since virtually no manual files were used, those repositories with manual files can expect to spend more than the one to three minutes reported here per added entry in doing straightforward authority work. Thirty-six percent of the added entries were found in the NOTIS authority file; 3% occurred in the bibliographic file and not elsewhere; 22% occurred in the LC Name Authority File, and fully 37% were not found in any of the sources consulted and had to be established according to AACR 2 as new name entries. No new names were added to the NOTIS authority file by the University Archives during this project, although it would undoubtedly be wise to incorporate this step into any future projects. What this may mean for repositories considering the cost of standardization is not yet clear, but it will need to be evaluated.

The University Archives's involvement in the RLG AMC RECON project began in October of 1984. When the project began, creation of approximately 1400 records was anticipated from four departments in the Northwestern University Library. In addition to nearly 150 collections from the University Archives, which included personal papers of many faculty members and records of institutions such as the Dearborn Observatory in Evanston, the project planned to draw on the manuscript collections of the Northwestern University Library's renowned Melville J. Herskovits Library of African Studies, the Department of Special Collections, and the Music Library. Each of these departments holds significant manuscript materials, which range from personal letters of John Maynard Keynes to papers and scores of conductor Fritz Reiner and composer John Cage. Needless to say, since this material has for years been housed, inventoried, and controlled by different departments in the library, a great variety of techniques for arrangement exists and the form and degree of thoroughness with which these collections have been described vary considerably. It was determined that 232 records were inappropriate for inclusion in the project, for reasons either of special cataloging problems or of low potential for scholarly use. In the end, 1168 records were created during the year the project ran.

There are two basic types of control for these scattered materials: a small proportion, approximately 170, of the targeted collections had full descriptive inventories, with thorough historical statements and summaries of the collections, in addition to container lists and, in many cases, lists of correspondents. These were, not surprisingly, in certain ways the least troublesome records to create. The AMC format permits lengthy statements and summaries which

could in many cases be drawn directly from descriptive inventories; notes on provenance and restrictions could also be taken directly from these full inventories. Added entries could in the best of circumstances be drawn directly from container or folder lists, and in general the establishment of subject headings is quite simple in these cases. It is possible to generate a fairly complete AMC record for a large manuscript collection by coding directly from a full descriptive inventory, since AMC has been devised to accommodate this form of record.

By far the greatest number of records targeted for the project (approximately 900) were in the form of card catalog records. The greatest proportion of these came from the Department of Special Collections and the Music Library, where most records comprised either single-item manuscripts, of one or many leaves, or small groups of manuscripts, of approximately two to three pieces each. These materials are controlled by single-item manuscript catalog descriptions, of the sort which was for years the preferred form. Most of these items merit individual description, and this approach to manuscript description is just as appropriate in the AMC format as is the collection level approach suitable for larger groups of materials. Unfortunately, no card file is perfect. The entries in the two main manuscript card files dated from several different eras and from several different cataloging regimes, both national and local. The style of cataloging varied greatly, as did its level: many of the records encountered were simply too sketchy to use as they were, lacking dates or other significant information. It was necessary to consult the manuscripts themselves to achieve minimal cataloging standards in 19% of the records. In another 23% of the records, reference work beyond the usual authority routine was required. In many cases either the form or the choice of main entry had to be changed to conform with AACR 2. This involved considerable record keeping, to avoid the problem of creating access to a collection or item in RLIN which remained accessible at Northwestern only in a highly eccentric way.

The project budget paid for one-half the time of one project archivist, and one work/study student at twelve hours per week. The project coordinator was initially to have spent one-half time on the project, but found that the demands of the schedule required nearly full-time attention. As work progressed, the project archivist spent nearly all project time coding worksheets. Some authority searching and most inputting were handled by a succession of work/study students. The project coordinator divided time among training, coding, authority work, review and proofing of all worksheets, and final proof of online records.

For those anticipating involvement in a RECON project, the two most important questions must be, How much will it cost? and How long will it take? It is extremely difficult to generalize from one institution's experience to another's. At Northwestern, costs ran somewhat higher than expected, largely because record creating took longer than anticipated. The original estimate of total cost per record (a figure that included RLIN charges as well as staff time) was approximately $18 per record; the true cost was closer to $21. Time spent on each record coding, searching, inputting, and reviewing was recorded. For the larger, multi-piece collections, which had an average of four added entries apiece, coding time averaged forty-one minutes, inputting approximately twenty-two minutes. Searching and reviewing added approximately three to five minutes to the total. For collections controlled by card file,

which generally constituted between one and twelve items per record, coding took an average of eleven minutes and inputting another nine minutes; again, searching and review added three to five minutes to the total. However, these figures pertain to records in which no difficulties presented themselves, and, as already stated, in nearly half of the cases (42%), it was necessary to consult the collection itself or to do some form of research on the material. Authority work or cataloging questions which presented problems were kicked out of the workstream and handled separately. If problems of evaluating material, tagging questions, and authority work could have been isolated or eliminated, the total per record figures of twenty-five and seventy minutes could have some value in establishing theoretical conversion rates. It should be noted, however, that if these theoretical rates are applied to the project we have just completed, we should have been able to finish within fourteen weeks what took us a full year. What is probably a more useful figure is one calculated on the basis of the total project: hours spent through the year divided by total output yields the figure of 2.7 hours per record (or about $21 per record).

