The Digitization of the "Michigan Pioneer and Historical Collections:" A Case Study

Michael E. Unsworth
Michigan State University, unsworth@msu.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Michael E. Unsworth, "The Digitization of the "Michigan Pioneer and Historical Collections:" A Case Study" (2009). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284314748>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

THE DIGITIZATION OF THE “MICHIGAN PIONEER AND HISTORICAL COLLECTIONS:” A CASE STUDY

Michael E. Unsworth, *Humanities Librarian, Michigan State University Libraries*

This presentation examines the digitization of an essential resource for the study of Michigan's early history. Despite extensive print- and microform-era cataloging performed on the serial and digitizing by four major projects, serious drawbacks in using this resource remain.

Why Should We Care

The “*Michigan Pioneer and Historical Collections*” (MPHC) is a critical source in researching the Wolverine State's early history. Riding a wave in historical consciousness brought on by the Centennial of the American Revolution, it first appeared in 1877 as the Pioneer Society of the State of Michigan's *Pioneer Collections*. Its initial contents were the “pioneer reminiscences, biographical sketches and memorials, and the proceedings of the local pioneer societies.” Over time, transcriptions of historical documents increasingly appeared.¹ Numerous scholars have relied on its abundant resources. Its importance became even more critical with the fiery loss of the State of Michigan Archives in 1951.²

Publishing History

Interestingly, this forty volume serial never displayed the title of “*Michigan Pioneer and Historical Collections*.” Instead, it underwent three major title changes (Table 1) and had three separate issuing bodies (Table 2) over its fifty-two year run. For the sake of simplicity, many authors have preferred to use *Michigan Pioneer and Historical Collections* in their citations, thus creating difficulties for those wanting to consult the actual source. Differing titles and names of issuing bodies sporadically appeared over the years (Tables 3 & 4).

The first two issuing bodies of the *Collections* had close relationships with the Michigan State Government which underwrote its printing.³ Due to its limited early distribution, the State funded a reprinting of Volumes 1-21 between 1900 and 1913. Confusingly, ten of the reprinted volumes were labeled “Second Edition.” In addition, there were numerous inconsistencies in the titles of the various volumes (Image 2). The only significant textual changes between the two editions were some cleanup of entries in each volume's index.⁴ The State's Michigan Historical Commission took over publishing from the Michigan Pioneer and Historical Society with Volume 39 in 1915.⁵ The financial shocks of the First World War and the Great Depression ultimately ended publication with Volume 40 in 1929.

Indexing

Another significant state involvement came in the form of the State Librarian supervising the indexing of each volume and ultimately a two-volume cumulated index for Volumes 1- 30 was published in 1904 & 1907.⁶ Entries were predominately for author and titles with some subject listings.⁷ A more expansive “List of Subjects, Authors, and Illustrations, Volumes 1 to 39 [1876-1915], Inclusive,” appeared in Volume 39 and ran from pages 459-582.

Over the years several outside bodies produced indexes, mainly for the benefit of genealogists. In 1939 the Michigan Works Projects Administration embarked on an effort to provide more complete subject access to the *MPHC*. A typescript "Classified Finding List" was completed in 1942 but it became a casualty of World War II. The Wayne University Press printed it in 1952.⁸

Bibliographic Control of the Set

A serial publication which has experienced three major title changes, three different issuing bodies, two separate editions, and is known by an informal title is going to present a challenge to bibliographic control. Regardless of what system of cataloging is used, the resulting record(s) will be long, obtuse, and mind-numbing to even the most motivated user.

Since the *MHPC* was a serial publication, it has been cataloged under both Latest Entry and Successive Entry formats. For pre-AACR2 records, the *National Union Catalog, Pre-1956 Imprints* has latest entry records for the initial forty volume edition⁹ (Image 3) and the twenty-one volume reprint¹⁰ (Image 2). For issuing bodies, these two records follow the then-contemporary cataloging practice of having only the latest form of a corporate body's name, instead of including all the pertinent names. Hence both records have entries for the Michigan State Historical Society, the successor to the historical societies that issued the *Collections* (Table 2). AACR2 online latest entry records in the OCLC database (Table 5) provide entries for the actual issuing bodies. The successive entry records (Table 1) also list the actual issuing bodies.

