2015

The Digital Humanities Imperative: An Archival Response

Pauleena MacDougall
Maine Folklife Center, Pauleena.MacDougall@umit.maine.edu

Katrina Wynn
Maine Folklife Center, katrina.wynn@maine.edu

The Digital Humanities Imperative:

An Archival Response

by Pauleena MacDougall and Katrina Wynn

INTRODUCTION

Digital humanities (DH) is an emerging field that involves the creation, use, and preservation of digital humanities resources. But it can also include advocacy, research, teaching, publishing, investigating, analysis, and synthesis. Digital humanities is a subset of the humanities that grew out of literary and linguistic computing in the 1970s (see sidebar for timeline). Digital humanities scholarship is organized in an international group of professional societies under the umbrella of the Alliance of Digital Humanities Organizations (http://adho.org/). During the last 20 years, DH research has led to the development of publications such as the University of Illinois Press book series on Topics in the Digital Humanities and the journals *Digital Humanities Quarterly, Digital Medievalist, Journal of Digital Humanities*, and *Literary and Linguistic Computing*. Centers in several countries devote their research to digital humanities, even at a time when funding for the humanities at colleges and universities is flat or even decreasing.

As an emerging field, DH education and research is an important addition to the education of Maine students. It is not enough to be born in a digital world; Maine students need additional skills in digital literacy, global awareness, and interactive communication as they join the workforce, whether they are working in business, libraries, education, or other fields. The future of digital humanities and its many convergent practices seems nearly infinite. Policies that will most likely effect the practice are statewide policies that can enhance the work and educational opportunities of Maine citizens such as access to broadband, access to computers and software in schools, as well as support for humanities education at all educational levels and support for libraries and archives.

The scope of DH is vast and rapidly growing, and practitioners find it necessary to develop and learn about archival standards, digital collections management, computing humanities, blogging humanities, scholarly tools, databases, copyright laws, and networked writing—all deeply intertwined with traditional humanities, but expanding and growing exponentially in tandem with the Internet.

THE DIGITAL IMPERATIVE SHAPES PRESERVATION AND ACCESS

The work of archivists and librarians focuses on two primary goals: preservation of their holdings and providing public access to the material in their collections. For the archivist, the greatest current challenge to this

---

TIMELINE

- First desktop computer (mass marketed) Hewlett Packard, 1968
- Intel microprocessor, 1971
- Electronic digital computer patented, 1973
- First Apple personal computer, 1976
- First IBM PC, 1981
- ARPANET (science network expanded by National Science Foundation), 1981
- First Listserv (Paris, France, by Eric Thomas), 1986
- World Wide Web (by British Tim Berners-Lee), 1989
- MOSAIC web browser, 1993
- L-Soft listserv software, 1994
- Private and commercial use of Internet (email, instant message), mid-1990s
- First database, 1996
- National Endowment for the Humanities Digital Humanities Initiative, 2006

(Kirschenbaum 2010)

work is technological change. For example, sound recordings have been made on wax cylinders, wire cylinders, an array of tape sizes in reel-to-reel and cassettes, and vinyl records. Texts have appeared on many materials, though most commonly paper. Photographs and negatives are on paper and film. The archivist needs to preserve all of these materials through a series of actions including proper care (climate control, archival grade storage materials, minimizing handling) and making access copies. As the digital age has rapidly changed the technology of how sound, texts, and images are created, the work has taken on added complexity. As a result, the digital archivist needs to provide digital preservation and access to collections such as those housed at the Maine Folklife Center (MFC) at the University of Maine in Orono.

The MFC oversees the Northeast Archives of Folklore and Oral History, founded by Professor Edward D. Ives in the 1950s. The 160-plus collections in the archives contain over 3,900 individual accessions, including roughly 12,600 photographs, 2,500 slides, 3,000 audio recordings, and 325,000 pages of printed materials, covering a broad range of topics. The materials were collected beginning about 1958 and include interviews with people who lived in the late nineteenth century. The collection is continually growing, as staff and members of the public conduct fieldwork projects in folklore and oral history. Northeast Archives is the state's only major repository for folklore and oral history materials.

