JANET SWAN HILL

The Northwestern Africana Project: An Experiment in Decentralized Bibliographic and Authority Control

The Northwestern Africana Project was designed to demonstrate that a system of decentralized bibliographic-control centers could be an effective method for building a comprehensive, consistent national database for use in a national library network. Utilizing funding granted by the National Endowment for the Humanities and by the Carnegie Corporation, Northwestern sought to show that a high-quality database, that is, one that conforms to Library of Congress cataloging standards and heading usage, could be prepared and maintained at a location remote from LC. This database would be transmitted to a central location (in this case the Library of Congress) for incorporation into a common database and for distribution.

INTRODUCTION

During the 1970s, Northwestern University Library (NUL) engaged in several activities involving cooperative creation and use of machine-readable data. As a COMARC (Cooperative MARC) participant we had observed the difficulties of attempting to ensure uniformity of cataloging standards among disparate cataloging agencies.[1] As the operator, on behalf of the National Library of Venezuela, of a special project in which a single file of bibliographic citations to works about or from Venezuela or by Venezuelan authors was created by consolidating information from multiple sources, we had noted the problems associated with creating a joint database and trying to impose uniformity of citation without the actual works in hand.[2,3] Our experience with these projects, combined with increased nationwide discussion of a national bibliographic network, led us to investigate ways to test the feasibility of producing a high-quality bibliographic database whose records would be consistent with LC practice, which could be submitted to a cen-

Janet Swan Hill is head of the Catalog Department, Northwestern University Library, Evanston, Illinois.

tral agency to form one component of the national data file.

Northwestern's Africana acquisitions provided an ideal body of materials with which to work, not only because they were readily identifiable and self-limiting, but also because of the relatively small number of Africana titles found on MARC tapes compared with other classes of works. Moreover, Northwestern University was already engaged in the collection and compilation of bibliographic information for Africana in connection with the publication of the *Joint Acquisitions List of Africana (JALA)*.[4] Northwestern's NOTIS (Northwestern On-line Total Integrated System)—a locally developed, integrated on-line computer system encompassing a full range of library services including acquisitions, cataloging, serial control, circulation, and accounting functions—provided the necessary computer support.[5]

The Library of Congress indicated that it would welcome such an experiment, and a project to "establish at Northwestern University Library a national center for the control of bibliographic data relating to African materials as a component in a national library network" was funded by the National Endowment for the Humanities with partial

matching funds from the Carnegie Corporation. Under its terms, Northwestern was to create and maintain cataloging and location data, both for its own Africana materials and for items acquired by libraries contributing entries to *JALA*. To assure the quality and consistency of the records, a project member stationed at the Library of Congress would search all titles and entries in LC's catalogs, oversee and code any authority records produced, and "negotiate" any problems with headings with the principal catalogers' offices. Eventually, many of the searching functions were to be carried out by staff at NUL using a cathode-ray-tube terminal (CRT) linked directly to LC's computer. Data created at Northwestern were to be transferred to the Library of Congress on tapes and eventually incorporated into the MARC database and distributed. The project ran from mid-1977 through 1979, at which time it was redescribed to focus on cooperative creation of authority records.

The first step taken was to create a database of 24,139 catalog records, representing materials in the Northwestern Africana collection cataloged since 1970. The records were extracted from the central NOTIS database by means of a computer search using location codes. Additional data began to be inputted almost immediately, both as a result of Northwestern's normal and Project-processing routines, and through the input of cataloging copy submitted by *JALA* contributors.

Another initial step sent the senior Africana cataloger and the head of the Catalog Department to visit the Library of Congress to consult with LC staff on specifications for the cataloging aspects of the project. These meetings resulted in a formal name for the project (Northwestern Africana Project, or NAP) and also brought the first intimations that certain hoped-for components might be either totally unavailable or significantly delayed. It was learned, for instance, that LC would not be able to incorporate NAP records into its MARC database during the project. Additionally, the computer link to LC would be delayed because of the difficulty of acquiring compatible equipment (an appropriate terminal was finally installed and operational on June 25, 1978). Once established, it would afford access only to the MUMS (Multiple

Use MARC System) Master Book File and APIF (Automated Process Information File) and could not access the various databases, such as authorities, that were not yet online at the Library of Congress. (Authorities eventually came up at LC and became available to Northwestern in late 1979.)

It quickly became clear that the Library of Congress' primary interest was in testing decentralized authority control, a matter which was only one of the NAP goals, although also of intense interest to Northwestern. Moreover, LC's internal requirements made the process of cooperating on authorities exceedingly complex.

PROJECT COMPONENTS

As work routines were set and modifications and enhancements to the NOTIS system were developed to meet project requirements, work began to be viewed as being of four types:

Input of Records from Northwestern

Provisional cataloging records for items received by Northwestern but not yet fully processed were selected for transfer to the NAP database according to potential interest or usefulness to other libraries. Records for Africana titles fully cataloged by Northwestern were automatically copied from NUL's central file to the NAP database.

