University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, UNL Libraries

Libraries at University of Nebraska-Lincoln

September 2004

An online guide to Walt Whitman's dispersed manuscripts

Katherine L. Walter

University of Nebraska-Lincoln, kwalter1@unl.edu

Kenneth M. Price

University of Nebraska - Lincoln, kprice2@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/libraryscience>

Part of the [Library and Information Science Commons](#)

Walter, Katherine L. and Price, Kenneth M., "An online guide to Walt Whitman's dispersed manuscripts" (2004). *Faculty Publications, UNL Libraries*. 135.

<https://digitalcommons.unl.edu/libraryscience/135>

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, UNL Libraries by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

An online guide to Walt Whitman's dispersed manuscripts

Katherine L. Walter
Kenneth M. Price

*University of Nebraska–Lincoln
Lincoln, Nebraska, USA*

Abstract

In November 2002, with funding from the Institute of Museum and Library Services, the University of Nebraska–Lincoln and the University of Virginia embarked on a project to create a unified finding aid to Walt Whitman manuscript collections held in many different institutions. By working collaboratively, the project team is developing a finding aid that is tailored to the needs of Whitman scholars while following a standard developed in the archival community, encoded archival description (EAD). XSLT stylesheets are used to harvest information from various repositories' finding aids and to create an integrated finding aid with links back to the original versions. Digital images of poetry manuscripts and descriptive information contribute to an ambitious thematic research collection. The authors describe the National Leadership Grant project, identify key technical issues being addressed, and discuss collaborative aspects of the project.

Keywords: Digital libraries, Archives

Submitted February 5, 2004; revised April 26, 2004; accepted June 12, 2004

Published in *Library Hi Tech* 22:3 (2004), pp. 277–282;
doi: 10.1108/07378830410560071
Copyright © 2004 Emerald Group Publishing Limited.
Used by permission.
<http://www.emeraldinsight.com/0737-8831.htm>

Introduction

Walt Whitman (1819–1892), a highly influential poet and one of the most innovative writers in the US history, is famous for his inclusive vision of democracy, for his celebration of ordinary people, and masterpiece, *Leaves of Grass*, which redefined American literature. Despite Whitman's centrality in American culture, his manuscripts have been little studied, and the poetry manuscripts, in particular, have never been collected and edited. Beginning in his teenage years, Whitman's manuscripts were scattered widely when documents were sent to friends and left with newspaper publishers. As a correspondent, he did not routinely keep copies of letters. Visitors to his house in Camden, New Jersey, often described Whitman dipping into the sea of paper that surrounded him, a seemingly endless source of manuscripts that were divided among three literary executors after his death. Many of the papers left with the literary executors were dispersed at auction and then further dispersed at subsequent sales. From an archival perspective, it is impossible to determine an original order for the entire corpus of Whitman manuscripts. The chaos of Whitman's papers was a point borne home to the project archivists by the fact that Whitman's manuscripts are now scattered in over 60 different institutional repositories, and poetry manuscripts have been located in 29 repositories. Because the materials are widely dispersed and irregularly documented, scholars or general readers interested in the development of Whitman's poetry—through multiple drafts to finished work—cannot locate and examine the relevant documents without great expense of time and money.

Whitman scholarship is complicated also by the fact that the poet only occasionally titled his manuscripts, and when he did, he often used a title different from that employed in any of the six distinct editions of *Leaves of Grass*. Furthermore, Whitman's drafts of ideas for his poems, his first treatment of key images, and his initial explorations of rhythmic utterances sometimes began as prose jottings that were gradually transformed into verse. For example, in the case of his great elegy for Lincoln, "When Lilacs Last in the Dooryard Bloom'd," Whitman jotted down bare lists of words that provided a kind of chromosomal code for the fully realized poem. Thus, for a number of reasons it is difficult to correctly identify and categorize Whitman's manuscripts.