The collection is especially strong in documentation of occupations, foodways, community histories, lore and legends, traditional music, social activities, ritual and worship, material culture, and expressive arts. Specific topics include the logging and lumbering industry; fishing and lobstering; women in Maine; country and western music; northeastern multi-ethnic culture; labor history; Native Americans; and tourism and hunting.

Collections in the Northeast Archives of Folklore and Oral History include those that have been generated through various class activities and grant-based initiatives, as well as materials that have been deposited by donors from outside the University of Maine community. Much of the material is in the form of oral history interviews, but there are also class papers, topical surveys, field journals, photographs, commercial publications and recordings, exhibit materials, printed ephemera, and other kinds of documentation.

The MFC first ventured into the digital world in 1992 when the University of Maine's Raymond H. Fogler Library received a grant to add bibliographic information about the first 2,000 or so accessions of the Northeast Archives of Folklore and Oral History to its online URSUS database. Prior to that date, anyone who wished to access the collection had to come to the MFC and search through card file indexes. At the conclusion of the cataloging project, researchers and students could, for the first time, access information about the collection through the Internet.

Then in 1998, MFC's archivist began converting the card files of the Northeast Archives and Oral History to a database and creating collections descriptions that were then put on a newly formed MFC website. At that time, libraries and archives still recommended analog rather than digital copies for preservation purposes. As a result, when the MFC decided to create preservation copies of audio files, they did so on preservation master tapes. The in-house database provided more opportunities for searching the collection by keywords, but was still only accessible to visitors who came into the MFC.

PRESERVATION

In April 1999, the MFC received a National Endowment for the Humanities Preservation and Access grant to make preservation copies of its older, more endangered materials. As preservation masters of original analog-formatted audio files were created, listening copies were copied onto CDs. It wasn't until the early 2000s that born-digital materials began to arrive at the center with the advent of affordable digital recorders and digital cameras. The arrival of digital recordings, digital photographs, and digital documents required new thinking about preservation and access.

Collections descriptions grew on MFC's website as the result of this preservation and access effort, and patrons increasingly requested digital copies of audio, photos, and documents which the MFC did not have. To meet that challenge, the board of advisors for the MFC came up with a plan to digitize the entire collection with the help of the American Folklife Center at the Library of Congress. The Library of Congress made an offer to purchase and preserve the original tapes, photographs, and documents from our collection in exchange for a sum that would support digitizing this important collection.

Over the course of two years beginning in 2012, 4,433 hours of audio tape were sent to George Blood

audio/video/film preservation lab in Philadelphia for digitization. The digitized files were returned to MFC on hard drives. Students, under the supervision of Archives Manager Katrina Wynn, scanned 178,150 pages of documents, 18,015 photographs, 13,864 negatives, and 2,517 slides. The details of the undertaking were complex. Wynn created a number of digitizing policies and procedures to complete the task. In addition, she had to check the MFC's database for accuracy throughout the process. In conducting this work, we recognized that the digital files being created would need to be of high quality so we consulted the Library of Congress's Digital Preservation website (http://www.digitalpreservation.gov/) and the archivist (first Michael Taft, later Bertrand Lyons) at the American Folklife Center with specific questions.

Digitization processes are complex and include the necessity of recording metadata for each item. Metadata is information (some of which is embedded in the digital files and some in separate files) describing when, where, by whom, and how the digital files were created. This information is necessary so that, in future, when files need to be normalized, emulated, or migrated to accommodate new technology, the archivist will know how to proceed. Migration, the most common option, means updating a digital file as new software develops (like taking an old Microsoft Word file and updating it with each new version of Microsoft Word). This means there will always be access to the contents, but migration can affect the formatting and appearance of a digital file. Aside from the conversion errors that can occur with migration, the method takes time and must be done on a fairly regular basis. A second option is normalization, which is a type of migration, specifically into one type or a very few types of file formats, often open-source, nonproprietary ones, in an attempt to beat technological obsolescence. For example, an archive could convert all its photographs into TIFF format, whether their original was TIFF, JPEG, or GIF. Emulation, almost the opposite of the first two options, involves preserving the old version and using technology to mimic the old environment in order to open the old file and see it as it (mostly) originally appeared (this is fairly common for people to do with old computer games). This method depends on greater technical skill than migration and also access to the hardware to read materials created by older technology.

