

*This is an Author's Original Manuscript of an article whose final and definitive form, the Version of Record, was published in American Journalism 31, no. 1 (Winter 2014): 4–25 [copyright Taylor & Francis], available online at: <http://www.tandfonline.com/> [10.1080/08821127.2014.875349].*

**Title:** Translation, Technology, and the Digital Archive: Preserving a Historic Japanese-language Newspaper

**By:** Kristin L. Gustafson, University of Washington-Bothell

**Keywords:** journalism history, technology, translation, marginalization, Japanese, newspaper, historiography, World War II, digital, archive, *Hokubei Hochi*, *North American Post*

## Introduction

The Nikkei Newspaper Digital Archive Project (NNDAP) began in July 2011. A joint project of the Hokubei Hochi Foundation and University of Washington Libraries, the NNDAP pursues a mission of interest and importance to journalism historians: the creation of a public digital archive for a Seattle-based, Japanese-language newspaper that published almost continuously across the twentieth century, from 1902 to the present moment, with a forced interruption during the World War II period of Japanese-American internment.<sup>1</sup> The *North American Post* or *Hokubei Hochi*—the two titles used for the newspaper after World War II (it was previously titled *Hokubei JiJi*)—is a rich resource for scholars, especially those interested in the history of the American West, the Japanese-American experience, Seattle, and, of course, ethnic newspapers in the US.<sup>2</sup> I joined the project in late 2012, intrigued to learn more about this historic newspaper with hyper-local, regional, national, and international value.

---

<sup>1</sup> The newspaper explains on its website that it uses the term Nikkei to refer to people of Japanese heritage.

<sup>2</sup> The newspaper will be referred to as the *North American Post* from here forward instead of referred to by its English and Japanese names. The University of Washington Libraries also lists the newspaper with the Kanji characters: 北米報知.

During the two years since its inception, the NNDAP has operated as a pilot project with the goal of creating an archive of the 1946–1947 newspaper issues (about 130 pages). The resulting digital archive will be used to recruit potential funders interested in preserving the entire run of this historic newspaper. Digitizing involves creating digital versions of the pages, coding, indexing, creating a metadata library with categories (such as people, buildings, events), and composing brief summaries of each page's contents in two languages. The NNDAP provides an opportunity for inter-disciplinary, inter-agency, and inter-campus involvement.<sup>3</sup> Board members work alongside undergraduate students in translating the Japanese-language text, and librarians work across institutional differences to create and meet rigorous archival standards.<sup>4</sup> The project builds bridges across academic interests, from computational linguistics to Japanese Studies to journalism history.<sup>5</sup>

Many national and state organizations are involved in creating digital archives of historic newspapers in the US. As most readers of this journal know, the National Digital Newspaper Program, a collaboration between the National Endowment for the Humanities, the Library of Congress, and state-level projects, has digitized more than 6.6 million pages of America's

---

<sup>3</sup> I extend my thank you to the two primary people organizing the Nikkei Newspaper Digital Archive Project for their roles in explaining the project: Elaine Ikoma Ko, the foundation's executive director, and Glenda Pearson, head of the library's Microform and Newspaper Collections, for their efforts in describing the intricacies of the project and processes. Ko started the Hokubei Hochi Foundation and serves as its executive director. Pearson is head of one of largest microform and newspaper library collections in North America.

<sup>4</sup> Pearson and other experts in the library—including Anne Graham, a senior computer specialist with the library's Digital Initiatives Program; Theodore Gerontakos, a metadata/cataloging librarian; Melanie Bolla, a graduate student in computational linguistics; and consultants in the East Asian Library—work together to coordinate newspaper scans, conduct metadata analysis, and more. The publisher of the *North American Post* and *Soy Source*, Tomio Moriguchi, also serves as the foundation's board president. Yohji Kameoka, another board member, helps with translation. Maiya Gessling is a NNDAP staff member and Hokubei Hochi Foundation student intern.

<sup>5</sup> KJ Hiramoto and Rena Kawasaki were the first undergraduate researchers to join the project from the University of Washington Bothell's School of Interdisciplinary Arts and Sciences.

newspapers published between 1836 and 1922 in 36 states and territories as of October 2013.

The resulting archive is fully searchable and made public via the *Chronicling America* website.<sup>6</sup>

Most of the newspapers being digitized are English-language papers, although the Library of

Congress has begun to include non-English newspapers. In addition, the Washington State

Library has a longstanding interest in digitizing newspapers (for example its online collection of indexed newspapers covering the state's territorial and early statehood).<sup>7</sup> It operates a state-level

project that is one of many across the country operating on behalf of the National Digital

Newspaper Program. Although the NNDAP is not part of the Washington State project, the state

project's coordinator, Shawn Schollmeyer, provides valuable consultation for the NNDAP.<sup>8</sup> In

doing so, she brings her knowledge of the high standards and expertise needed to preserve and

digitize historical newspapers.

The NNDAP appears to be the first time a historic Japanese-language newspaper has been translated and digitized in the US. The point is to use digital technology to create a public and easily accessible archive of a newspaper that served a large Japanese-American community for more than a century—and to make that archive useful to both English and Japanese readers.

---

<sup>6</sup> The Library of Congress, "Chronicling America: Historic American Newspapers," *Chronicling America*. Accessed 23 October 2013. <http://chroniclingamerica.loc.gov/>. These numbers get updated continually on the website. The program's "historically significant newspapers" include those published in English, French, German, Italian, and Spanish.

<sup>7</sup> Office of the Secretary of State, "Newspapers at the Washington State Library," *Washington State Library*, Accessed 23 October 2013. [http://www.sos.wa.gov/library/newspapers\\_wsl.aspx](http://www.sos.wa.gov/library/newspapers_wsl.aspx). The Washington State Library began a pilot project to create a digital collection of the Pioneer Newspaper collections in 2005. Laura Robinson, "The Evolution of Newspaper Digitization at the Washington State Library," *Microform & Imaging Review* 39, no 1 (March 2010): 24–27. This effort preceded the library involvement with the National Digital Newspaper Program beginning in 2008.

<sup>8</sup> Ingrid Pharris, "Chronicling Washington: WSL Receives Additional Funding for Newspaper Digitization Project," *Our Corner: Washington Secretary of State Blogs*, August 17, 2012. Accessed 26 January 2013.

<http://blogs.sos.wa.gov/FromOurCorner/index.php/2012/08/chronicling-washington-wsl-receives-additional-funding-for-newspaper-digitization-project/>.

Of course, digitizing the *North American Post* is not an easy task. It is not simply a matter of scanning a page and making that page available online. Rather, the project requires creativity and innovation in explaining the archived content through metadata. And it requires a keen understanding of written Japanese, which combines characters and syllabaries.<sup>9</sup> Those translating the *North American Post* must be fluent in this complex system of written language.