The Walt Whitman Archive, an ambitious online scholarly project conceived by a team of scholars headed by Kenneth M. Price, University of Nebraska–Lincoln, and Ed Folsom, University of Iowa, began in 1995. It is a thematic research collection (Palmer, 2004) that sets out to make Whitman's vast work electronically accessible to scholars, students, and general readers. The site located at <http://whitmanarchive.org> is maintained on a server at the University of Virginia's Institute for Advanced Technology in the Humanities (IATH). The goal of the *Whitman Archive* is to create a dynamic site for research and teaching that will grow and change over the years. Editorial work on the poetry manuscripts is supported by the National Endowment for the Humanities. In order to advance the editing project and to increase the understanding of encoded archival description (EAD), a complementary project was undertaken by the University of Nebraska–Lincoln and the University of Virginia. This project, funded by the Institute of Museum and Library Services (IMLS) from 2002 to 2004, is entitled "An Integrated Finding Guide to Walt Whitman's Poetry Manuscripts." The purpose of the IMLS-funded project is to make an inventory of Walt Whitman manuscripts in various repositories and to provide access to the manuscripts through the *Walt Whitman Archive*. Aided by such standard references as *Walt Whitman: a Descriptive Bibliography* by Joel Myerson, *American Literary Manuscripts*, and *Archival Resources*, the *Whitman Archive* team has identified an estimated 70,000 manuscript items produced by or relating to Walt Whitman. Several thousand of these manuscripts are poetry manuscripts. The scholars and archivists working on the project fully expect that other manuscripts will appear as private collections pass into institutional hands or are offered on the auction market. Though all of the manuscripts located will be included in the finding aids of individual repositories online, we are particularly focusing on enhancing the descriptions of poetry manuscripts. One of the goals of the IMLS project is to increase the public's understanding of Whitman as a foundational figure in American culture. The enhanced finding aids and the accompanying digital images developed as part of this project help the public gain new insight into the development of Whitman's poetry, providing a wide audience new understanding of the creative process that brought about some of the most moving and memorable poems ever written in the US. Readers who otherwise have little access to manuscript reading rooms are able to see that Whitman, who often praised spontaneity, was himself an incessant reviser: his works did not magically appear fully formed but instead reached their often

majestic state through complex processes of trial and error and painstaking reiterations and revisions. Whitman scholars are adding to the archival descriptions of poetry manuscripts to help readers understand those processes by situating the archival material in its wider intellectual context. As part of the freely accessible *Walt Whitman Archive*, the images and finding aids provide teaching and scholarly opportunities not otherwise available.

The University of Nebraska–Lincoln's project objectives are to:

- create a unified guide to Whitman's manuscripts, with special enhancement of the descriptions of poetry manuscripts;
- produce several thousand digital images from poetry drafts found in various collections;
- establish best practices for using EAD to identify manuscripts for dispersed, cross-repository collections;
- develop a model for scholar-archivist collaboration; and
- build a search interface.

Collaboration

From the beginning, the project was conceived as one in which the scholars, archivists, and librarians would work together collaboratively. Each community brings special skills to the development of the online integrated guide. There are really two teams of individuals who have been involved. The first team is the overall research group, including consultants on the project.

- *Scholars*. Kenneth M. Price, UNL; Ed Folsom, University of Iowa; John Unsworth, University of Virginia's IATH; Brett Barney, UNL; Andrew Jewell, UNL; and other graduate students in English departments at all three institutions.
- *Librarians*. Katherine Walter and Brian Pytlik Zillig, UNL; Terry Catapano, Columbia University; and Heike Kordish, New York Public Library.
- *Archivists*. Mary Ellen Ducey, UNL; Daniel Pitti, IATH; Kris Kiesling, University of Texas at Austin; Steve Hensen, Duke University; and Anne Van Camp, Research Libraries Group.

The second team—the UNL EAD project team—is a subset of the larger group. This smaller group is composed of the UNL and IATH faculty and staff noted above.