Regardless of what type of record that was consulted by the future digitizers, none provided sufficient information for Internet searching. This will be discussed in detail below.

Microform reproductions

Despite its importance in Michigan history, the *MPHC* attracted little attention from microform publishers. It was only in 1988 did the Ann Arbor-based University Microfilms International (UMI; now ProQuest) filmed it as part of its ambitious *Genealogy and Local History* microfiche series (series numbers LH5401 - LH5404). UMI inputted records for four titles into OCLC (Table 6). A careful examination of the microfiche in question (LH5401 & LH5402) reveal that the UMI cataloger erred and should have followed the example of the catalogers of the printed set in having records for three titles (Appendix 1). In 1994, the University of Michigan Library's Preservation Office produced a microfilm version (Table 7) and used successive entry cataloging. Thus, at the dawn of the digital era, this acid-rich resource was deteriorating and available only at Michigan libraries and out-of-state research collections.

The Seductive Appeal of Digitization

By the late 1990s, rapidly developing information technologies promised a breakthrough in using the *Collections*. Digitization could increase its availability, provide hi-tech replacements of fragile originals, easily produce an indivisible full run of a perennial favorite, and, most intriguingly, create straightforward online access to the contents of a serial that had defied comprehensive indexing. Michigan State University professor Justin Kestenbaum, a leading figure in the historiography of Michigan and an

enthusiastic adaptor of information technology, was exploring production of a CD-ROM version of the *MPHC* at the time of his death in 1995.¹¹

Enter the Digital Era

Some organizations selectively digitized parts of the *Collections* (Table 8). Four major international projects digitized more substantial portions. Google Books and the Internet Archive did mass digitizations; the University of Michigan and the Library of Congress efforts were thematically focused (Table 9).

The prime result of these projects is that all forty volumes are available on the Internet for the use of researchers. Thirty-nine volumes of the first edition and fourteen volumes of the reprint/second edition were converted. Since the reprints contained only “minor corrections,” their content is virtually identical to the first edition.¹²

Only one project, the University of Michigan, digitized all forty volumes. As detailed below, these projects did not meet reasonable expectations in accessing and utilizing the *MPHC*. On the contrary, it is hard to locate and utilize the resource on the Internet.

Obviously a Book?

First of all, none of the projects approached the resource as a serial. Each volume was scanned as a free-standing monograph, resulting in several negative implications. There was no quality control to insure that all volumes were digitized. In fact, participating libraries in two projects digitized the same volumes twice: Google (4) and Internet Archive (2) (Table 9). Of course, if Google Book Search had recognized that the *Collections* is a serial, then it wouldn't have included it;¹³ possibly Google Scholar should have handled its contents.¹⁴

As it turns out, thirty-nine of the “first edition's” forty volumes were digitized by at least one project; 14 of 21 for the reprint/second edition. No project digitized Volume 9 of the first edition; all four did convert Volume 9 of the second edition.¹⁵ As noted above, the absence of a digitized first edition for Volume 9 is not crucial since the text is essentially the same in second edition.

The Library of Congress digitized only nine volumes of the reprint/second edition.¹⁶ This was due to the guidelines of its *Pioneering the Upper Midwest* which selectively picked volumes from the collections of the Michigan, Minnesota, and Wisconsin historical societies. The reasoning was that these volumes would buttress the monographs that formed the bulk of the project.¹⁷

Searching Within the *MPHC*

A more critical problem of having forty free-standing volumes is that researchers must repeat the same search for each volume rather than being able to search the entire set. What one can do rapidly for a multi-titled JSTOR entry¹⁸ or in the *Making of America's* “War of the Rebellion” set¹⁹, one must tediously key-in forty separate times. While the compiled indexes mentioned above have also been digitized, a researcher must diligently navigate through clunky software. One can only speculate why this was done.

Searching for the *MPHC* on the Internet

Three titles changes, several variant titles, and three issuing bodies make locating this set difficult in any environment. Each digitizer complicated access by its decisions on titles and issuing bodies. None used "*Michigan Pioneer and Historical Collections*" as an access point.

