

Commentaries & Case Studies

DEAN DeBOLT and JOEL WURL, *Editors*

The Commentaries and Case Studies department is a forum for sharply focused archival topics that may not require full-length articles. Commentaries and Case Studies articles generally take the form of analyses of archivists' experiences implementing archival principles and techniques within specific institutional settings, or short discussions of common theoretical, methodological, or professional issues. Members of the Society and others knowledgeable in areas of archival interest are encouraged to submit papers for consideration. Papers should be sent to Managing Editor, *The American Archivist*, Society of American Archivists, 600 S. Federal, Suite 504, Chicago, IL 60605.

The NHPRC Data Base Project: Building the "Interstate Highway System"

RICHARD A. NOBLE

The second edition of the National Historical Publications and Records Commission's *Directory of Archives and Manuscript Repositories in the United States* will be published in the spring of 1988.¹ The volume describes the manuscript, archival, and special media holdings of 4,200 institutions, a 50 percent increase over the first edition. The *Directory* is the only single-volume reference work covering the nation's archival and manuscript institutions. Yet it began as part of a much more am-

bitious project that has an interesting and significant history.

Chronological Overview

The *Directory* effort began in 1951 with a decision by the National Historical Publications Commission to compile a central register of manuscript collections. The purpose was to improve administration of the commission's documentary publications program. A staff headed by the commission's executive director, Philip M. Hamer,

¹The Oryx Press of Phoenix, Arizona, is publisher of the *Directory*'s second edition.

The author, archivist with the Machine-readable Branch of the National Archives and Records Administration, previously worked for the National Historical Publications and Records Commission as a grants analyst and an editor of the Directory of Archives and Manuscripts Repositories. This article is a revised version of a paper presented at the 1987 annual meeting of the Society of American Archivists.

gathered collection-level data in the 1950s. The result was the commission's *A Guide to Archives and Manuscripts in the United States*, which was published in 1961 and covered 1,300 repositories nationwide.² The Library of Congress's *National Union Catalog of Manuscript Collections* (NUCMC) also was first prepared during the late 1950s, and its first volume was published in 1962, a year after the *Hamer Guide*.³

After deciding in 1974 to revise the *Hamer Guide*, the commission quickly recognized that the amount of repositories and records holdings had grown dramatically in the thirteen years since publication of *Hamer's* volume, making a full update of *Hamer's* collection-level information prohibitively time consuming and unwieldy. The commission instead decided to prepare a repository-level directory and to compile the collection-level data over a longer time period. The *Directory* was intended to provide an overview of the nation's historical records repositories—a matter of increased interest to the commission when its records grant program was inaugurated in 1975.⁴

The commission chose to automate production of the *Directory* so that future editions could be produced more easily and selected the SPINDEX computer program package for this purpose. SPINDEX, which stands for Selective Permutation Indexing, was developed in the late 1960s primarily

by the Library of Congress and the National Archives, initially as a means of indexing finding aids. By 1975 it was able to generate descriptive guides with sophisticated indexes in an attractive typescript format. SPINDEX, however, had the disadvantage of being an off-line, batch mode system.⁵

Starting in 1976, the commission's *Directory* staff canvassed 10,000 repositories by mail and telephone. Published in 1978, the *Directory* described the holdings of 2,700 repositories—over twice as many as appeared in the *Hamer Guide*. But, unlike the *Hamer Guide*, the *Directory* provided only repository-level information. Reviews of the *Directory* from the archival and library community were almost uniformly positive.⁶

The *Directory* was conceived to be part of the commission's plan for a national collection-level data base on archives and manuscripts. The data base was designed to use the same SPINDEX programs, standard descriptive elements, and thesaurus of index terms as the *Directory*. Due to the flexibility of the SPINDEX programs, the data base could generate a variety of records guides, such as to specific states, regions, types of repositories, or types of records.⁷

The commission staff planned that the data base would be implemented in a piece-

²Frank G. Burke, "A Proposal for Revision of *A Guide to Archives and Manuscripts in the United States*," unpublished, 15 August 1975, 1-2; Philip M. Hamer, ed., *A Guide to Archives and Manuscripts in the United States* (New Haven: Yale University Press, 1961).

³Library of Congress, *National Union Catalog of Manuscript Collections 1959-61* (Ann Arbor: J. W. Edwards, 1962). Subsequent editions have been as follows: 1962 and *Index 1959-1962* (Hamden, Conn.: Shoe String Press, 1964), and 1963-1985 (Washington: Library of Congress, various dates).

⁴With the inauguration of the records program, the National Historical Publications Commission (NHPC) was redesignated the National Historical Publications and Records Commission (NHPRC). Larry J. Hackman, Nancy Sahli, and Dennis A. Burton, "The NHPRC and a Guide to Manuscript and Archival Materials in the United States," *American Archivist* 40 (April 1977): 201-03.