One of the first meetings of the overall research group (University of Nebraska–Lincoln, University of Virginia, and collaborators) was held in Lincoln, Nebraska and

facilitated by IATH's Daniel Pitti. As the group discussed the issue of enhancing descriptions, most of the archivists on the team noted that special funding is typically required to provide item-level descriptions or calendar-level information, and that such information is not usually needed by most users. The scholars and the archivists recognized, however, that a digital thematic research collection that centers on a national icon like Whitman may require item-level descriptions, whereas other collections may not merit the staff time. In such instances, the collaboration between scholars and archivists can be very valuable.

The overall research group participates in an archived e-mail discussion list, and various individuals have had both face-to-face and phone meetings. Daniel Pitti and others from collaborating institutions have responded to encoding models proposed by members of the EAD project team at UNL.

The EAD project team, including Ducey, Jewell, Barney, and Pytlik Zillig, worked to coordinate EAD implementation across collections in a way that ensures the interoperability of records produced by different institutions and participants. All members of the Nebraska team meet weekly to discuss issues concerning the unified finding aid. Frequent communication among the librarians, archivists, and scholars has enriched the project and kept it on track.

Technical issues

Descriptive information

Descriptive information about Walt Whitman's manuscripts in many repositories was received in various forms, such as:

- photocopies of older catalog cards for single items, apparently never digitized;
- hand-typed finding aids;
- finding aids produced with word processing software, but not encoded;
- finding aids produced in HTML; and
- fully-encoded EAD finding aids (XML).

Not surprisingly, some finding aids, containing only a limited number of original Whitman manuscripts, describe the papers of Whitman associates or Whitman collectors. An early decision was to limit, at least for now, the descriptions created for the *Walt Whitman Archive* to those manuscripts actually written by Whitman himself. Someday it may be possible to add other archival materials, including documents by associates and collectors.

One of the most important decisions was to develop repository-specific finding aids, with the intent to harvest

the descriptions of poetry manuscripts into a single unified finding aid later in the project. As mentioned earlier, many of the individual repositories' finding aids or catalog cards were not in digital form, and the ones that were in digital form were not necessarily encoded following current archival standards.

In order to harvest data, the finding aids had to be encoded. EAD is a standard for encoding archival finding aids using SGML or XML, and it is this standard upon which the primary work of the Whitman project is based. As a standard for archival description, EAD is designed to encode finding aids in such a way that the contents of various collections can be searched uniformly online. Consequently, Nebraska developed a project-specific model for institutions not contributing their own encoding. Mary Ellen Ducey, UNL archivist, and Andrew Jewell, a graduate student in English at UNL, developed draft EAD documents for institutions without electronic finding aids and received recommendations for changes from Daniel Pitti, IATH, and Kris Kiesling, University of Texas at Austin. This model was shared among the major participating institutions and accepted with slight revisions by the participating archivists. By establishing a project-specific model, the team ensured that the ability to harvest specific fields from the repository finding aids using XSLT would be facilitated.

As part of the project, we request digital images of poetry manuscripts from the holding repositories to allow scholars to enhance the finding aids. Repositories are asked to provide 24-bit color TIFF images with a minimum resolution of 600 dpi, presented in context. Thus, when a poem is written on the back of a letter or an envelope, or when it is one of a group of related pages, images of the contextual materials are also obtained. We also seek permission to post derivative JPEG and thumbnail images on the *Walt Whitman Archive*.

With the image in hand, the individual repository's finding aid is reviewed to determine if there is additional information that would be useful to the scholarly community. For example, in the University of Tulsa's collection of Walt Whitman Ephemera, there is a manuscript called "[Poem describing a perfect school.]" This leaf has writing on both sides, as noted by the archivist at the University of Tulsa: "Written in pencil on 8vo sheet with portion of another poem, also in his hand, on verso." We have identified the poetic lines written on the verso as part of an extremely important Whitman poem, "To Think of Time," and have enriched the description as follows: "The verso lines, beginning 'The three or four poets are well,' were included, in a revised form, in Whitman's poem 'To Think of Time,' first published without a title in the 1855 edition of *Leaves*

of *Grass*, as 'Burial Poem' in 1856, 'Burial' in 1860 and 1867, and under its final title in 1871."