Digital preservation has its own technical challenges. Hardware and software can become obsolete, making access next to impossible. In addition, digital bits (0s and 1s) reside on a storage medium (CD or hard drive), and these bits can degrade over time. Furthermore, data loss for digital files is even more catastrophic than for analog because it can easily include the loss of an entire recording, rather than a minor loss of sound, for example. Technical issues that must be addressed include the cost and capacity of digital storage; the complexities of archiving and preserving website content and video; finding ways to automate processes; maintaining a technical infrastructure to support long-term digital preservation; the difficulty of preserving a wide variety of formats; and keeping up with ever-developing standards and new technology. Archives are also continually challenged by the lack of funding to maintain skilled staff.

To meet these challenges, staff at the MFC made copies of transcripts as PDF files along with printed copies on acid free paper. The digitized sound files had already been copied onto tape, so we also preserved the tape copies. We placed all of the completed digital files on hard drives and transferred copies to the University of Maine's Fogler Library computer server, with a backup copy at University of Maine at Fort Kent. These files are regularly migrated and checked for degradation. However, if all of these preservations fail, we can always go to the Library of Congress to retrieve the originals. Clearly, there are limitations to any of the methods we are using for preservation, but we hope the redundancies will overcome any weak links in the preservation plan. Still, as the MFC is the repository for most audio recordings from the state's historical societies and libraries, it is essential that we follow best practices with multiple redundancies for preserving the historical record. To this end, we created a digital preservation policy that dictates how often and by whom the digital files need to be checked and their care reevaluated.

DIGITAL ACCESS

In addition to preservation, we are also working to provide greater access to the collection. The primary work of creating accession descriptions and abstracts of the interviews in our collection is time consuming and ongoing. As we complete each description, it is put in an online database that can be viewed and searched on our website (http://umaine.edu/folklife/archives/collections/).

To increase awareness of and access to the collection, the MFC, with funding from the National

Endowment for the Arts, developed an online exhibit for the University of Maine's Digital Commons called "The Maine Song and Story Sampler."[1] The exhibit presents audio and video recordings of traditional songs and stories as MP3s and MPEG4s, JPGs of associated photographs with accompanying transcriptions of the text and musical notation as PDFs and contextual information—thus creating a global audience for these unique traditions. Visitors to the site can click on a Google map of northeastern North America, find their Maine or Canadian Maritime community, and click on a link to a song or story collected in their community. Each song and story includes contextual information: the name of the singer and collector, date and place of collection and other interpretive information. MFC staff and students listened to various recordings and chose selections with interesting content and good sound quality. (See sidebar on next page for an example.)

The variety of songs and stories chosen reflect the diverse occupations and ethnicity of the people, as well as a range of both inland and coastal communities, large cities and small towns. Teachers interested in using the materials in the classroom will also find a free downloadable curriculum.

PERMISSIONS/COPYRIGHTS

In selecting materials for the sampler, MFC staff attempted to contact interviewers and interviewees for permission wherever possible. By law, the creator of any object (text, audio file, or photograph) holds the copyright unless and until the creator turns that right over to someone else. Audio interviews belong to both the interviewer and the interviewee. Although materials in our collection are accompanied by a release form that essentially turns copyright over to the MFC, the form does not explicitly give permission to put the materials on the World Wide Web. We followed the spirit of the guidelines of Fair Use Copyright Law to the best of our ability. The purpose of the exhibit was educational, and no fees are involved for interaction with the exhibit. We avoided using works by commercial artists and made selections short. We also included information about the artists and researchers involved and provided bibliographic information for users. Because we were unable to find people to grant permission in every instance, we placed a statement on the site stating that we would take materials down if copyright holders objected to

the resources being placed there. So far, no one has requested that materials be removed. Anyone creating a documentary or even a website using archival photographs or audio and visual materials needs to be aware of copyright law. The law states, "the distinction between what is fair use and what is infringement in a particular case will not always be clear or easily defined. There is no specific number of words, lines, or notes that may safely be taken without permission. Acknowledging the source of the copyrighted material does not substitute for obtaining permission" (http://www.copyright.gov /fls/fl102.html).