Handling of Contributed Copy

Copy forwarded from cooperating libraries for inclusion in *JALA* was in card form. Cards were searched against the NAP computer file, new records were input, and duplicates were coded as added locations. Cataloging copy was edited as much as possible for correctness and consistency, while verification of headings was done separately as a part of the authentication procedures described below.

Verification of Headings

Each heading used in a record on the project database was sent to a NAP agent stationed at the Library of Congress. The agent checked each heading against LC's authority files (the Official Catalog) and advised the Evanston staff of the results of the search. All headings on records in the NAP file were made to conform with an established form in the Official Catalog, if one were found. Any

NAP record on which all headings agreed with those in the Official Catalog or established by the project was coded as "fully authenticated." Because a heading cannot be established without the work in hand, contributed records that had any unestablished heading could not receive the fully authenticated coding. Headings from Northwestern's provisional and in-process records were verified and coded as for contributed copy.

During the course of the project, more than 10,000 headings were searched, of which approximately 59 percent were names, 9 percent series, and 32 percent subjects. Of all headings searched, 77 percent were already established in the Official Catalog, and 23 percent were new. More than 95 percent of subjects searched were found, while 76 percent of series and 66 percent of names were already in LC's catalog.

Transmittal of Data
to the Library of Congress

Specifications for types of records to be included on the database, special coding conventions, and the format for record transmittal had been settled at the first meetings and in later correspondence. Each quarter, beginning in April 1978, a tape was made of all in-process, provisional, or fully cataloged records added during that period, plus any corrections, updates, or deletions (except where the only correction made was the addition of a location) that had been made in the interval. The selection of records was controlled by a system-maintained record creation or update date in the bibliographic record. Tapes were sent to LC in MARC communications format, using the new spanned-block format that the Library of Congress had requested; were converted by LC's Catalog Distribution Service into the Library of Congress internal format; and were reconverted to communications format prior to distribution.

PROJECT RESULTS AND PRODUCTS

The two most visible products of the Northwestern Africana Project were the computer production of an enhanced *JALA* and the availability for sale through the Library of Congress of the NAP database.[6] Other results were realized as well, and are described below.

Joint Acquisitions List of Africana

JALA has been published bimonthly for twenty years by the Melville J. Herskovits Library of African Studies of Northwestern University Library. It lists Africana acquisitions of nineteen cooperating libraries. Beginning in 1978 with volume 17, number 1, *JALA* was produced from the NAP database. Records were selected based on a cataloging date included in the holdings record. A modification of the NOTIS card-production program caused the machine generation of a two-column printed format. Other enhancements to the publication included the availability of Northwestern in-process information, the inclusion of multiple-location data, editing of bibliographic data for consistency and for correctness of headings, and regularity of production.

In January 1980, responsibility for production of *JALA* was transferred back to the Melville J. Herskovits Library. *JALA* continues to be computer produced, the database continues to be maintained and, although editing of cataloging and verification of headings cannot be controlled as carefully as before, *JALA* continues to include NUL in-process data and multiple-location information. Additionally, the capability for production and cumulation of specialized indexes or arrangements of entries through computer production is being utilized by G. K. Hall in its publication of annual cumulations of *JALA* beginning with the 1978 numbers.

Distribution of Tapes

Although the Library of Congress was unable to incorporate NAP data into its own MARC tapes, it offered NAP tapes for sale to the public through the Catalog Distribution Service beginning in April 1979. At least two bibliographic utilities indicated that they would purchase a tape subscription, and one of them, UTLAS, announced it would be mounting the tapes on its database in 1980.[7] Even if there had been no market for these records, the production of tapes would not have been simply an exercise. When the Library of Congress is able to incorporate non-LC records into its MARC tapes, it may be that mechanisms used by Northwestern for selecting and transmitting records to LC will provide a design for inclusion of noncen-

trally produced cataloging in a national bibliographic database.

Decentralized Creation of Authority Records

Throughout the term of the project, NAP staff searched and established names, subjects, and series in LC's Official Catalog. Although the rate of heading production was increasing as the first phase of the project neared conclusion, the actual number of contributed headings was not large. The low total for thirty months' operations of 68 new, 337 retrospectively converted, and 300 corrected name headings; 6 series treatment cards; and 2 topical subject headings is largely attributable to the complex operational procedures that were included in the project specifications. Heading verification and establishment procedures were as nearly identical to those followed by Library of Congress staff as a thousand-mile separation would allow, and this "equality" of approach contributed substantially to the ultimate costliness of the authorities segment of the project. LC catalogers, for example, are not permitted to interpret the presence of a heading on a bibliographic record as evidence that the heading has been established. Similarly, even if a heading has once been verified, the cataloger may not assume that it remains unchanged; the heading must be reverified before being used again. This system of safeguards may be appropriate when a cataloger can actually visit LC's Official Catalog, but when headings must be mailed to an agent for checking, returned to the cataloger for preparation of any required authority records, and mailed back to the agent for second verification and coding, the system is burdensome and far from cost-effective. So long as LC's Official Catalog continued in manual form, however, and so long as the name authorities database was not available to NUL online, LC saw no way in which its procedures could be shortened.

