

**Author(s):** Graham Hukill, 2011

**License:** Unless otherwise noted, this material is made available under the terms of the **Creative Commons Attribution 3.0 License**: <http://creativecommons.org/licenses/by/3.0/>

**We have reviewed this material** in accordance with U.S. Copyright Law and have tried to maximize your ability to use, share, and adapt it.

Copyright holders of content included in this material should contact [open.michigan@umich.edu](mailto:open.michigan@umich.edu) with any questions, corrections, or clarification regarding the use of content.

For more information about **how to attribute** these materials visit <http://open.umich.edu/education/about/terms-of-use>.

Graham Hukill  
School of Information – IMLS Winter 2011 Internship  
In Collaboration with Open.Michigan &  
the Program in American Culture

## **Final Report**

### **Background and Overview**

The idea for this digital preservation project was born between Pieter Kleymeer from Open.Michigan and Bobby Glushko from the University Library's mPublishing division. Both had previously worked with Professor Margaret Noori, Director of the Comprehensive Studies Program and Lecturer in American Studies, and in speaking with her, had come to appreciate the considerable amount of materials related to her studies, from published texts and notes, to visual and audio recordings, of Ojibwe language and culture. In the early planning stages of the project, before specifics had been outlined, there was talk of conducting an inventory of what kind of materials existed, what might be good candidates for digitization, and begin to look at, inspired in large part by Open.Michigan's involvement, what could be turned into useful and meaningful online educational resources. The evolution of the project's goals and scope from these early planning stages is important to map, as it illustrates some of the challenges we faced setting up a digital collection of language related audio materials.

Open.Michigan is "a University of Michigan initiative that enables faculty, students, and others to share their educational resources and research with the global learning community." Soon celebrating their 3<sup>rd</sup> anniversary, Open.Michigan has hit the ground running and is working actively with faculty across many disciplines to get their materials online, accessible, and free of copyright concerns. Materials they reformat and host are usually organized by particular courses, which create the umbrella collection for a variety of different materials and types. So while Professor Noori's materials are clearly of great scholarly value, and everyone involved believed would make excellent educational language resources, there was not a "course" to intellectually start grouping these files together. Nor was there much precedent within Open.Michigan of approaching faculty with topical interests as opposed to course specific. This was an interesting wrinkle in the project, and would influence the rest of the work and nature of the project itself.

The original project scope was ambitious. Working collaboratively with Open.Michigan, mPublishing, and faculty from the American Culture Studies department, to paraphrase the original proposal, we hoped to evaluate Professor Noori's collection of materials and complete a preservation assessment, create a processing/organization plan, make recommendations for metadata, and assess what significant properties of the materials need to be preserved to make them future renderable and useful as online education resources.

Though we had envisioned an assessment of materials Professor Noori had at the time, we realized early on the value in focusing on a particular set of materials; we could branch out from there if time allowed. We chose to focus on some audio recordings that had already been digitized by the Language Resource Center, and see about getting them formatted, described, and deposited into a digital repository. This was a pivotal point, as the decision moved us away from

assessing and cataloging Professor Noori's materials, effectively ending the involvement of mPublishing and the notion of preparing materials for print publication.

Pieter Kleymeer, as my mentor for the project, was nothing short of dedicated and enthusiastic throughout. He was sensitive to the balance between preservation and access of digitized materials, while also honoring Open.Michigan's interests in the project as well, with the goal of creating an established path to Professor Noori's materials by alleviating them of copyright concerns, and getting them online, described, and accessible. Because Open.Michigan does not host files in-house, we began looking into other file hosting and collection management options. This was also a pivotal point in the project, as it became clear the involvement of a few distinct groups, all with different requirements, familiarity with the materials, and technological expertise, would add complexity to finding a solution that worked for everyone.