The NNDAP pilot project makes the voices of Japanese Americans returning to a city after being held in internment camps during World War II accessible beyond print and microform archives. And the project creates an opportunity to learn more about the challenges of digitizing America's journalism history. In this essay, I will describe three critical issues that emerged during my experience with the project in 2013: (1) the design decisions that shape digital newspaper archives and influence how they can be used in the future; (2) the role preservation methods play in future historians' use of the archives and interpretation of material in the archives; and (3) the selection of newspapers to include in the archives. First, decisions made by non-scholars influence how users access the digital archives in later years. The meta-data reflect practical decisions about the information that is included and excluded, and the summarizing nature of coding and indexing reflects the archivists' decisions about which audiences might use the archives in the future. Second, preservation methods influence future interpretation and thus require careful thought. All archives contain narratives. Making the Japanese-language content from the 1940s accessible to present-day Japanese and English speakers requires a certain amount of interpretation that leaves a contemporary and narrative

---

<sup>9</sup> Kanji are Chinese characters, and Hiragana and Katakana are syllabaries, or sets of written symbols for syllables. Hiragana are used to help readers understand the Kanji and to write words for which there is no Kanji equivalent. Katakana are used primarily to write foreign words and to create emphasis. I thank Rena Kawasaki for helping me understand these distinctions.

fingerprint on the archives. And third, the selection of newspapers we preserve requires resources of labor and money. The archival selection is clearly influenced by which newspapers have survived into the present. In the case of Japanese-language newspapers in the US, some newspapers did not survive the Japanese-American community's displacement through internment. United States Copyright Law also affects which newspapers are selected for preservation. In this essay, I will explore some of the barriers and opportunities to projects such as the Nikkei Newspaper Digital Archive Project with the hope of encouraging and information similar newspaper digitization projects in the future.

### ***North American Post* and the Nikkei Newspaper Digital Archive Project**

For more than 100 years, the *North American Post* served as a valuable resource for preserving community and culture. According to the newspaper's website and the Hokubei Hochi Foundation, the *North American Post* is the largest and oldest Japanese-language newspaper published in the Pacific Northwest.<sup>10</sup> This claim is based on the newspaper's affiliation with a daily newspaper, the *Hokubei JiJi*, which began publishing in 1902. The *Hokubei JiJi* continued to publish until the beginning of World War II, when all Japanese-language newspapers in Seattle closed as Japanese American citizens were moved to internment camps. After the war, as some of these citizens returned to the city, the *North American Post* began publishing its newspaper from the same address that the *Hokubei JiJi* had used before the war. The *North American Post* published weekly in 1946 and began publishing daily in 1950. In

---

<sup>10</sup> The newspaper, as mentioned earlier, was not published during World War II when Japanese Americans were interned.

1981, it began publishing two or three days a week through the present moment.<sup>11</sup> As of 2013, the *North American Post* publishes an online and print edition.<sup>12</sup>

The decision to begin the digitization project with the post-World War II years of 1946–1947 aligns with the mission of the Hokubei Hochi Foundation and its executive director.<sup>13</sup> In 2011, the foundation began to promote and preserve the culture of the Nikkei and Japanese community through educational and cultural projects and activities in the Pacific Northwest. Its executive director, Elaine Ikoma Ko, has served several leadership roles in the region.<sup>14</sup> The foundation work also has a personal dimension, as Ko’s grandfather, Sadahiko Ikomo, was the newspaper’s editor and publisher in 1946. He started the newspaper after he was released from the Minidoka internment camp in Idaho and returned to Seattle. Ikomo had been sent there with

---

<sup>11</sup> It is identified in many sources as a Japanese-language newspaper. However, it should be noted that the newspaper contains some English-language content as well, including the introduction of an “English Section” in the January 1, 1947, issue, and the newspaper’s contemporary bi-lingual, Japanese- and English-language content. In an editorial in the January 1, 1947, issue, Sadahiko Ikoma, the editor and publisher, said that the English Section was making its “initial bow,” and that the hope was “to enlarge its English Section and finally [publish] ‘all in English’” as the new generation of Japanese Americans comes of age. That English-language page followed eleven pages in Japanese language in a twelve-page that issue.

<sup>12</sup> In 2012, the *North American Post* purchased a Japanese-language newspaper competitor called the *Soy Source* and publishes this newspaper alongside its own.

<sup>13</sup> The aim was to include a year’s worth of publishing. Some recent materials discussing the project now describe the goal of digitizing 1946 and the year 1947 is not included.

<sup>14</sup> Elaine Ikoma Ko was the InterIm Community Development Association’s executive director and operations director, the founding director of the International District Housing Alliance, the director of the City of Seattle’s Office for Women’s Rights, and the coordinator for the King County Women’s Program. René Brogan, “Elaine Ikoma Ko: Challenging Careers in Civic Activism,” *University of Washington Department of Communication Alumni News*, August 9, 2012. Accessed 1 April 2013. <http://www.com.washington.edu/2012/08/elaine-ikoma-ko-challenging-careers-in-civic-activism/>. Amanda Ma, “Alumni Hall of Fame: Advocate for Social Justice Elaine Ikoma Ko.” *University of Washington Department of Communication Alumni Profile*, October 18, 2012. Accessed 1 April 2013. <http://www.com.washington.edu/2012/10/alumni-hall-of-fame-advocate-for-social-justice-elaine-ikoma-ko/>.

his family, including Ko's mother, during World War II.<sup>15</sup> One of Ko's stated hopes is that she can learn more about Ikoma through the archived issues of the newspaper.<sup>16</sup> Tomio Moriguchi, the newspaper's current publisher, is also the foundation's board president.<sup>17</sup>

As head of the University of Washington Libraries Microform and Newspaper collections, which is one of the largest in North America, Glenda Pearson provides the archival expertise needed for the NNDAP.<sup>18</sup> The collections demonstrate, in her words, "a sustained interest in ethnic, minority, and special audience newspapers."<sup>19</sup> Pearson and other experts in the library work together to coordinate newspaper scans, conduct metadata analysis, and more.

I first learned of the effort to archive the *North American Post* and the NNDAP in August 2012 (about a year after the project began) during a chance encounter with Ko.<sup>20</sup> She briefly described the nascent archive project that was underway. By November, a news story published details about the foundation-library partnership and project.<sup>21</sup> At that time, I began discussions with two university undergraduate students in an interdisciplinary program—one in Media and Communication Studies and one in Community Psychology—about how to engage with the

---

<sup>15</sup> Kirsten Johnson, "Bilingual North American Post to be digitized at UW Libraries." *University of Washington Department of Communication Alumni News*, November 13, 2012. Accessed 13 November 2012. <http://www.com.washington.edu/2012/11/bilingual-north-american-post-to-be-digitized-at-uw-libraries/>.

<sup>16</sup> Ibid.

<sup>17</sup> Tomio Moriguchi, the publisher of the *North American Post* and *Soy Source*, is the son of Fujimatsu Moriguchi, who founded Seattle's Uwajimaya Asian Grocery & Gift Markets, a local and family-owned specialty supermarket chain.

<sup>18</sup> Pearson and the Digital Initiatives Program worked with the Hokubei Hochi Foundation from the beginning.

<sup>19</sup> Personal correspondence with Kristin Gustafson, February 28, 2013.

<sup>20</sup> Ko stopped by the office during interviews for a separate research project.

<sup>21</sup> Kirsten Johnson, "Bilingual North American Post to be digitized at UW Libraries." *University of Washington Department of Communication Alumni News*, November 13, 2012. Accessed 13 November 2012. <http://www.com.washington.edu/2012/11/bilingual-north-american-post-to-be-digitized-at-uw-libraries/>.

NNDAP. By January 2013, the students began pursuing undergraduate research credits for their participation in the NNDAP.

### **The Digital Turn and Media History**

Making newspapers and periodicals available digitally changes how we *do* media history. This *doing* is influenced by the form these media take once digitized, the historical methods digitization allows, and issues of access. Nicholson points out that digitization of historical records has changed even in the past decade, as historians have significantly more titles available online in 2012 than in 2002.<sup>22</sup> Scholarship on methods for digitizing historical primary sources is relatively new. Yet the changes in *how* we do historical research now are comparable to the “revolution” in scholarship in the late 1980s when historians began writing “new kinds of history based on a growing awareness of the power of culture and language in shaping past experiences and societies.”<sup>23</sup> These new discussions about digitization occur across disciplines and bring some disciplines closer in unexpected ways. Nicholson suggests that cultural history—a discipline grounded in “the close examination of written and spoken texts” and “the power of representation and place of discourse”—has distanced itself from computer-based research in past years. Yet it has much to gain from using keyword search and related technology “to trace the development and movement of ideas and discursive formations in ways that were once impossible.”<sup>24</sup>

Some newspaper digitization projects have been criticized for failing to capture text in its original form. Examples of this preservation problem can be found in digitization projects for

---

<sup>22</sup> Bob Nicholson. “The Digital Turn: Exploring the Methodological Possibilities of Digital Newspaper Archives,” *Media History* 19, no 1 (2013): 59–73.