Cumbersome though the process was, the experience of working through an agent and adhering strictly to LC's internal requirements was valuable, so long as it led inescapably to the conclusion that cooperative creation of an authorities database is not financially feasible if all restrictions that apply to LC's own staff, especially including on-site

consultation of a central official catalog, are placed on non-LC authority contributors. Further, the low heading total in no way interfered with the demonstration that high quality, compatible authority work can be performed remote from the Library of Congress as long as all relevant rules and interpretations are available.

Cataloging of Africana Materials

Because of additional project responsibilities taken on by Northwestern's Africana cataloging staff and the difficulty of attracting and holding temporary project staff, it was correctly assumed that even the temporary availability of another professional cataloger would not add to the number of titles fully cataloged by Northwestern for the NAP tapes. Our expectations were more than realized when the project cataloger left after seventeen months. Although we were able to replace some production with a copy cataloger, the amount of original Africana cataloging produced by Northwestern staff suffered. This would not have been a problem as the project was originally envisioned, but as the focus on authority records increased, the limited amount of original cataloging generated was regarded as a regrettable, though unavoidable, condition.

Online Access to the LC Databases

The online access to the LC data files, which was part of the original project design, was intended to be only the first of many such connections to be set up. When congressional action placed a moratorium on the installation of other online connections to LC,[8] it made NUL's computer link virtually unique in the nation. Because of the stage of development of LC's online system and limitations for its use in authority creation, as well as because of the initial delay in setting up the connection, our ability to query LC's databases online had limited impact on the first eighteen months of the project. Acquisition and installation of the terminal itself (an OMRON) were complicated, and once in place the terminal was plagued by maintenance problems. The eventual development of component word searching,[9] however, plus the online availability of name authority data and the replacement of the sophisticated OMRON terminal with a

simpler dial-up model, greatly increased the utility of the LC link, and consultation of the LC database is playing an increasing role in NUL's own processing routines. Because of the current limitations on making online hookups available to other institutions, the uses that Northwestern makes of its LC connection will undoubtedly be of considerable interest to the Library of Congress and the library community.

Increased Knowledge and Adherence to LC Standards

Since a major objective of the project was to test whether cataloging of a quality suitable for inclusion in a national bibliographic database could be created in a decentralized manner, it was necessary for Northwestern to have access to LC's internal pronouncements on cataloging policy and practice. Accordingly, Northwestern received copies and continuing updates of several Library of Congress documents. These manuals (*Subject Cataloging Manual, Descriptive Cataloging Manual,* and *Rule Interpretations and Official Revisions*), combined with ready access to LC databases through the online computer link, proved a great benefit not only to project staff but to all cataloging personnel.

Of special significance was the availability of information on creation and interpretation of machine-readable authority data. This information assisted us substantially in formulating procedures and setting standards for NUL's subsequent project to convert its manual name, series, and subject authority files to the NOTIS database. The experience continues to be valuable in light of Northwestern's continuing contribution of authority records to the Library of Congress in connection with the Name Authorities Cooperative (NACO) project, a preliminary program for cooperative creation of a national authorities database.[10]

Submissions to the Register of Additional Locations (RAL) and National Union Catalog (NUC)

Because of the cooperative links with LC that were established for NAP, and because of Northwestern's machine capability, NUL was asked by the Library of Congress to participate in an experiment to see if LC could accept machine-readable submissions to *NUC*

and *RAL*. Two test tapes were submitted, one containing only additional location information, the other containing all types of records. Manual submissions continued to be made in a parallel system. It is our hope that LC will be able to derive the information it needs from these machine-readable contributions, so that in the future NUL and other libraries can report to *NUC* and *RAL* in this simpler, more efficient manner.

Program Modifications and Enhancements to the NOTIS System

The NAP database was created as a totally separate file from Northwestern's central catalog, and its holdings information differed somewhat from usual NOTIS structure. Therefore, the existence of Africana data on a separate file did not make the maintenance of Africana records on the main file unnecessary. Because of the requirement for keeping nearly identical records on two separate files, and in order to avoid duplicate keying for input, a mechanism was developed for copying data from one file to another. The procedure, which was perfected early in the project, was crucial to NAP and has been of considerable use in other NUL processing activities since its implementation.