## Research and Process

An article by Erway and Shaffner entitled, *Shifting Gears: Gearing Up to Get Into the Flow*, ended up inspiring some real strides towards the tail end of the project. In retrospect, it would have been wise to more actively consider questions it raises about digitization projects and workflows, particularly in the early stages. The article asks: to what extent we should optimize procedures for preservation and access vs. *primarily* for access, and what is the appropriate level of metadata? Section 2 of the article, "Selection has Already been Done", suggests that sometimes what is currently available, and where time has already been spent, should be the first materials you look to for digitization and preservation. We were still wavering between focusing on cataloging what she had in her possession, and once that was accomplished, where and how it should be preserved vs. a smaller "pilot" group of materials. It was around this time that we learned of the digitized and partially described recordings the LRC had created. We identified this as a great group of materials to start with, and would eventually, at least at this stage in the project, comprise the totality of the materials we have formed and deposited. As we realized it made sense to refine the scope of the project to these files only, the next step was deciding how and where these files would reside, where a collection would be born.

An *ad hoc* research phase of the project was spent looking into cataloging and content management systems that might aid in both organizing currently known materials *and* ingesting future found or created ones. It touched on an interesting chemistry when dealing with an endangered language, where preservation of materials does not directly correlate to preservation of the language itself. We researched and evaluated some metadata schemas, weighing these considerations against what schemas potential repositories were employing. And maybe most challenging, was working to balance these considerations among parties invested in the project. We identified a fundamental obstacle: this was the problem of finding a repository, and a workflow, that would accept large initial deposits, until the majority of currently held materials were accounted for, but would *then* allow for simple and ongoing deposits of individual, or small batches, of materials to maintain and build the collection. The initial files from Professor Noori that we would describe, preserve, and offer access to, are certainly valuable in their own right, but it was establishing a well defined route to preservation and access that we felt might be the greatest outcome of the project as a whole.

## Outcome

After meeting with James Ottaviani from the University's Deep Blue digital file repository a couple times, we decided on it as the primary repository for our pilot group of files. We worked to clean up the metadata and organize the actual files. We also looked into standards for the 15 primary Dublin Core fields, compared to Open.Michigan's use of these same Dublin Core fields, and migrated the existing metadata into a new spreadsheet for depositing. The collection, after some back and forth about collections names and minor edits, can be found under the "Research Collections" / "Anishinaabe Language & Literature", with the enduring URL: <http://deepblue.lib.umich.edu/handle/2027.42/83716>. There are approximately 95 unique recordings, each comprised of an AIFF file for preservation and an MP3 file for access and dissemination, associated metadata for them all within the Deep Blue database (some more detailed than others, where depth of description will be more robust in future deposits), and a Attribution-ShareAlike Creative Commons license ("BY-SA", <http://creativecommons.org/licenses/by-sa/3.0/>), reviewed by both Open.Michigan and Professor Noori.

Most of what is described above represents phases in the project, some of the challenges we faced as we went along. But in addition to and often influencing these logistical decisions were some very interesting contextual and abstract considerations. A good example is the Ojibwe.net website that currently hosts material from Professor Noori. It had been assumed that preservation and depositing of these recordings into an online collection would only augment the website, and not detract from it. But Professor Noori, in one of our early meetings, explained the time and effort they had put into the website in an attempt to contextualize the recordings, something that not only might be lost in a repository like Deep Blue or the Internet Archive, but might even undermine the website itself. It was difficult, but ultimately illuminating, to account for how preservation of digital materials – often copies available for access – can complicate the life and value of those same materials in other forms and locations. This unique property of digital materials, particularly when dealing with contextually sensitive topics such as Native American culture or endangered languages, can certainly be challenging, as was encountered over the course of the project. Even selecting a collection name in Deep Blue presented complexity, balancing the imposed structure of a database versus the connotations of a discrete name.

While this pilot group is only a fraction of what we ultimately hope to digitally preserve and offer access to as open and online educational resources, it nonetheless represents a considerable amount of research, careful consideration of access versus preservation, the delicacy and context in which materials of an endangered language should be handled by, and paves the way for a robust and ongoing collection. The project is ongoing, with a substantial amount of recordings that did not make it into this first deposit slated to be deposited over the course of the 2011 summer, with additional help describing and translating some of the materials with the help of someone fluent in the language. Professor Noori, in her work, studies, and teaching, is producing an extremely valuable body of work that deserves proper care and attention so that it may endure, and that those interested will have continued and unfettered access. While there is still much to do, we have laid the groundwork for a valuable, dynamic, and enduring digital collection of Anishinaabemowin language and literature resources.