DIGITAL DICTIONARY

Another digital humanities project at the MFC is a collaborative project with Native programs, the Penobscot Nation, and the American Philosophical Society to complete a Penobscot language dictionary. The current *Penobscot Language Dictionary* consists of a 500-page dot-matrix-printed manuscript of Penobscot words with English translations. The original manuscript is the work of researcher Dr. Frank T. Siebert and his assistants. They began creating the dictionary for the Penobscot Nation during the 1980s by writing words on index cards and then alphabetizing the cards and typing the words from the cards on sheets of paper in alphabetical order. Towards the later part of the 1980s, project staff began entering the dictionary words on a computer and saving the information on a computer hard drive with backups on 5¼-inch floppy discs. Upon Siebert's death, the manuscript and disks went to the American Philosophical Society in Philadelphia as part of Siebert's collection.

At the request of members of the Penobscot Nation's language revitalization group in 2012, MFC staff approached the National Endowment for the Humanities for a grant to complete the dictionary. The current project seeks to make both an online database version of the dictionary for the Penobscot Nation and a final published print version. Working with a part of the dictionary that was digitally entered onto floppy disks in the 1980s, the American Philosophical Society contracted with George Blood, an audio/video/film preservation lab in Philadelphia, to transfer the digital files onto current media. Then the project's linguist, Conor Quinn, transferred those digital files into an XML-structured digital manuscript, with every original component separately tagged.

EXAMPLE FROM THE MAINE SONG AND STORY SAMPLER

Title: Bye-Bye Longjohns

Narrator: Jim Cahill and Dot Ruppell

Location: Bingham, ME

Document Type: Song

Collector: Jeff McKeen

Date: 1991

NA: 2245

CD: 2172

Abstract

"Bye-Bye Longjohns" is a musical representation of how most Mainers feel by the time March rolls around. For some, this feeling comes even earlier. The song was written in western Maine during the late twentieth century.

Description

"Bye-Bye Longjohns" is a musical representation of how most Mainers feel by the time March rolls around. For some, this feeling comes even earlier. The song was written in western Maine during the late twentieth century. The song began as one verse by the nephew of Grange member Myrtle McKinney, who was also the official musician in Norridgewock, ME. After singing the short song, Jim Cahill asked Dottie Abbott, a resident of The Forks and member of the Grange in Bingham, to write additional verses. The fact that the song has multiple authors likely explains the change in refrain from the first to second verses. Though even more verses may have been written since, the three verse song heard here was the full version as of 1991. Also note that the song follows the tune of "Bye-bye Blackbird."

The Patrons of Husbandry was founded by Oliver Hudson Kelly in 1867 as a secret society of agriculturists concerned with education, economic cooperation, political lobbying, and fraternal association. The first Grange in Maine was established in Hampden in 1873, and by 1887 the state had the largest Grange membership in the nation. The chief function of the Grange in Maine has always been social – to improve the quality of life for farm families. From the very beginning, music played a central role in Grange activities. Recent Grange musicians retained many older styles of music and traditional entertainment. From the "official" Grange piano music to songfests and harmonica tunes, the music of an earlier age is still heard in Grange halls throughout the state.

Transcription

1. I put them on October 1, that was orders from my hon, How long, longjohns? They'll keep me warm all winter long, I want to tell you in this song, How long, longjohns? They were closer to me than a friend, next year to Sears again I'll send; I kind of miss this underwear, for several months we were a pair, Longjohns, bye-bye!

2. They got smelly towards the end, it was even hard to bend, Bye-bye longjohns! I bought red but they turned black, I should really take them back, Bye-bye longjohns! I must admit that I was glad to shed them, but for months they really warmed my aft end; There were times I thought I'd freeze, especially when I felt a breeze, Longjohns, bye-bye!