⁵*SPINDEX II: Report and Systems Documentation* (Washington: National Archives and Records Service, 1975), 1-7; Hackman, Sahli, and Burton, "The NHPRC and a Guide," 203-04.

⁶*Directory of Archives and Manuscript Repositories in the United States* (Washington: National Historical Publications and Records Commission, 1978). Examples of favorable reviews include: "'A' rating—important for even a small basic reference collection in this subject" (*Wilson Library Bulletin* 65 [April 1979]: 583, 587-88); "Most academic libraries will need at least one copy" (*Choice* 16 [July/August 1979]: 652); "The directory is a gold mine of information for scholars, researchers, and other clientele of academic and research libraries" (*Booklist* 75 [1 July 1980]: 1627).

⁷Hackman, Sahli, and Burton, "The NHPRC and a Guide," 204-05.

meal but methodical manner. In 1977 the staff described the data base concept as follows: "Like the interstate highway system, as each little bit is completed it can be used. Someday all of the pieces will be there, integrated into a single whole."⁸

To initiate the national data base, the commission from 1977 to 1982 provided over \$1 million in grant funding for several descriptive projects. These projects described historical records, using standard data elements, for inclusion in the national data base. There were three NHPRC-funded statewide survey projects—in Kentucky, New York, and Washington. These projects gathered collection-level and sometimes series-level data from all known repositories by means of site visits. The commission also funded a cooperative descriptive project by three midwestern state archives—Illinois, Minnesota, and Wisconsin.

In 1980 the commission issued a report on the feasibility of creating the national data base. The report projected that almost 20,000 repositories and over 700,000 collection descriptions would be included in the national data base. (It is interesting to note that, today, after only three years of data entry, the Research Libraries Information Network (RLIN) data base already includes over 100,000 archives and manuscript descriptions.) The NHPRC report projected that creation of the data base would cost \$15 million, over \$10 million of which would come from the commission. The report acknowledged that a large increase would be needed in the commission's established records grant budget of \$2 million a year.⁹

By 1982 several factors had led the commission to abandon the national data base. Of greatest importance were the dramatic cuts in the commission's grant funding and

staffing implemented at the beginning of the Reagan administration. Another factor was that the General Services Administration and National Archives stopped supporting the SPINDEX programs and, in particular, did not develop promised new programs to permit more efficient manipulation of data. Moreover, the batch mode requirements of SPINDEX made it less desirable than the on-line descriptive systems offered by the Research Libraries Information Network (RLIN) and On-line Computer Library Center (OCLC).

Because of these problems, the commission in 1982 decided to limit the data base's goals to producing the second edition of the *Directory*. Now that the second edition is a reality, the commission is considering contracting with a publisher to update the data periodically and to publish future editions.

Results and Impact

Though its original primary objectives were not attained, the NHPRC data base played a role in the development of the US-MARC Archival and Manuscripts Control (AMC) format. The format was developed by the Society of American Archivists' National Information Systems Task Force (NISTF), which was active from 1977 to 1983. NISTF was asked to investigate establishing one or more national automated information systems for description of archives and manuscripts.

In 1977 the NHPRC data base was the only national archival descriptive system that was automated and used standard data elements; however, as indicated above, the SPINDEX programs carried disadvantages. Another candidate for a national data base was NUCMC, which had the drawback of being a manual system. Finally, there were the two major on-line library bibliographic

⁸Ibid., 205.

⁹"Staff Report: NHPRC Data Base" (Unpublished, 1980), 14.

networks: RLIN and OCLC. In 1977, however, RLIN and OCLC were devoted almost exclusively to the description of books and periodicals using the MARC format for published materials.

Possibly because of the inadequacies of the several potential national information systems, NISTF soon gave up on selecting a particular one and decided instead to create standard data elements to permit information exchange among the existing systems. From 1982 to 1984, NISTF, its successor the SAA Committee on Archival Information Exchange, and the Library of Congress developed the standard data elements, named the USMARC AMC format and formally approved in 1985. The RLIN and OCLC data bases, using the USMARC AMC format, now have emerged as the dominant national data bases for archives and manuscripts.¹⁰

The NHPRC data base, through its use of standardized data elements, helped pave the way for the USMARC AMC format. In fact, a number of the NHPRC data elements are compatible with the AMC format, and an automated conversion can be done between the NHPRC and AMC formats. The Midwest guide project and the New York state project, for example, have successfully converted their records into RLIN.

Although the NHPRC national data base never was realized, it left a legacy of useful survey projects across the country. Table 1 is a summary of the projects, and the appendix is a bibliography of the guides published by the NHPRC data base.