<sup>23</sup> *Ibid.*, 63.

<sup>24</sup> *Ibid.*, 63.

newspapers published across at least three centuries. For example, publishers of nineteenth-century newspapers and other periodicals routinely bound them into annual or semi-annual volumes to preserve so-called “ephemeral” objects of print culture. Today, digital copies are often created from these bound volumes, yet these volumes displaced content that appeared in the original single-issue copies, such as advertisements and covers.<sup>25</sup> This content has been lost to the contemporary scholar who does not have access to the original issues. A similar problem occurs in many twentieth-century digital press archives. For example, the Minnesota Historical Society carefully selected a range of historical documents for its *Duluth Lynchings Online Resource*, including newspaper content. Yet many of the newspaper articles included in the archive are divorced from the original newspaper itself. In other words, the articles made available through the computer screen lack the context of the page and pages through which the original audience understood them. As we lose this historical context, we also lose the ability to ask and answer historical questions that depend on this context.

Another concern regarding digitization has to do with users and access. Increasingly, students, scholars, historians, and others rely on search engines and online archives for their research, and those who rely exclusively on such tools may overlook critical texts not included in the body of online material.<sup>26</sup> In addition, many archives created by profit and nonprofit companies are not freely available to a wide public and are instead available only to those who belong to institutions, such as universities, with the resources to afford them.<sup>27</sup> And some digital archives present unique search issues that might stymie users. An analysis of *Multicultural*

---

<sup>25</sup> Laurel Brake, “The Longevity of ‘Ephemera,’” *Media History* 18, no. 1 (2012): 7–20.

<sup>26</sup> Patrick Leary, “Googling the Victorians.” *Journal Of Victorian Culture (Edinburgh University Press)* 10, no. 1 (2005): 72–86.

<sup>27</sup> James Mussell, “Ownership, Institutions, and Methodology,” *Journal of Victorian Culture* 13, 1 (spring 2008): 94–100.

*Canada*, a Canadian community newspaper digitization project designed “to digitize and provide free access to the heritage of Canada’s immigrant people,” mentions difficulties the site’s users might experience when searching Chinese-language content. Instead of using the traditional search method of finding “all the words” in an archive or document, they learned that it worked better to use a “phrase” search for the Chinese character-based language and its “pairs or multiples of characters forming a meaningful equivalent to an English word.”<sup>28</sup> Finally, content-selection decisions typically rest in the hands of librarians, funders, and the commercial companies doing the digitization, not researchers, argues Nicholson: “By the time we view a digital archive, its contents have been through a complex process of transformation; the questions we can ask of a digital source are, in this sense, determined before we even login to the database.”<sup>29</sup>

Although digitization projects raise various problems and concerns for historians, they also provide significant opportunities. They have often been praised as a “practical revolution” through which research has been made “faster, easier, more convenient and more productive.”<sup>30</sup> In some instances, historians no longer have to travel to distant libraries, struggle with microfilm machinery, and guess at which publication addressed a particular issue and when—and media historians in particular benefit, as newspapers, magazines, and periodicals make up the majority of digital materials as opposed to books, theater scripts, and letters.<sup>31</sup> Journalism historians, for their part, have been actively discussing the merits of digital archives. For example, in the spring

---

<sup>28</sup> Lynn Copeland, “Exposing Canada’s Historical Ethnic Newspapers Through the Multicultural Canada Project,” *World Library and Information Congress: 76<sup>th</sup> IFLA General Conference and Assembly*, August 10–15, 2010. Accessed 1 April 2013. <http://conference.ifla.org/past/2010/102-copeland-en.pdf>.

<sup>29</sup> Nicholson, 65.

<sup>30</sup> *Ibid.*, 61.

<sup>31</sup> Nicholson.

2013 issue of *Clio*, the newsletter of the History Division of the Association for Education in Journalism and Mass Communication, Keith Greenwood referred members to various online resources, including the *Center for Civil War Photography* online archives, and Kimberly Wilmot Voss informed members how to use social media to access *The Museum of American History* online archives.

Keyword searches are also important benefits of digital archives. Using keywords to locate information in texts can shift the research process in several ways. For example, it can change the direction of the research approach from a top-down to a bottom-up approach. Researchers can search full-text documents for the keyword “domesticity” and receive broader results than if they searched only document titles for the same keyword. Keyword searches (and searches using synonyms of those keywords) of full-text content thus help researchers understand that content from the bottom up.

Keyword searches also help historians establish quickly whether a particular newspaper meets the research needs. As Bingham has suggested, they help historians situate content by allowing them to make comparisons across time, between newspapers, and between different types and forms of newspaper content. Keyword searches can also invite historians to “conduct similar searches in different types of digital archive, placing, for example, newspaper editorials alongside speeches made in Parliament. Historians can raise their horizons and their research questions can be more ambitious; they can range more freely between periods and across borders.”<sup>32</sup>

With NNDAP, it became clear that decisions made by those creating the digital archive would affect how users accessed and made sense of the archives. The Hokubei Hochi Foundation

---

<sup>32</sup> Adrian Bingham, “The Digitization of Newspaper Archives: Opportunities and Challenges for Historians,” *Twentieth Century British History* 21, no. 2 (April 2010): 225–31, 229.

needed an ally and advice in order to make content searchable by keywords and available to a broad public via a host site, and thus the involvement of the University of Washington Libraries was critical.<sup>33</sup> The Digital Initiatives Program of the university libraries offers three areas of consultation generally, all of which are relevant to the NNDAP specifically: (1) content creation issues, which include “the creation of metadata for description of and access to digital files”; (2) software to display, store, and archive digital files; and (3) issues of preservation, which include deciding “appropriate standards for the long-term viability of digital files.”<sup>34</sup> Of the three software packages that the Digital Initiatives Program works with to provide access to and help display digital files, the NNDAP uses CONTENTdm.<sup>35</sup>

Decisions regarding the creation and use of NNDAP’s metadata are still in process and being worked out by the Digital Initiatives Program, the library, and the foundation.<sup>36</sup> But these issues—practical choices about what information should get left in and out and the summarizing nature of coding and indexing—are important for journalism historians to consider if they get involved in planning or using digital archive projects.

The metadata created for the NNDAP include personal names, locations, keywords, and corporate names found in every article and advertisement on a page. This information needs to be recognizable for a broad audience. To help make this happen, the NNDAP keywords are compared with those in the Library of Congress Authorities—which provides lists of subject headings used in reference sections of libraries—and revised in order to make the two lists

---

<sup>33</sup> According to their respective websites, the Seattle Foundation gave The Hokubei Hochi Foundation \$24,000 in 2011, a portion of which went to the NNDAP, and 4Culture gave the foundation \$2,600 in 2012 for the “Hokubei Hochi Newspaper Preservation Project.”

<sup>34</sup> University of Washington Libraries, “Digital Initiatives Program,” Accessed 3 September 2013. <http://www.lib.washington.edu/digital/>.