Aside from the fact that many repositories did not have finding aids per se to provide, a further complication was that the levels of description received from various repositories ran from extremely sketchy to very detailed. The Nebraska project team produced finding aids as best it could, based on the existing description and, typically, digital surrogates of the items, though some repositories continue to be slow in providing scans. The following description taken from the paper finding aid of the Livezey-Walt Whitman Collection at the Bancroft Library of the University of California–Berkeley, demonstrates how little the project sometimes had to work with:

"Wood Odors" (poem) Holograph Ms.

Working from an image of this manuscript and using other supplemental information from reference works on Whitman, scholars on the project were able to elaborate this description considerably. The description as it now appears in the *Whitman Archive* follows:

Item: 1

Title: "Wood Odors"

Date: ca. 1875

Physical Description: 1 leaf, handwritten

A draft of a poem unpublished in Whitman's lifetime entitled "Wood Odors." The poem was apparently written as Whitman was making notes for his 1882-1883 book, *Specimen Days*. Specifically, the poem appears to respond to the visit he made to the Stafford farm in New Jersey in the mid-1870s. Some have argued that this draft is not a poem at all, but a list of phrases toward the composition of *Specimen Days* (see David Goodale, "Wood Odors," *Walt Whitman Review* 8, [March 1962], 17). "Wood Odors" was published first in *Harper's Magazine*, 221 (December, 1960), 43.

Integrated guide to dispersed Walt Whitman manuscripts

The importance of uniform approaches to EAD became very evident when working on the integrated guide. In March 2003, IATH convened a group of scholars, archivists, and librarians in Lincoln, Nebraska. Scholars Price and Folsom described the value of developing a resource where one would be able to find information on all of Whitman's manuscripts, and, especially, on the various manuscript drafts and notebook versions of each of the more than 300 poems Whitman published in *Leaves of Grass* and of the additional poems (approximately 125)

that he did not include in his masterpiece. Led by Daniel Pitti of IATH, one of the architects of EAD, representatives from the University of Virginia, New York Public Library, Columbia University, the University of Texas at Austin, Duke University, the University of Iowa, and the Research Libraries Group discussed the encoding needed to develop a unified finding aid to dispersed manuscripts and articulated the desired outcomes of the project.

As described above, the wide dispersion of Whitman's manuscripts throughout his lifetime and after his death makes it impossible to determine an original order. In an article entitled "Disrespecting Original Order," Frank Boles notes that the concept of original order is less relevant for collections of personal papers than for governmental or institutional records. He argues that "original order is to be respected when it is usable, but ... a theory of simple usability can guide archivists when original order becomes inadequate" (Boles, 1982, p. 32).

In the case of Whitman's manuscripts, the research group concluded that scholars would be best served by creating a single, integrated guide to Whitman's poetry manuscripts. As envisioned, the *Walt Whitman Archive* would display the images of the poetry manuscripts (work by the author) with enhanced descriptions (work by librarians or archivists and scholars), and a citation and link to the individual holding institution's finding aid (work by archivists). Thus, an image of an 1881 corrected proof of Whitman's "The Dalliance of the Eagles" would be accompanied by the citation "Library of Congress, Charles A. Feinberg Collection," by additional scholarly descriptions or notes concerning the poem, and by a link to the Library of Congress's finding aid on the *Walt Whitman Archive*. Then, using XSLT stylesheets, poetry manuscripts for each of the poems would be united into a single alphabetical list regardless of location. In effect, the online unified guide creates a virtual order for Whitman's poetry manuscripts.

Though in concept and design the unified guide was simple, its development proved to be more complicated. Planning how to unite the finding aids for poetry manuscripts described in over 30 finding aids at 29 repositories (some repositories have more than one Whitman collection) has offered some interesting challenges. As the project progressed, the UNL EAD team had to address how to assign uniform titles to various drafts. This process is described in the "Work identification..." section of the article that follows. Once this issue was resolved, the team was able to develop a series of stylesheets to create an integrated guide. The steps are shown in Figure 1, entitled "Integrated Guide to Walt Whitman's Poetry Manuscripts: the XSLT transformations."