The first NHPRC data base project to

receive funding was a statewide survey by the Washington State Historical Records Advisory Board, beginning in 1977. The project director noted that this effort was the country's first "since the WPA Historical Records Survey of the 1930s to survey historical records on a statewide basis simultaneously in a wide variety of records sectors."¹¹ Among other products, the Washington project produced a typeset collection-level guide to 250 repositories in the state. After NHPRC grant funding expired in 1983, the Washington project drastically reduced its personnel and maintained the data base at a minimal level by obtaining voluntary updates from the state's major repositories. The project now is considering conversion of the SPINDEX data to the USMARC AMC format for inclusion in RLIN, OCLC, or the Western Library Network (WLN), a third bibliographic network.¹²

A second survey project—by the Kentucky Department for Libraries and Archives—was funded by the commission between 1978 and 1983. Since expiration of commission funding, the project has received support from the U.S. Department of Education. The project finished surveying the state in 1984, having covered 6,000 collections. In 1986 the project published the *Guide to Kentucky Archival and Manuscript Repositories*, which provides repository-level information for the state's 285 archival institutions. The project also has been entering collection-level information into the data base and plans to publish a series of collection-level volumes. The project is considering switching from

¹⁰Richard H. Lytle, "A National Information System for Archives and Manuscript Collections," *American Archivist* 43 (Summer 1980): 423–26; David Bearman, "Toward National Information Systems for Archives and Manuscript Repositories," *American Archivist* 45 (Winter 1982): 53–56; Richard H. Lytle, "An Analysis of the Work of the National Information Systems Task Force," *American Archivist* 47 (Fall 1984): 357–65; Nancy A. Sahli, "Interpretation and Application of the AMC Format," *American Archivist* 49 (Winter 1986): 11–12.

¹¹John F. Burns, "Statewide Surveying: Some Lessons Learned," *American Archivist* 42 (July 1979): 295.

¹²Author's conversation on 15 June 1987 with David W. Hastings, Chief of Archives, Washington State Archives.

Table 1

NHPRC Data Base Projects: Summary Statistics					
Institution	Repositories Described	Collections & Record Groups	NHPRC	Funding* Other Federal**	State
Kentucky Department for Libraries & Archives	285	6,000	\$354	\$167	
Cornell University, New York Historical Resources Center	1,000	20,000	\$321	\$363	\$500
Washington State Historical Records Advisory Board	1,000	25,000 [†]	\$457	\$25	
State Historical Society of Wisconsin (Midwest Guide Project)	3	13,000 [†]	\$100		
NHPRC <i>Directory</i> (1978)	2,675	—	federal non-grant funds		
NHPRC <i>Directory</i> (1988)	4,600	—	federal non-grant funds		
TOTAL	6,888 ^{††}	64,000	\$1,232	\$555	\$500
<p>* In 100's of thousands. Figures are approximate.</p> <p>** Includes funding from the National Endowment for the Humanities and the Department of Education.</p> <p>† Includes series descriptions.</p> <p>†† Excludes 1978 NHPRC <i>Directory</i>.</p>					

SPINDEX to an in-house automated system, as well as transferring the SPINDEX data base to RLIN or OCLC.¹³

The third NHPRC survey project covered New York State and was administered by Cornell University. For each county, the project produced a computer printout describing historical records at the collection level. After NHPRC funding ran out in 1983, the project was able to continue the survey with funding from the National Endowment for the Humanities, matched by an appropriation from the New York State legislature. In 1986, with a grant from the U.S. Department of Education, the Research Libraries Group (the parent institution of RLIN) wrote a computer program to convert New York's 12,000 SPINDEX collection descriptions into RLIN. The New York project now has abandoned SPINDEX and enters all of its data directly into RLIN; a total of 20,000 New York descriptions are now in the system. The project has thereby ensured national access to the records of many small repositories that probably never would have participated in a national data base on their own. Cornell is negotiating for the New York State Archives to take over maintenance of the data base when the statewide survey is completed.¹⁴

A final NHPRC-funded project, the Midwest State Archives Guide Project, was administered by the State Historical Society of Wisconsin. The project endeavored to establish a cooperative data base among the state archives of Illinois, Minnesota, and Wisconsin and to publish a guide to the records of the three institutions. The project failed in the latter goal, although it

produced a series-level guide to a representative sample of records. More importantly, the project achieved close cooperation among the three participating state archives and thereby produced a consistent multi-institutional data base. Unfortunately, soon after NHPRC support ended in 1982, the Midwest guide project was discontinued and the common data base abandoned; however, the Midwest project, like the New York State project, was successful in transferring most of its data to the RLIN data base through an automatic conversion program.¹⁵