<sup>35</sup> The Seattle Foundation, “Nonprofit Organization Profiles: Hokubei Hochi Foundation,” Accessed 19 August 2013. <http://www.seattlefoundation.org/npos/Pages/HokubeiHochi.aspx>.

<sup>36</sup> The site is not yet active or ready for public use.

compatible. The US Library of Congress maintains a thesaurus of subject headings that helps libraries collect, organize, sort, and share information in ways that go beyond the title.<sup>37</sup>

To the credit of the volunteers, students, and professionals working on the NNDAP, there is a strong commitment to maintaining the integrity and context of the *North American Post* in the digitization process. However, some keywords used in the newspaper do not match perfectly with Library of Congress subject headings, and this means that some information is lost through the so-called “practical revolution” in digitization. For example, the keywords for “Japanese American Concentration Camps” do not fit the LOC subject headings, which would list this as “Japanese Americans—Evacuation and Relocation, 1942–1945.” So, as Glenda Pearson with the University of Washington Libraries points out, “You cannot, authoritatively, START with Japanese Americans—Concentration Camps or START with Concentration Camps—Japanese Americans.” This is because, as she explains, the U.S. federal government did not use the “C-word” to describe its internment of Japanese Americans once the German concentration camps became known.<sup>38</sup> A different example of the keyword problem that demonstrates how words might not translate across cultures or industries perfectly is an ad that appeared in the October 23, 1946. The advertisement was political in nature, a campaign ad for a candidate running for Pierce County sheriff. While “political advertisement” is the best LOC Authorities keyword for this item, there is no Japanese language equivalent for the English term “political advertisement.”<sup>39</sup>

---

<sup>37</sup> This process of organizing subject headings is not beyond critique. Sanford Berman was among the first to critique biases embedded in the listings. He worked as a head cataloger for Hennepin County Library in Minnesota. Sanford Berman, *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People* (Metuchen, NJ: Scarecrow Press, 1971).

<sup>38</sup> Personal correspondence with Kristin Gustafson, December 4, 2013.

<sup>39</sup> I thank Rena Kawasaki for explaining this process in greater detail.

Another metadata issue involves the creation of a thesaurus using personal names, locations, keywords, and corporate names found in the individual articles and advertisements. The gathering of the individual information for the names, locations, and keywords requires close attention to nuances found in the stories and language. But creating the thesaurus itself is also a complex undertaking. The Online Dictionary for Library and Information Science describes the intricacies and differences of the terms metadata, index, and thesaurus. Metadata is “structured information” that describes information, objects, and artifacts, which can serve different purposes of descriptive, structural, and administrative.<sup>40</sup> An index in a work alphabetically arranges headings with personal names, places, subjects, and page numbers (in the case of written work) and can rely on a “controlled vocabulary” that is developed over time.<sup>41</sup> And a “thesaurus” provides “synonyms and near-synonyms in a written language, usually arranged conceptually.” In library and information science, the term is sometimes equated with the term “controlled vocabulary.”<sup>42</sup> The aim of the NNDAP during its pilot stage is to create one thesaurus in English and one thesaurus in Japanese for the 1946–1947 content. The English version of the thesaurus relies on standards set by the LOC Authorities. However, there is no

---

<sup>40</sup> Joan M. Reitz, “Online Dictionary for Library and Information Science,” *CLIO*, Accessed 9 October 2013. [http://www.abc-clio.com/ODLIS/odlis\\_A.aspx](http://www.abc-clio.com/ODLIS/odlis_A.aspx). Descriptive metadata is intended to facilitate the “discovery, identification, and selection” of an information resource (oftentimes to aid usability); structural metadata is intended to describe “internal organizing” of that information resource (such as tying together various components); and administrative metadata is intended to facilitate management of an information resource (such as identifying restrictions to access or when a document was created).

<sup>41</sup> *Ibid.* A keyword index is a “type of subject index in which significant words, usually from the titles of the works indexed, are used as headings.” A subject index alphabetically arranges the list of headings selected by an “indexer” to represent subject content and can include names, geography, and authors.

<sup>42</sup> *Ibid.* A controlled vocabulary provides an “established list of preferred terms from which a cataloger or indexer must select when assigning subject headings or descriptors in a bibliographic record, to indicate the content of the work in a library catalog, index, or bibliographic database.” The LOC Authorities described in this essay provide standards, which are a controlled vocabulary.

equivalent to the LOC Authorities for the Japanese keywords and thesaurus. These are challenges that can be overcome. But the solutions are not always obvious and have the potential to affect the way future researchers might access the original text.<sup>43</sup>

The issues I have just described highlight how critical the creation of metadata is in a newspaper digitization project and underscore the value of inter-disciplinary, inter-agency, and inter-campus collaboration in archival digitization projects. The people involved with NNDAP bring knowledge and perspectives that work together to explore the intricacies and address the complexities of the project. For example, a conversation between journalism historians and software experts might focus on the industry norms regarding political advertisements in the 1940s and determine how best to classify these in a digital file for future use. And a conversation between a Japanese studies librarian and a project translator working directly with the newspaper text might explore solutions to a translation problem that has few obvious precedents.

### **Philosophies and Methods of Preservation**

Researchers know that archives are much more than collections of texts and other material remains of the past. When we look closely at archives, we can see embedded in them fingerprints, falsehoods, gaps, and assumptions about what matters most in our historical record (for example, “great men”). Archives are socially constructed collections of socially constructed products, and so they reflect the same power, ideologies, stories, and rituals found in other

---

<sup>43</sup> Again, Rena Kawasaki provided help in explaining the process. One solution considered for building the Japanese-language thesaurus was to use a Japanese-language history book as a reference. However, most of these focus on Japan’s history, whereas the newspaper’s content centered on US history or Japanese history as seen from the lens of Japanese Americans. Another solution offered by Azusa Tanaka, who is a Japanese Studies Librarian from the University of Washington Libraries East Asia Library and consulted on the NNDAP project, might be to use a Japanese-language book of U.S. history. This is one example of the kind of decisions that digitizing archives requires for creating metadata.

socially constructed artifacts. Some materials are left out while others are elevated—and this selection process necessarily shapes the stories historians tell about the past. Of course, we know that a researcher is “acted upon by contextual pressures, influenced by prevailing intellectual trends, anchored in tradition or torn between traditions, and shaped by his/her own training, by ideology and societal prejudices.”<sup>44</sup> Yet the contents and shape of archives also act upon a researcher. Archives represent a “lived process.” They can be used as sources but they can also be viewed as subjects or the focus of study.<sup>45</sup>

Archives have narratives. While this is not a new idea, it is worth mentioning in order to contextualize the methods of preservations discussed in this essay. Archivists produce descriptive guides, machine-readable cataloging records, encoded archival descriptions, and other documents to help researchers understand the stories about the contents of archival collections.<sup>46</sup> Understanding the chain of custody of materials as well as the original order or arrangement of the materials plays a central role in archiving.<sup>47</sup>

This way of thinking about archives and artifacts is familiar to media historians. News content—such as newspapers and TV newscasts—represent organizational products created by media organizations; analysis of news content can provide insight into the process of “how something is actually created and put together.”<sup>48</sup> It is useful to remember, too, that the media

---

<sup>44</sup> Marha P.Y. Cheung, “The Mediated Nature of Knowledge and the Pushing-hands Approach to Research on Translation History,” *Translation Studies* 5, no. 2 (2012): 156–171, 156.

<sup>45</sup> Lucille M. Schultz, “Foreword,” *Beyond the Archives: Research as a Lived Process*, ed. Gesa E. Kirsch and Liz Rohan (Carbondale: Southern Illinois University Press, 2008), vii.