3. I'd have shed my underwear, I don't care I'll go bare, Bye-bye longjohns! They were very close to me, they tickled me, tee hee hee, Bye-bye longjohns! If you see them you'll know where to find me, how I miss that old trap door behind me; I have shed my underwear but I don't care I'll go bare, Longjohns bye-bye!

Keywords: Jim Cahill, Dot Ruppell, Jeff McKeen, Bingham, Maine, Grange, Patrons of Husbandry, Myrtle McKinney, Norridgewock, Dottie Abbott, The Forks, longjohns, folksong, Bye-bye Blackbird

Disciplines: Ethnomusicology | Folklore | Oral History

Recommended Citation

Cahill, Jim and Dot Ruppell. 1991. "Bye-Bye Longjohns." NA2245, CD2172.13. Maine Folklife Center, University of Maine.

XML is a free open software that can be read by both humans and machines; it also supports Unicode, a universal character set that can be used in any human language. This makes XML democratic and usable by anyone with access to the Internet. It ensures interoperability—a concept important to archivists and librarians who must transfer data from older media to newer media. The use of XML presents a model for creating dictionaries without expensive software programs or extensive programming knowledge. Since most indigenous languages are preserved by a small number of people with few resources, this method is an important contribution to dictionary making for marginalized peoples.

Currently Quinn, Pauleena MacDougall, and several student assistants are correcting errors and entering missing lexical items, with the goal of having a robust and complete digital manuscript of the dictionary. Having the materials in an XML document make editing easier, as much can be automated while keeping all formatting out of the database. Eventually the digital manuscript will be placed in a design program for final publication by the University of Maine Press. In the meantime, any number of versions of the dictionary and sub-dictionaries (place name index, for example) can be easily constructed. Versions of the dictionary are preserved in stages as the process moves forward. This will allow for back-checking in case of errors and also will provide other scholars with a record of the process.

In parallel to the dictionary work, Quinn has also collaborated with language teachers for the Penobscot Nation's Department of Cultural and Historic Preservation in building a Penobscot language curriculum. In consultation with the Listuguj Educational Directorate's Mi'gmaq language program, Quinn developed critical design principles for language teaching. In particular, Quinn took a minimalist approach by establishing a small set of core vocabulary to illustrate basic grammar—grounding the introduction of both with extensive visual reinforcement. All of this dictionary work has all been accomplished rather rapidly because we can work in a digital realm and can quickly and easily correspond via email with all of the principals involved. Collaborative work of this kind is only possible because of widespread access to the Internet and is illustrative of the potential for digital humanities teaching and scholarship.

Our experiences with digital humanities illustrate a range of ways that humanities scholars work in the digital world. Creating more accessible versions of previously undigitized materials, preserving and conserving digital humanities resources, creating new digital resources (as in the case of the dictionary), and sharing scholarship across the Internet using new and better software tools are just some of the ways that the digital world is changing the humanities. ✍

ENDNOTE

1. The Maine Song and Story Sampler may be found on UMaine's Digital Common's site (http://digitalcommons .library.umaine.edu/) by entering "Maine Song and Story Sampler" in the search field.

REFERENCES

Kirschenbaum, Matthew G. 2010. "What Is Digital Humanities and What's It Doing in English Departments?" ADE Bulletin 150:1–7.

Pauleena MacDougall is director of the Maine Folklife Center and a faculty associate in anthropology at the University of Maine. MacDougall edits the Maine Folklife Center's annual monograph series, *Northeast Folklore*, and is currently working with staff in collaboration with the Penobscot Indian Nation in preparing the *Penobscot Indian Language Dictionary* for publication by the University of Maine Press in 2016.

Katrina Wynn has served as the archives manager of the Maine Folklife Center since 2012 and teaches the "Introduction to Folklore" course at the University of Maine. She is working on a graduate certificate in digital curation from UMaine, a program she pursued to aid her in her digital duties at the MFC.