In addition to development of file-to-file copying, NAP required several other additions or alterations to be made to the NOTIS system and its programs. Although the work was originally done for NAP, its benefits have extended beyond the project. The several programs for selection of records for transmittal or transfer are of continued use, and programs controlling the machine production of *JALA* are being used in connection with a variety of other ongoing programs.

Machine-to-Machine Transfer of Data

Soon after the project began, it became apparent that a very costly component of our activities was the keying of contributed cataloging data into the NAP file. Since most of the copy was already in machine-readable form in another database (OCLC, RLIN, or LC), investigations were made into the possibility of receiving contributed copy in machine-readable form either from an institution's archival tapes or through machine-to-machine transfer of data. Although represent-

atives of OCLC and the Research Libraries Group (RLG) expressed interest in development of machine-to-machine data transfer, the limited duration of NAP did not enable us to carry negotiations to a point where an experiment could be carried out. Northwestern continued its own investigations, however, and RLG's professed interest in developing this capability played an important part in NUL's decision to join that partnership.

Subsequent to the project's end, the Council on Library Resources funded a joint project of LC, RLG, and the Washington Library Network (WLN) to address the problems of sharing authority data among databases in an online mode as a requirement for the formation of a national authorities database. [11]

CONCLUSION

The Northwestern Africana Project was designed to test a method for creating a consistent national bibliographic database. Because of LC's inability to incorporate non-LC data into its own MARC tapes the test could not be completed. Despite the absence of the final step, the project still touched on and influenced several vital issues leading to such a database. The coordination, selection, and submission of bibliographic records to the Library of Congress; the test of machine-readable tape submissions to the *National Union Catalog* and the *Register of Additional Locations;* the discovery of the value of LC's internal manuals; and the usefulness of having an online link to the Library of Congress computer system are all of national interest.

Probably the most important part of the project, however, was the experimentation in methods for cooperative creation of authority records. The experience of NAP not only showed that high-quality, consistent authority work can be performed in a decentralized manner, but also confirmed that successful creation of a consistent national bibliographic

database depends heavily on the successful and efficient sharing of authority data.

On January 1, 1980, the Northwestern Africana Project took a dramatically different turn as remaining funds began to be used to support telecommunications costs and training for Northwestern's participation in LC's newly begun Name Authorities Cooperative Project. There is no way to know to what extent the Library of Congress' experience with Northwestern influenced the design of NACO, but it undoubtedly had some impact. We are pleased to note that the requirements for authority creation and verification that were most irksome in NAP have been modified or eliminated from the new venture. On-site consultation of LC's Official Catalog is no longer required, and participants may assume that a heading that appears on an LC bibliographic record has been established. Additionally, staff at participating libraries have full responsibility for preparation of authority records, while central LC staff perform only the same sort of review given to LC-based catalogers. Even if no program for decentralized authority creation had continued after the cessation of NAP, the lessons of the project would have been valuable to learn. It is extremely satisfying, however, to see the emergence of this program that is intended to be a prelude to eventual cooperative creation of a national authorities database, and to have been asked to be one of its first participants.

NAP had its share of frustrations, but difficulties encountered were all superable, and we can regard the completed phase of the project with satisfaction, feeling that we have contributed in several important ways not only to the field of Africana scholarship through the dissemination of bibliographic information, but also to the field of librarianship through our work to create a high-quality compatible database at a location removed from the Library of Congress.

REFERENCES

1. Lenore S. Maruyama, "COMARC Meeting Report," mimeographed (1979).
2. "Northwestern University," *College & Research Libraries News* 41:37 (Feb. 1980).
3. Melitta Cutright, "NU Library Completes Project for Venezuela" (memorandum, Northwestern Univ., Dec. 1979), p.7.

4. *Joint Acquisitions List of Africana* (Evanston, Ill.: Melville J. Herskovits Library of African Studies, Northwestern University Library, bimonthly).
5. Karen L. Horny, "NOTIS-3 (Northwestern On-line Total Integrated System): Technical Services Applications," *Library Resources &*

Technical Services 22:361–67 (Fall 1978).
6. "Africana Bibliographic Records Available on Tape," *Library of Congress Information Bulletin* 38, no.13:109 (March 30, 1979).
7. "Priorities for New Source Files," *UTLAS Newsletter* 4, no.12:1–2 (Dec. 1979).
8. U.S., Congress, House, "Legislative Branch Appropriation Bill, 1979," *Congressional Record*, V.124, 1979, p.H5503.
9. Patrick Frazier, "Alien in the Reading Room,"

American Libraries 11:538 (Oct. 1980).
10. "LC, Three University Libraries Agree on Name Authority Cooperation," *Library of Congress Information Bulletin* 39, no.7:53 (Feb. 15, 1980).
11. "An Integrated Consistent Authority File Service for Nationwide Use," *Library of Congress Information Bulletin* 39, no.28:244–48 (July 11, 1980).