<sup>46</sup> *Ibid.*

<sup>47</sup> *Ibid.*

<sup>48</sup> David L. Altheide, *Qualitative Media Analysis: Qualitative Research Methods*, 38 (Thousand Oaks, CA: Sage Publications, 1996), 4.

represented in these archives present socially constructed narratives.<sup>49</sup> Media disseminate information as well as portray and confirm “a particular view of the world.”<sup>50</sup> For Carey, news reading and news writing are dramatic and ritual acts through which “reality is produced, maintained, repaired, and transformed.” In reading or watching the news, audiences confront “not pure information but a portrayal of the contending forces in the world.”<sup>51</sup>

Narratives can also be found in archives and in the digitizing process. One of the fingerprints often found in the process of digitizing newspapers is optical character recognition (OCR). Although OCR was a method originally considered for NNDAP, the condition of the printed archives and other factors such as the complexities of the written language required this option to be rejected. The original University of Washington Libraries’ microfilm holdings of the *North American Post* were missing some issues and were in poor condition. The university retrieved some newspaper issues and re-filmed a portion of the archive in the 1990s. However, it did not re-film everything.<sup>52</sup> Filming is relatively expensive. For example, it costs the university \$500 to get one reel of positive film ready for public use, and currently the budget covers about a third of what it has already committed to filming.<sup>53</sup> The *North American Post* files that were not re-filmed in the 1990s were in poor condition for the OCR process. The issue of film condition is especially important, as oftentimes the original newspapers used for today’s microfilm might be

---

<sup>49</sup> Peter L. Berger and Thomas Luckmann, *The Social Construction of Reality: A Treatise in the Sociology of Knowledge* (Garden City, NY: Doubleday & Company, Inc., 1966). Norman Fairclough, *Discourse and Social Change* (Cambridge, United Kingdom: Polity Press, 1992). Stuart Hall, “Culture, the Media and the ‘Ideological Effect’,” in *Mass Communication and Society*, ed. James Curran, Michael Gurevitch and Janet Woollacott, (London, United Kingdom: The Open University Press, 1977), 315–347.

<sup>50</sup> James W. Carey, *Communication as Culture: Essays on media and society* (New York: Routledge, 1989 [reprinted 1992]), 20, 23.

<sup>51</sup> *Ibid.*, 20, 23.

<sup>52</sup> Personal correspondence with Kristin Gustafson on December 4, 2013.

<sup>53</sup> *Ibid.*

discarded due to storage constraints. This was not the case when the University of Washington Libraries re-filmed the *North American Post* two decades ago. At that time, the library borrowed printed copies from the publisher's run of the newspaper. However, not all newspapers have the resources or ability to maintain such archives. And most libraries now discard paper copies after filming, Pearson says. For example, almost all of the currently published newspapers that the University of Washington Libraries has committed to microfilm—including many smaller state newspapers, some out-of-state newspapers, and some foreign publications that no one else is filming—get recycled after the filming process because there is not enough storage space.<sup>54</sup> This practice is common.

Still, despite the condition of the archives, most digital newspaper projects will consider the use of OCR, so it is important to understand how it works. OCR mechanically or electronically converts a scanned image into a searchable, text-based digital document with key words and/or text. Its history can be traced to 1809, but the OCR technology came into its own in the 1950s when the U.S. Department of Defense created a device that could read Morse Code. The technology developed further in the late 1990s with the accessibility and affordability of personal computers and high quality scanning equipment.<sup>55</sup> Put simply, OCR allows for keyword searches. When it comes to digital newspaper archives, digital methodology assumes that the hard copy “is fundamentally different” from the digitized version.<sup>56</sup> The OCR process makes this

---

<sup>54</sup> Ibid.

<sup>55</sup> Simon Tanner, Trevor Muñoz, and Pich Hemy Ros, “Measuring Mass Text Digitization Quality and Usefulness: Lessons Learned from Accessing the OCR Accuracy of the British Library’s 19<sup>th</sup> Century Online Newspaper Archive,” *D-Lib Magazine* 15, no. 7/8 (2009). <http://www.dlib.org/dlib/july09/munoz/07munoz.html>. Tanner et. al point to Schantz’ scholarship on reading aids to the blind, which were patented in 1809. Herbert F. Schantz, *The History of OCR, Optical Character Recognition* (Manchester Center, VT: Recognition Technologies Users Association, 1982).

<sup>56</sup> Nicholson, 64.

difference even more apparent. Nicholson explains that the obvious difference is that “one source is made from paper” and “the other exists as billions of 1s and 0s.”<sup>57</sup> But digitizing text is much more than this material difference. As Nicholson points out, “Unlike microfilming, the creation of a digital newspaper does not simply produce what archivists term a ‘surrogate’, or stand-in, for the original. Instead, it creates something new; sources are ‘remediated’ and not just reproduced.”<sup>58</sup> And the “fingerprints” embedded in these newly created or remediated digital newspapers as a result of digitization become part of that archive’s narrative. The OCR software “is still not highly accurate because of the quality of the original source material, particularly text on poor quality newsprint” explains Glenda Pearson.<sup>59</sup> Seventy percent accuracy is considered “good” at this point. “In speaking about how the digital differs from the original, this seems an important point, particularly when depending on keyword searching to find articles, giving the researcher the impression that he/she has found all instances of the word(s) searched, and NOT taking more time to browse,” she says. To do it well, a good OCR database has employed the software first and then the database is corrected manually by “a real live human being who does a side-by-side comparison” of the two texts.<sup>60</sup>

Methods of preservation challenge not only how we put together archives and how we use archives, but they also challenge our epistemological stance. Martha P.Y. Cheung describes this challenge in her work on translation history: “The idea that meaning is interpreted, constructed and refracted by the political or ideological stance of a researcher and his/her view of

---

<sup>57</sup> Ibid., 64.

<sup>58</sup> Ibid., 64.

<sup>59</sup> Personal correspondence with Kristin Gustafson on December 4, 2013. Pearson suggests that these challenges support an argument for digital image databases, which would make the entire published page viewable and readable, “thus allowing for more serendipity, browsing, and almost demanding that the context of a particular article be considered, or at least making it easily possible to consider.”

<sup>60</sup> Ibid.

history is unsettling to many historians because it challenges, at a fundamental level, the very existence of the discipline.”<sup>61</sup> The process of translation surfaces questions such as whether the source is primary or secondary or a “hyphenated category of its own”; whether it should be dismissed or ignored by native speakers; and whether “knowledge derived from translated sources” is sufficient for a historian to speak with authority on a given subject.<sup>62</sup>

The process of *how* metadata are created is one that is typically hidden from the view of the end user much in the same way a student might not see the planning of a college course or a visitor might not see the architectural drawings of a house. But this process is important. One example from the NNDAP is illustrative. The character-based portion of the Japanese language, known as Kanji, has changed over time. When Kawasaki began working as a student researcher on the NNDAP project, she noticed how her use of Kanji in 2013 differed at times from the use of Kanji in the 1946 issues of the *North American Post*.<sup>63</sup> Specifically, the characters used in the 1946 text included more strokes than they did in 2013. As a result, the 1946 characters had a more intricate and complex meaning. To complicate matters further, at times there were variations in meaning between the old and new Kanji characters—such as the old 聯 and the new 連—used interchangeably within the same *North American Post* article.<sup>64</sup>

---

<sup>61</sup> Marha P.Y. Cheung, “The Mediated Nature of Knowledge and the Pushing-hands Approach to Research on Translation History,” *Translation Studies* 5, no. 2 (2012): 156–171, 158.

<sup>62</sup> *Ibid.*, 157.