Direct benefits of such projects are difficult to measure. Ideally, a repository embarking on a retrospective conversion project would gain better control of the material described by imposing basic standards of description on all the existing finding aids. Rationalization of terms—either by use of LC authority files or by the imposition of a local authority system—is another potential benefit which can greatly improve access to materials even without automation. Wider access to information regarding a repository's holdings is a major goal of any retrospective conversion project, but this wider access may not bear fruit for some time. Both major national bibliographic utilities, Online Computer Library Center, Inc. (OCLC) and RLIN, are available directly to patrons only in rare instances; these utilities are designed primarily for use by librarians in acquiring mass-produced materials. Therefore, AMC records will generate few reference inquiries immediately. Ultimately, AMC records on a national bibliographic database will become more and more available for direct patron access, but their presence there now will not be likely to generate much new activity for some time to come.

AMC can replace numerous manual files within a repository, but there is no reason that it need to. The ease of maintaining an online accessions record may be offset by the possible difficulties of using such records where only one or two terminals can be purchased and where they may not always be available for reference or processing activities. In many repositories, where collections are not of overwhelming size, maintenance of manual files, in addition to or instead of using AMC for "archival control," may be preferred, particularly where no local system exists to complement a national or regional utility. In such cases, transactions costs can often seem not worth the gain in speed of access and easy file maintenance.

Repositories wherein AMC can be most readily integrated with existing files will certainly automate most painlessly and profit more quickly from doing so. Where, for example, an authority file for headings is already in place, modification to suit national standards need not be prohibitively time-consuming. On the other hand, where varying qualities of description and idiosyncrasies abound, converting to AMC may prove immensely disruptive. It hardly needs to be pointed out, however, that for most, the first experiences with AACR 2

and MARC will involve a great deal of learning by doing, a time-consuming but valuable process.

*ABOUT THE AUTHOR:* Patricia Cloud is Assistant University Archivist at Northwestern University. This essay is a revised version of a paper presented at the meeting of the Midwest Archives Conference, Columbus, Ohio, October 1985.

## NOTES

1. Katharine D. Morton, "The MARC Formats: An Overview," *American Archivist* 49 (Winter 1986): 21-30.
2. Elaine Engst, "The Development of the RLIN AMC System," *Documentation Newsletter* 9 (Fall 1983): 1-6; Richard H. Lytle, "An Analysis of the Work of the National Information Systems Task Force," *American Archivist* 47 (Fall 1984): 357-65; Mary Alice Ball, "The Development of the USMARC Archival and Manuscript Control Format: The Politics of Negotiation" (Masters thesis, The University of Chicago Graduate Library School, 1984); and "AMC Offers New Access to Nation's Archival Resources," *Research Libraries Group News* 5 (September 1984): 3-5.
3. Ball, "The Development of the USMARC . . . ," 24.
4. Engst, "The Development of the RLIN AMC System," 3-4; "AMC Offers New Access," 4.
5. Engst, "The Development of the RLIN AMC System," 3-4.
6. Concerning the special projects undertaken jointly by RLG and these two institutions, see *MAC Newsletter* 12 (January 1985): 8-9; Research Libraries Group, press release, January 18, 1985.
7. *RLIN Archival Control Manual* (Stanford, Ca., 1984), 15.
8. For a detailed description of the workings of these fields in RLIN, see Alan M. Tucker, "The RLIN Implementation of the MARC Archives and Manuscript Control Format" (RLG, n.d.).
9. *RLIN Operations Update*, 32:11-12.
10. RLG, press release, January 18, 1985.
11. *Anglo-American Cataloging Rules*, 2nd ed. (Chicago: American Library Association, 1978).
12. See Lytle, "An Analysis of the Work . . . ," 360-61; Tucker, "The RLIN Implementation . . . ," 3.
13. Tucker, "The RLIN Implementation," 4; "Standard for Archives and Manuscripts in RLIN," *RLG Document* 84-78 (July 1984), and revision, *RLG Document* 85-67 (December 1985).
14. *RLIN Operations Update*, 32:11-12.
15. Steven L. Hensen, *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies and Manuscript Libraries* (Washington, D.C.: Library of Congress, 1983).
16. RLG AMC Task Force, minutes, January 8-9, 1984, 4.
17. Hensen, *Archives, Personal Papers*, . . . , 6.
18. Now current to within 24 hours, due to the establishment of a direct link between the RLIN and LC authorities databases. *LC Information Bulletin* 44 (October 28, 1985): 305.