In summary, the NHPRC data base was unique in several ways. Its central product—the NHPRC *Directory*—provides summary information, in one volume, about several thousand repositories of all types nationwide. The NHPRC data base was also the first attempt to compile a national automated data base for archives and manuscripts. Under its aegis, the first comprehensive records surveys since the 1930s were undertaken in three states. The guides produced by these survey projects have greatly increased access to historical records in the three states. Additionally, the on-site visits to repositories by trained field surveyors allowed records custodians to gain better in-house control of their collections.¹⁶

On the negative side, the NHPRC national data base was heavily dependent upon federal grant money, which has proved to be an unstable funding base in the 1980s. The data base projects also depended upon the federally funded NHPRC data base staff to promote standardization of data elements

¹³Barbara A. Teague, "The Burden of Batching: Current Uses of SPINDEX at Kentucky's Public Records Division" (Paper presented at the Fiftieth Annual Meeting of the Society of American Archivists, August 1986).

¹⁴G. David Brumberg, "From Batch Mode to On-Line: The New York Historical Documents Inventory and the Growth in Archival Automation" (Paper presented at the Fiftieth Annual Meeting of the Society of American Archivists, August 1986).

¹⁵The Midwest Guide Project converted 12,000 SPINDEX collection/record group descriptions into RLIN. Marion Matters, "SPINDEX—The Mother of Invention" (Paper presented at the Fiftieth Annual Meeting of the Society of American Archivists, August 1986).

¹⁶"Final Report [to NHPRC], Kentucky Guide Project, Phase II, Grant #81-45" (Unpublished, 1986); Burns, "Statewide Surveying," 296.

and index terms. When that staff was eliminated in 1982, the projects no longer were able to maintain common descriptive standards.

As a result of their continued dependence upon federal funding, the three surviving data base projects—in Kentucky, New York, and Washington State—have a somewhat uncertain future. This is one reason the inclusion of the survey data in the current bibliographic networks is so important. At least the data will be available in the networks, even if the individual data base projects do not continue.

The future is particularly tenuous for the SPINDEX computer system, because of its batch mode requirements. In fact, few institutions are still using SPINDEX.¹⁷ But the NHPRC data base, and its use of SPINDEX, did help pave the way for development of the USMARC AMC format and its implementation in national bibliographic systems. In this respect, there is still validity to the “interstate highway system” metaphor used by the NHPRC staff. The RLIN, OCLC, and WLN data bases are gradually realizing the goal envisioned ten years ago by the NHPRC.

¹⁷Besides the Kentucky and Washington state archives, other institutions still using SPINDEX include the Portland (Oregon) City Archives and the South Carolina Department of Archives and History.

Appendix

NHPRC Data Base: Published Guides

National Historical Publications and Records Commission:

Directory of Archives and Manuscript Repositories in the United States (Washington, D.C.: NHPRC, 1978). \$25.

Directory of Archives and Manuscript Repositories in the United States (Phoenix, Ariz.: Oryx Press, 1988). \$55.

Kentucky State Survey Project:

The Guide to Kentucky Archival and Manuscript Repositories (Frankfort, Ky.: Kentucky Department for Libraries and Archives, Public Records Division, 1986). \$12.

In preparation: collection-level guides, arranged alphabetically by city, ca. 8 to 10 volumes, with a consolidated index.

New York State Survey Project:

Fifty-five county-wide records guides have been published. (New York has 62 counties.) Example: *Guide to Historical Resources in Allegany County, New York, Repositories* (Ithaca, N.Y.: Cornell University, New York Historical Resources Center, 1980). Prices range from \$3 to \$16; price list available.

Washington State Survey Project:

Historical Records of Washington State: Guide to Records in State Archives and its Regional Repositories (Ellensburg, Wash.: Washington State Historical Records Advisory Board [WASHRAB], 1981). \$25.

Historical Records of Washington State: Records and Papers Held at Repositories (Ellensburg, Wash.: WASHRAB, 1981). Out of print.

Historical Records of Washington State: Guide to Public Records Held by State and Local Government Agencies [computer output microform (COM)] (Ellensburg, Wash.: WASHRAB, 1984). \$10.

Historical Records of Washington State: Private Records and Papers not held in Archival Custody [COM] (Ellensburg, Wash.: WASHRAB, 1984). Distributed only to major repositories, which were asked to screen researchers wanting access to the materials. Out of print.

Genealogical Resources in Washington State (Ellensburg, Wash.: WASHRAB, 1983). \$10.

Midwest State Archives Guide Project:

“Prototype Guide” (1978). A “test” typeset guide including a sample of records descriptions from each of the participating Midwest state archives and a consolidated index. Out of print.