<sup>63</sup> Kawasaki, Rena, “How Language Use Indicates Acculturation and Enculturation Processes of Native Japanese and Japanese Americans: Examples from the 1946 issues of *Hokubei Hochi*,” unpublished work (2013).

<sup>64</sup> Kawasaki explained that in an October 23, 1946, *North American Post* article, the first version of the word “consumed” was written as 盡 in its old form of Kanji and 尽 in the new. She deconstructed the old Kanji. The older form is composed from 筆, which means brush, and 皿, which means plate. The literal meaning, therefore, is to brush off whatever is on the plate until it is empty. This notion of void or nothingness is central to “Zen” a branch of Japanese philosophy. However, in the second form of the word “consumed,” this ideology is entirely omitted. When the Kanji is simplified to 尽, the character still possesses the meaning of “corpse,” which partly

This change in language—either within the same 1946 text or across decades as the metadata created for the original text—affects the NNDAP metadata. The project not only undertakes Japanese-to-English translation of the Japanese content and English-to-Japanese translation of the English content, but also involves making words used in 1946 understandable to contemporary and future audiences. For example, newspaper headlines written with the older Kanji might be misunderstood or unintelligible to a contemporary researcher or web user, so one step taken in the NNDAP is to update the metadata material so that it makes sense for a contemporary keyword search. A researcher can still identify the original headline because the metadata ultimately is linked to the original text. As this discussion illustrates, information can get lost without careful thought to how future users link to these archives.<sup>65</sup>

Presentism or present-mindedness is a practice historians strive to avoid. This is the practice of “viewing the past in terms of the present” rather than more accurately presenting the past by capturing the spirit, feelings, persuasions, emotions, forces, and things of the past within context.<sup>66</sup> Avoiding present-mindedness is part of a larger mission among U.S. professional historians to approach history objectively.<sup>67</sup> Archivists, too, strive to avoid present-mindedness in their preservation activities. But this is an aspirational ideal that is often impossible to meet, as the NNDAP project illustrates. Making the Japanese-language content from the 1940s accessible

---

shares the meaning of nothingness, but has lost its totality. Kawasaki cited the Kanji dictionary as her guide in understanding this translation: Tadashi Kamada and Torataro Yoneyama, *Kangorin*, (Tokyo: Taishūkan Shoten, 1987).

These changes occurred in the time frame when the language transformation movement took place in Japan. “After the Language Reform Policy, certain parts of Kanji characters were simply omitted in some Kanjis. In others, they were replaced by simplified components that often do not portray the meanings of the original characters.” Kawasaki 2013, 7.

<sup>65</sup> Again, I thank Rena Kawasaki for explaining this process in greater detail.

<sup>66</sup> James D. Startt and Wm. David Sloan, *Historical Methods in Mass Communication* (Northport, AL: Vision Press, 2003), 51.

<sup>67</sup> Peter Novick, *That Noble Dream: The “Objectivity Question” and the American Historical Profession* (New York, NY: Cambridge University Press, 1998).

to present-day Japanese and English speakers requires a certain amount of interpretation that then leaves a contemporary narrative fingerprint on the archives.<sup>68</sup> Perhaps the lesson of NNDAP is that interpretations made as part of preservation methods require careful thought.

### **Complexities and Selection in Archiving: A Closer Look at an Asian American Newspaper**

If successful in its attempts to get additional funding, the NNDAP could produce a lasting archive of a historically important newspaper serving a historically important ethnic community in the United States across more than a century. If unsuccessful, the voices that the project seeks to preserve might fail to be included with the more familiar, dominant, and already archived voices. The latter option is familiar to archivists if profoundly undesirable. As Jobs and Lüdtkke have discussed, archivists are not ready and waiting for every project researchers might pursue: “The assumption of *the* archive as a natural being also propels the impression that archivists would be ever ready for any request researchers may venture.”<sup>69</sup>

Why is the *North American Post* so important to include in the increasingly large digital repository of historically important US newspapers? The answers lie in the history of the newspaper’s region and community. In the late nineteenth century, the two main ports of entry to

---

<sup>68</sup> Trying to determine how to translate the metadata for a contemporary audience without losing the rich meaning of the past takes time. At an early stage in the project, two undergraduate researchers worked together to capture nuances in language. One person took the Kanji character and translated what it meant literally; the second person tried to make it make sense of that translation in the context of 2013. For example, the term for refrigerator was written in Kanji as “cold-container” in 1946, explained Rena Kawasaki. It would have been translated to the modern Kanji term of “cold-storage” to make sense for a contemporary audience using the metadata. Kawasaki 2013.

<sup>69</sup> Sebastian Jobs and Alf Lüdtkke, “Unsettling History: Introduction,” in *Unsettling History: Archiving and Narrating in Historiography*, ed. Sebastian Jobs and Alf Lüdtkke, (Frankfurt: Campus Verlag, 2010), 7–25, 15–16. Italics in original.

the US mainland were San Francisco and Seattle, and Washington State's Japanese American population grew significantly at the turn of the century alongside the whole population of the state. The Japanese population grew from 360 (out of 357,232 total) in 1890, to 5,617 (out of 518,103) in 1900, to 12,929 (out of 1,141,990) in 1910. By 1940, the population was 14,565.<sup>70</sup> Seattle became the "headquarters" for these new immigrants with local businesses serving day-to-day needs, and Japanese Americans worked inside and outside the city, including hotels, railroad construction, farming, sawmills, and canneries.<sup>71</sup> In 1917, 85 Japanese Americans worked for one of three daily newspapers, according to a survey by the Ministerial Federation.<sup>72</sup>

Laws and rulings have affected the lives of Japanese Americans, often negatively, and Seattle was no exception. Japanese immigrants in the Pacific Northwest "were well-versed in intensive cultivating farming methods and played a key role in developing the area's agricultural industry,"<sup>73</sup> and in 1920, farmers with Japanese heritage provided three quarters of the region's vegetables and half the milk.<sup>74</sup> However, despite these contributions and important role in the fabric of the region, Japanese Americans faced several forms of discrimination. They were not allowed to own land; they were required to live in certain areas; they were severely restricted in their ability to become naturalized U.S. citizens; and they suffered various discriminatory social practices, including segregated seating in theaters and exclusion from public beaches.<sup>75</sup> One

---

<sup>70</sup> U.S. Census "Characteristics of the Population" reported these numbers. Bureau of the Census. "Population: Volume II Characteristics of the Population," in *Sixteenth Census of the United States: 1940* (Washington, DC: United States Government Printing Office, 1943), 304.

<sup>71</sup> Doug Chin, *Seattle's International District: The Making of a Pan-Asian American Community* (Seattle, WA: International Examiner Press, 2001).

<sup>72</sup> There is no indication of which newspapers. The Seattle Ministerial Federation, *Report of the Committee on Orientals* (Seattle, WA: Seattle Ministerial Federation, June 4, 1917).

<sup>73</sup> Chin, 32.

<sup>74</sup> David A. Takami, *Divided Destiny: A History of Japanese Americans in Seattle* (Seattle, WA: University of Washington Press and Wing Luke Asian Museum, 1998).

<sup>75</sup> Ibid.

devastating harm began February 19, 1942, when President Franklin Roosevelt signed Executive Order 9066 during World War II. This government order authorized military commanders to evacuate persons of Japanese ancestry from their West Coast homes and businesses. When Japanese Americans returned to the area, many of them found they had permanently lost their jobs, homes, investments, savings, and businesses. And some of their former friends had joined the nation in considering the Japanese race as the enemy.<sup>76</sup>

Journalism historians in the United States have paid some attention to Japanese Americans, their history, and the press publications that served their communities. Historical research addresses the internment, discrimination, and World War II. This history includes attention to the news coverage,<sup>77</sup> film and government documents,<sup>78</sup> pre-war media attention,<sup>79</sup> and newspapers published in the camps constructed to contain Japanese Americans during World War II.<sup>80</sup> Yet there is more historical work to be done.