Figure 1 Sequence of XSLT transformations

Essentially, item-level information is drawn from many different levels of the constituent finding aids (from <c01> through <c05>) and redeployed in a “flat” file structure, so that all item-level information about poetry manuscripts is expressed at the same level (<c01>). For a more detailed illustration of the stylesheet, and to see how the component EAD files are gathered see Figure 2.

Next, a second stylesheet (Figure 3) organizes related manuscript items as <c02>s within <c01>s.

A third stylesheet transforms the EAD Integrated Guide to HTML for display in the browser. In this way, we are able to display and group all drafts of a particular poem together.

To see the Integrated Guide to Walt Whitman’s Poetry Manuscripts, go to <http://whitmanarchive.org> and click on “Manuscripts.”

Work identification or uniform titles

The first attempt to generate the integrated union finding aid was exciting (i.e. the stylesheets flattened the files as desired), but grouping the poetry drafts was impossible without some means of identifying like drafts. As noted earlier, Whitman was amazingly prolific, and he did not consistently name his poetry drafts in ways that logically or meaningfully grouped the drafts. To address this

Figure 2 First stylesheet gathers all component EAD files and creates a flat <c01>

```

<xsl:for-each select="document('.../docs/filename.xml')/did/parent::*[@level='item']">
  <xsl:element name="c01">
    <xsl:attribute name="level" value="item"/><xsl:attribute>
    <xsl:copy-of select="."/></did/>
    <xsl:copy-of select="./scopecontent"/>
    <xsl:element name="dao">
      <xsl:attribute name="href" value="filename.xml"/></xsl:attribute>
    </xsl:element>
  </xsl:element>
</xsl:for-each>

```

problem (and for other practical reasons), Price, Folsom, Pitti, and Brett Barney, an encoding specialist on the project, developed a system of identification which embeds in the union finding aid relationships between manuscripts of Whitman poems and the conceptual “work” they

Figure 3 Second stylesheet organizes related manuscript items as <c02>s within <c01>s

```

<xsl:template name="c01">
  <xsl:for-each select="//unitid[@type='work'][not(=following::unitid)]">
    <xsl:variable name="workfile">../.../xmlworkfiles/<xsl:value-of
      select="."/>.xml</xsl:variable>
    <xsl:variable name="currentNode">
      <xsl:value-of select="."/>
    </xsl:variable>
    <xsl:choose>
      <xsl:when test="contains('##')">
        <!-- hide any temp workIDs -->
      </xsl:when>
      <xsl:otherwise>
        <c01 level="class">
          <did>
            <unititle type="uniform">
              <xsl:element name="title">
                <xsl:value-of select="document($workfile)/titleStmnt/title[@type='uniform']"/>
              </xsl:element>
            </unititle>
            <unitid type="work">
              <xsl:value-of select="."/>
            </unitid>
          </did>
          <!-- the following generates the correct c02s within the appropriate c01s -->
          <xsl:for-each select="ancestor::ead/unitid[@type='work']">
            <xsl:choose>
              <xsl:when test="contains('##')">
                <!-- hide any temp entries -->
              </xsl:when>
              <xsl:when test="contains(',$currentNode')">
                <c02 level="item">
                  <did>
                    <xsl:copy-of select="ancestor-or-self::c01/descendant::unitid[@type='WWA']"/>
                    <xsl:copy-of select="ancestor-or-self::c01/descendant::unititle"/>
                    <xsl:copy-of select="ancestor-or-self::c01/descendant::unitdate"/>
                    <xsl:copy-of select="ancestor-or-self::c01/descendant::physdesc"/>
                  </did>
                    <xsl:copy-of select="ancestor-or-self::c01/descendant::scopecontent"/>
                    <xsl:copy-of select="ancestor-or-self::c01/descendant::dao"/>
                  </c02>
                </xsl:when>
                <xsl:otherwise>
                  <!-- hide all nodes not equal to value of the current node -->
                </xsl:otherwise>
              </xsl:choose>
            </xsl:for-each>
          </c01>
        </xsl:otherwise>
      </xsl:choose>
    </xsl:for-each>
  </xsl:template>