Two digitizers simply relied on the latest entry cataloging records for the first and second editions for their access points. The University of Michigan had the "Michigan Historical Commission" for the author and "*Michigan Historical Collections*" as the title (Image 3) while the Library of Congress used "Michigan State Historical Society" and "*Historical Collections*" (Image 2). The implications of their decisions are obvious. Earlier issuing bodies and titles can not be used in searching the Internet.

The Internet Archive did a little better. It chose to list the Michigan Historical Commission and Michigan State Historical Society as issuing bodies with "*Pioneer Collections*" and "*Historical Collections*" for titles. Thus researchers searching the Internet for "Michigan Pioneer and Historical Society," "Pioneer Society of the State of Michigan," or "*Michigan Historical Collections*" would not find all the volumes the Internet Archive digitized.

Google Books apparently followed a "Let a Thousand Flowers Bloom" philosophy. Google has four issuing bodies and five titles. Unfortunately, it did not match up its choices with the information on actual title pages (Table 10).

Coping With a Mess

Thus, intellectual access via Internet searching to this important resource is still impaired by the projects' poor decisions on the nature of the *Collections* and choice of titles and issuing bodies. It is highly unlikely that another digitization project will tackle this resource. Thus, the library community has to do what it has always done: make sense out of an existing unfortunate situation.

The first step towards creating some order is to include an entry for "*Michigan Pioneer and Historical Collections*" to existing bibliographic records. After the author spent half an hour explaining the *MPHC* intricacies to a Michigan State University Libraries' serial cataloger, the OCLC records that have MSU holdings have been modified (Images 4 and 5).

Another contribution would be to have 856 links for the individual volumes in the *Collections*' bibliographic records. Both the University of Toronto²⁰ and, to a lesser extent, the University of Michigan²¹ have tried to do this with some degree of success (Tables 11 and 12) for their latest entry records. Regrettably, neither library listed all forty volumes though they are available. Thus, another library could input comprehensive records into OCLC, preferably for successive entry records.

A More Comprehensive Approach

The best way to address the above confusion is to create a web site that discusses the history, bibliographic organization and problems of the *Collections* and lists all the online versions. The author has done this with "Michigan Pioneer and Historical Collections: A

Research Guide.”²² As currently organized, the site gives information of the *MPHC*’s value to researchers, its publication history, indexes, and online volumes.

Preventing Such Mistakes in the Future

The author will then engage the audience in crafting strategies that could prevent such situations in the future and on coping with existing problems of this nature.

Summary

The story of the *Michigan Pioneer and Historical Collections* demonstrates the pitfalls of the “vacuum cleaner” approach of mass digitization projects. Apparently digitizers will rely on whatever cataloging information is available and then use the easiest way of churning it into “metadata.” Complexities and subtleties will be cheerfully ignored in pursuit of the Goddess of High Production. However, it will create opportunities for librarians to cast some light on the resulting messes. This assumes that such work will be valued and funded.

¹ Clifford L. Lord, “By Way of Background,” *Keepers of the past*, edited by Clifford L. Lord. (Chapel Hill, University of North Carolina Press [1965]), 4; George Newman Fuller, “A Sketch of Historical Societies in Michigan,” Michigan Historical Commission. *Bulletin* No. 3 (May 1914): 12-13.

² David J. Johnson, “Out of the Ashes,” *Michigan History* Vol. 85, no. 1 (2001): 22-27.

³ Michigan. *Public Acts* (1881), pp. 72-73 quoted in John Christian Larsen, “A Study in Service: the Historical Development of the Michigan State Library and [sic] Its Historical Predecessor, the Legislative Council Library, 1828-1941, Ph.D. Diss. University of Michigan, 1967, 150.

⁴ Fuller, “A Sketch of Historical Societies in Michigan,” 15, footnote 18.

⁵ Michigan. *Public Acts* (1913), p. 525 quoted in Larsen, “A Study in Service,” 273.

⁶ Marcy C. Spencer, “Preface,” in *Index to the Reports and Collections of the Michigan Pioneer and Historical Society. I-XV 1874-1890* (Lansing: Robert Smith Printing, 1904), [1].