---

<sup>76</sup> Chin, 69.

<sup>77</sup> Ronald Bishop, “‘Little More than Minutes’: How Two Wyoming Community Newspapers Covered the Construction of Heart Mountain Internment Camp,” *American Journalism* 26, no. 3 (summer 2009): 7–32. Patricia A. Curtin, “Press Coverage of the 442<sup>nd</sup> Regimental Combate Team (Separate Nisei): A Case Study on Agenda Building,” *American Journalism*, 12, no. 3 (summer 1995): 225–241. Hamm, Bradley J., “Redefining Racism: Newspaper Justification for the 1924 Exclusion of Japanese Immigrants,” *American Journalism*, 16, no. 3 (summer 1999): 53–69.

<sup>78</sup> Thomas B. Christie and Andrew M. Clark, “Framing Two Enemies in Mass Media: A Content Analysis of U.S. Government Influence in American Film during World War II,” *American Journalism*, 25, no. 1 (winter 2008): 55–72. Betty Houchin Winfield and Janice Hume, “Historiographic Essay: Shhh, Do Tell! World War II and Press-government Scholarship,” *American Journalism* 12, no. 3 (summer 1995): 367–383.

<sup>79</sup> Thomas H. Heuterman, “‘We Have the Same Rights as Other Citizens’: Coverage of Yakima Valley Japanese Americans in the ‘Missing Decades’ of the 1920s and 1930s,” *Journalism History*, 14, no. 4 (winter 1987): 94–103.

<sup>80</sup> Lauren Kessler, “Fettered freedoms: The Journalism of World War II Japanese Internment Camps,” *Journalism History*, 15, no. 2–3 (summer/autumn 1988): 70–79. Catherine A. Luther, “Reflections of Cultural Identities in Conflict: Japanese American Internment Camp Newspapers During World War II,” *Journalism History*, 29, no. 2 (summer 2003): 69–81. Takeya Mizuno, “The Federal Government’s Decisions in Suppressing the Japanese-Language Press, 1941–42,”

Washington State and Seattle offer media historians a rich selection of media to study. The Japanese YMCA published *The Report* from 1899 until about 1904. And from 1902 to 1942, at least nine other Japanese-language or Japanese community newspapers published in Seattle—the *Hokubei JiJi* (*The North American Times*), *Northwestern News*, *Asahi Shinbun*, *Gei-Bi-Jin* (or *Man-Ji-Ho*), *Great Northern American Times*, *Kobushi*, *Donchike*, *Japanese-American Review*, and the *Japanese-American Courier*.<sup>81</sup> However, it is important to note that no Japanese-language newspaper published in the city from 1942–1945; they all ceased publication when the US government forced Japanese Americans into internment camps. In 1946, some Japanese Americans returned to their homes and began to publish newspapers again. The first was the *North American Post* (also known then and now as the *Hokubei Hochi*), which resumed publishing in 1946 after four years of silence.<sup>82</sup>

This history is relevant to the preservation of newspapers and to the NNDAP. The selection of any archive we preserve requires labor and money. The work requires hours and hours of effort, and most of the labor invested in the NNDAP since 2011 comes from volunteers, interns, and professionals lending a hand. And the funding is limited. However, there are two

---

*Journalism History* 33, no. 1 (spring 2007): 14–23. Takeya Mizuno, “Journalism Under Military Guards and Searchlights: Newspaper Censorship at Japanese American Assembly Campus During World War II,” *Journalism History* 29, no. 3 (fall 2003): 98–106. Takeya Mizuno, “Self-censorship by Coercion: The Federal Government and the California Japanese-language Newspapers from Pearl Harbor to Internment,” *American Journalism* 17, no. 3 (summer 2000): 31–57. John D. Stevens, “From Behind the Barbed Wires: Freedom of the Press in World War II Japanese Centers,” *Journalism Quarterly*, 48 (summer 1971): 279–87.

<sup>81</sup> Four of the newspapers were listed with variations in the name in a 1920 report of Seattle businesses by Phillip Tindall and the Seattle City Council: *Great Northern*, *North American Times*, *Ashi News*, and *Japanese American Review*. U.S. Congress, “Part 4: Hearings at Seattle and Tacoma, Wash.; Appendix; Index,” *The Committee on Immigration and Naturalization House of Representatives, Sixty-sixth Congress, Second Session, July 26, 27, 28, 29; August 2 and 3, 1920*, (Washington, DC: Washington Government Printing Office, 1921), 1119.

<sup>82</sup> Greg Lang, “Seattle’s First Japanese Newspaper, The Report, is Issued in 1899,” *History Link: The Free Online Encyclopedia of Washington State History*, June 10, 2001. Accessed 5 September 2009. [http://www.historylink.org/index.cfm?DisplayPage=output.cfm&file\\_id=3352](http://www.historylink.org/index.cfm?DisplayPage=output.cfm&file_id=3352).

concerns related to availability that supersede the need for labor and money. The availability of newspapers for research can be affected by a community's displacement through internment.

And access to the newspapers can be affected by U.S. Copyright Law. Both of these issues might make it difficult for traditional archival projects to invest in preserving the *North American Post* or newspapers like it through digitization.

To understand the problem of access, consider the National Digital Newspaper Program's efforts nationally and locally. This grant program, as mentioned earlier, is a collaboration between the Library of Congress and the National Endowment for Humanities. In two decades, the program will create a "national, digital resource of historically significant newspapers from all the states and U.S. territories published between 1836 and 1922."<sup>83</sup> People can now engage with newspapers from around the country using a searchable, free Internet database maintained permanently at the Library of Congress. Every year, the National Digital Newspaper Program digitizes 100,000 more pages in each of its state programs.<sup>84</sup> Washington State is one of 36 states involved in the national program, as of October 2013.<sup>85</sup> Since 2008, the Washington State projects have received \$935,424.<sup>86</sup> Hosting such a project is expensive in terms of technology, as is maintaining the site in years to come. The selection and digitizing efforts require funding as

---

<sup>83</sup> Washington Secretary of State Wikis, "WA State National Digital Newspaper Program Wiki," Accessed 26 January 2013. <http://wiki.secstate.wa.gov/ndnp/MainPage.ashx>.

<sup>84</sup> The 2005 and 2007 awards were for years 1880–1910. In 2008, the focus was 1880–1922. In 2009, the years were 1860–1922. And in 2010, 2011, and 2012, the years were 1836–1922. The Library of Congress, "Chronicling America: Historic American Newspapers," *Chronicling America*. Accessed 26 January 2013. <http://chroniclingamerica.loc.gov/>.

<sup>85</sup> The Library of Congress, "Chronicling America: Historic American Newspapers," *Chronicling America*. Accessed 23 October 2013. <http://chroniclingamerica.loc.gov/>.