```

contribute to. The title of the “work” is derived from the final manifestation of the poem, most often the version Whitman published in his final, or “deathbed,” edition of *Leaves of Grass* (1892). For example, manuscript drafts of poetic lines that were later incorporated into “Song of Myself” are flagged with the ID for that poem. This innovative encoding of an individual manuscript’s relationship to a Whitman “work” allows us to enable a highly-valuable, automated organization of the integrated finding guide: we can group all dispersed manuscripts that relate to a specific poem. The use of “work” identifiers, for example, will allow a user of the integrated guide to use a well-known title, such as “When Lilacs Last in the Dooryard Bloom’d,” to perform searches and retrieve several manuscripts – from notes to lists of words to trial lines to corrected proofs – that may be held at different repositories.

Conclusion

On the access side, the project is providing researchers an opportunity to experiment with methods for virtually reintegrating dispersed collections of Whitman manuscript materials using the standard for archival description, EAD; on the social and intellectual side, the project offers an unusual opportunity to experiment with a deeper engagement between scholars and archivists, in which scholars might enrich the item-level descriptions of archival materials. Through the coordinated effort of a large number of libraries and institutions, the project has demonstrated how best to utilize and integrate EAD records made and maintained at disparate institutions by different creators.

By applying EAD consistently and by taking advantage of XML and XSLT, the project team has been able to develop a virtual collection of Whitman poetry manuscripts. Though archivists and librarians may not typically construct virtual collections (Westbrook’s, 2002), there may be times when manuscripts of a particular individual are so scattered that a virtual collection may be the only way in which to make sense of chaos. The collaborative team for the Whitman’s project has found that the different perspectives brought from each of the communities (scholars, librarians, and archivists) has enriched discussions and offered a creative approach to uniting split collections.

Whether a particular collection in an archive has a finding aid with a single summary description or a more detailed finding aid is determined by the priorities of individual repositories. While scholarly and general user interests are a factor in setting such priorities, many factors may be considered in making decisions concerning collection processing and retrospective conversion of finding aids. Rarely does an archive have the luxury of creating item-level descriptions for manuscript or other collections.

Occasionally, however, there are collections of such importance that item-level descriptions are essential. We believe that Whitman’s role as a foundational figure in the US merits this approach. Few of America’s great writers continue to generate as much interest in the wider culture as the poet of *Leaves of Grass*. Over a century after his death, Whitman is a vital presence in cultural memory: television shows and films depict him, musicians allude to him, advertisers appropriate him, schools and bridges are named after him, and politicians invoke him. Truck stops, think tanks, summer camps, corporate centers, and shopping malls bear his name. He is part of the very fabric of American life, its past, present, and no doubt future as well.

Electronic access to manuscripts offers opportunities for studying his scattered works that were unimaginable in the past. When our project began, the *Walt Whitman Archive* was receiving an average of 3,000 visits per day. As of January 2004, the site averages 10,000 visits per day. Based on the experience of the *Walt Whitman Archive* and the usage of the site, we believe that the general framework for developing the unified finding guide, with enhanced descriptive work by scholars, is worth replicating in other research communities to enrich resources pertaining to other highly significant writers or topics.

References

- Boles, F. (1982), “Disrespecting original order,” *The American Archivist*, Vol. 45 No. 1, pp. 26–32.
- Palmer, C. L. (2004), “Thematic research collections,” in Schreibman, S., Siemens, R., Unsworth, J. (Eds.), *A Companion to Digital Humanities*, Blackwell, Malden, MA, .
- Westbrook, B. D. (2002), “Prospecting virtual collections,” *Journal of Archival Organization*, Vol. 1 No. 1, pp. 73–80.