⁷ *Classified finding list of the collections of the Michigan Pioneer and Historical Society*. (Detroit: Wayne University Press, 1952), [1].

⁸ *Classified finding list of the collections of the Michigan Pioneer and Historical Society*. (Detroit: Wayne University Press, 1952), [1].

⁹ “Michigan Historical Commission. *Michigan Historical Collections* (Library of Congress Card Number 0-3910 rev 2*) “*National union catalog, pre-1956 imprints* ([London], Mansell, 1968-1981), Vol. 382, p. 166, Column 3.

¹⁰ Michigan State Historical Society. *Historical collection. Collections and researches made by the Michigan pioneer society* (Library of Congress Card Number 20-15298 Revised) “*National union catalog, pre-1956 imprints* ([London], Mansell, 1968-1981), Vol. 382, p. 423, Column 1.

¹¹ David T. Bailey, e-mail message to author, Nov. 3, 2008.

¹² Henry S. Bartholomew, “Preface to the Second Edition, Volume IX,” in *Collections: Report of the Pioneer Society of the State of Michigan...Vol. IX, Second Edition* (Lansing: Wynkoop, Hallenbeck Crawford Co., 1908), vi.

¹³ Unfortunately, Google Books does not define what it considers a “book.” “About Google Books,” <http://books.google.com/intl/en/googlebooks/about.html> (accessed Oct. 26, 2009).

¹⁴ “About Google Scholar,” <http://scholar.google.com/intl/en/scholar/about.html> (accessed Oct. 26, 2009)

¹⁵ The title of the first edition was *Pioneer Collections*; simply *Collections* for the second edition.

¹⁶ [http://memory.loc.gov/cgi-](http://memory.loc.gov/cgi-bin/query/S?ammem/lhbumbib:@OR%28@field%28AUTHOR+@od1%28Michigan+state+historical+society+%29%29+@field%28OTHER+@od1%28Michigan+state+historical+society+%29%29%29)

[bin/query/S?ammem/lhbumbib:@OR%28@field%28AUTHOR+@od1%28Michigan+state+historical+society+%29%29+@field%28OTHER+@od1%28Michigan+state+historical+society+%29%29%29](http://memory.loc.gov/cgi-bin/query/S?ammem/lhbumbib:@OR%28@field%28AUTHOR+@od1%28Michigan+state+historical+society+%29%29+@field%28OTHER+@od1%28Michigan+state+historical+society+%29%29%29)

¹⁷ “Pioneering the Upper Midwest :About the Collection” (

<http://memory.loc.gov/ammem/umhtml/umabout.html>; accessed Oct. 3, 2009)

¹⁸ “Advanced Search,” (<http://www.jstor.org.proxy1.cl.msu.edu/action/showAdvancedSearch>; accessed Oct. 25, 2009)

¹⁹ <http://digital.library.cornell.edu/m/moawar/waro.html>

²⁰ “Michigan historical collections,”

(http://search2.library.utoronto.ca/UTL/index?N=0&Nr=p_catalog_code:683582&showDetail=first :

accessed Oct. 26, 2009). This record lacks vols. 12, 24, 29, 34, & 40. It lists Volume 15 twice. The second Volume 15 is actually the index to Vols. 1-15

(<http://www.archive.org/details/michiganindex1to15michuoft>; accessed Oct. 26, 2009).

²¹ “Michigan historical collections,” (<http://mirlyn.lib.umich.edu/Record/000534625/Holdings#0>; accessed Oct. 26, 2009). This record has a “Google Online (Available Online)” link which goes to Volume 17

(http://books.google.com/books?id=HRLiAAAAMAAJ&printsec=frontcover&source=gbv_ViewAPI#v=onepage&q=&f=false; accessed Oct. 26, 2009) and a “HathiTrust Digital Library listing that only displays vols. 1-2; 7-11; 14; 32 (listed twice both links go to <http://hdl.handle.net/202>), & 33). This partial listing is amazing since the U. of Michigan has digitized all forty volumes.

²² (<http://libguides.lib.msu.edu/MPHC>; accessed Oct. 26, 2009).