<sup>86</sup> National Endowment for the Humanities, "Funded Projects Query Form," Accessed 23 October 2013. <https://securegrants.neh.gov/publicquery/main.aspx>. Grant period: 7/1/2008–8/31/2014. The original grant in 2008 was \$341,424, and supplement grant in 2010 was \$284,000 and in 2012 was \$310,000.

well. In deciding which newspapers to digitize, decision-makers must necessarily undertake financial cost-benefit analyses.<sup>87</sup>

Another major consideration in selecting newspapers for digital archives is US Copyright Law. By deciding to digitize newspaper issues published prior to 1922, the National Digital Newspaper Program avoids the extra work of securing permissions to archive newspaper content protected by copyright. And then there is the availability of and access to the newspaper itself. Newspapers that archivists and researchers have identified in past years as valuable and worth preserving make it easier for these same newspapers to be used for future digital preservation projects. This point is obvious, and yet it is worth saying because newspapers such as the *North American Post* might not have been recognized in years past, as our sense of history as diverse and inclusive continues to evolve. When we consider both of these issues—US Copyright Law and the availability of newspaper content—and compare the National Digital Newspaper Program with the Nikkei Newspaper Digital Archive Project, we can easily see how issues of access and copyright law might lead to significant exclusions from the archival record. Of the newspapers selected for the state of Washington’s National Digital Newspaper Project, not one was created for a Japanese-American community.

The decision to select newspapers published before 1922 for the state and national digital archive programs likely excludes newspaper operations similar to the *North American Post*. Part of the problem is that many issues of the *North American Post* newspaper published before

---

<sup>87</sup> The expectation is that program participants “digitize primarily from microfilm holdings for reasons of efficiency and cost, encouraging selection of technically-suitable film, bibliographic completeness, diversity and ‘orphaned’ newspapers (newspapers that have ceased publication and lack active ownership) in order to decrease the likelihood of duplicative digitization by other organizations.” The Library of Congress, “Chronicling America: Historic American Newspapers,” *Chronicling America*. Accessed 26 January 2013. <http://chroniclingamerica.loc.gov/>.

World War II are missing despite the newspaper's, the foundation's, and the University of Washington Libraries' repeated efforts to recover them. The closure of the Japanese newspapers in Seattle—including the *North American Post*'s predecessor, the *Hokubei Jiji*—and the removal of these Americans from their homes and businesses for the duration of World War II—disrupted the newspaper's continuous record of publication.<sup>88</sup> And the internment, with its radical disruption of community practices, likely contributed to the loss of earlier issues of the newspaper. Although members of the Japanese American community in Seattle—bound by culture, geography, or language—surely valued the newspaper, others in the broader non-Japanese American community might not have recognized the importance of the newspaper in the region's shared history and thus missed out on preservation opportunities.<sup>89</sup>

## Conclusion

If the NNDAP is successful in digitizing the extant run of the *North American Post*, it will radically expand access to past issues of the newspaper. Researchers and community members will no longer be restricted to pulling a hard copy issue of the newspaper out of archival boxes or viewing reels of microfilm at a library. Instead, the information will be accessible to people with Internet access anywhere in the world. And end users will be able to

---

<sup>88</sup> Continued efforts of the University of Washington Libraries to secure a complete run of the newspapers have proven unsuccessful. There are several gaps in the *Hokubei Jiji*'s publication record that would fit the pre-1922 publication criteria of the National Digital Newspaper program. The library is missing record of all of the issues from 1902–Oct. 14, 1916, and then most of the 1920s issues (with the exception of January–March 1920).

<sup>89</sup> Glenda Pearson points out that it was not until Lucy Maynard Salmon's *The Newspaper and the Historian* in 1923 “that historians, at least in the U.S., really began to see newspaper content as significant resources to studying history.” Librarians have worked to preserve “a diversity of publications, almost none of which would have ‘qualified’ for preservation if more conservative voices had prevailed,” Pearson says. Lucy Maynard Salmon, *The Newspaper and the Historian* (New York, NY: Oxford University Press, 1923). Personal correspondence with Kristin Gustafson on December 4, 2013.

search and sort archives using carefully constructed metadata in Japanese and English. So a newspaper that once served only the readers of a Japanese-language newspaper in a specific community in Seattle could potentially cross the borders of city, state, nation, and continent—and cross barriers of language and engage interests across generations. This kind of access invites inter-cultural and intra-cultural conversation in Seattle and beyond.

The NNDAP helps us consider important issues related to digital newspaper preservation projects: (1) the design decisions that shape digital newspaper archives and influence how they can be used in the future; (2) the role preservation methods play in future historians' use of the archives and interpretation of material in the archives; and (3) the selection of newspapers to include in the archives. This essay has attempted to explore these issues in the interests of serving others who are interested in digitizing historic newspapers.

And the NNDAP offers many lessons about the importance of community-university partnerships in knowledge production and the value of engaging students in inquiry-based research and learning that employs innovative technology. The students involved in the NNDAP learned first-hand about the histories they discussed in their homes but were not often studied in their classrooms. Consider the reflections of KJ Hiramoto, one of the undergraduate students who worked on the project.<sup>90</sup>

As a second-generation Japanese American who attended American public schools all my life, my education relating to World War II came from a mainstream American context. In some history classes, we occasionally learned about the devastating experiences of the Holocaust from memoirs written by Jewish survivors, but very rarely did we learn about the experiences of World War II through a Japanese American's perspective.

---

<sup>90</sup> K.J. Hiramoto, "My Discovery with Nikkei Heritage," *North American Post* (Seattle, WA), June 27, 2013. Accessed 27 June 2013. <http://www.napost.com/2013/06/27/my-discovery-with-nikkei-heritage/>.

The multicultural nature of the project also presented an opportunity to engage students with Japanese-language and Japanese-culture skills in a way that brought their knowledge to the center of the learning. The two students I worked with directly spent two quarters working with bi-lingual metadata constructed through the NNDAP and then applying what they learned to their community psychology, journalism, and history coursework to explain the patterns in newspaper content and nuances in a community's language.

The history of archives plays an important and parallel role to the history of journalism. Archives and journalism become repositories of our past—sometimes the distant past, sometimes the recent past. Although the cultural turn has debunked a notion of a singular past,<sup>91</sup> there are an infinite number of ways to continue to challenge notions of a singular past, a singular public, and a singular journalism—and a singular archive of it all.

While historians are wise to the fact that these repositories have reflected the interests of the powerful for ages, a wider public might not know how decisions such as those discussed in this essay affect *how* and *why* we can access certain archives and not others.<sup>92</sup> That lack of awareness and the subsequent exclusion of some newspapers from digitization efforts are problems we must work to overcome. Projects such as NNDAP continue to challenge the notion of one universal archive as well as the assumption that archives are ready for the asking of researchers. In this essay, I have explored at least some ideas about how we might best continue to preserve historically important newspapers for scholars and the public to study in the years that stretch before us. Journalism history depends on it.

---

<sup>91</sup> Jobs and Lüdtké.

<sup>92</sup> Philipp Müller, "Ranke in the Lobby of the Archive: Metaphors and Conditions of Historical Research," in *Unsettling History: Archiving and Narrating in Historiography*, ed. Sebastian Jobs and Alf Lüdtké, (Frankfurt: Campus Verlag, 2010), 109–125.

**The following corrections were made to the manuscript after submission to the publisher.**

Citation 19 should read: Glenda Pearson, personal communication to author, February 28, 2013.

Citation 36 should read: "The site became active in February 2014."

Citation 38 should read: Glenda Pearson, personal communication to author, December 4, 2013.

Citation 52 should read: Glenda Pearson, personal communication to author, December 4, 2013.

Citation 59 should read: Glenda Pearson, personal communication to author, December 4, 2013. The remainder of the citation, beginning with "Pearson suggests...", remains as written.

Citation 68, should read "...tried to make the translation make sense in the context of 2013..." instead of "...tried to make it make sense of that translation in the context ..."

Citation 69 should only include the page ranges of "7–25" and delete the page ranges of "15–16."

Citation 89 should begin the same as written, with "Glenda Pearson points out ... 1923)." However it should conclude with this revised citation: Glenda Pearson, personal communication to author, December 4, 2013.