Digitizing Manuscripts in Trying Times
Presented 9 May 2004 by Heather E. Ward
39th International Congress on Medieval Studies at Western Michigan
University, Kalamazoo, MI

[Slide #1] The University of Oregon Libraries do not have the breadth of

medieval manuscript holdings that some larger libraries have at such universities as UC

Berkeley and Columbia. Although we have some generous donors, strong development

offices, and unique manuscripts we do not have that immediate recognition that larger

collections have. The state of Oregon has also had steadily declining revenue for a

number of years. So, it's necessary for us to pursue creative solutions and make a strong

case for digitizing our collection where larger institutions might find funding more

readily available. By tying a venture, such as this first manuscript digitization project, to

faculty projects outside the library, we are demonstrating our determination and

willingness to cooperate both to the university administration and to potential donors

outside the university. Future expansion of the project will be more attractive to those

donors because we will also have proven that we have the necessary skills and knowledge

and that we are capable of making the project relevant to current university teaching and

research needs.

It is crucial to reach out to other departments on campus and to share the technical

expertise we have developed. Budgets are tight and we cannot afford to reinvent the

wheel each time a different department takes on a similar project. This is also important

to the grant writing process. Both the library and department will benefit by collaborating

on grant requests. Library faculty and staff should encourage teaching faculty members to

talk with them before completing a grant request for a project that could involve library

staff and collections. This way the proposal can take into account existing facilities and expertise and the resulting funds can support the actual needs of faculty members and the library. This is one of the many areas where communication among university programs and the library is vital to the success of a project.

[Slide #2] At the University of Oregon, we have developed particularly strong collegial relationships between the Library, the Medieval Studies Program and the Feminist Humanities Project through a variety of channels including the Special Collections & University Archives department, the medieval studies library liaison—that's, me—and the Metadata and Digital Library Services Department. A history of collaborative projects provided a strong foundation on which to build our digitization efforts. Since I had worked with both Medieval Studies faculty and the Digital Library Group, I felt that I was in a good position to act as a liaison between the two. So, I began making inquiries into the status of our medieval manuscripts and plans to digitize them.

I had been interested in the digitization of medieval manuscripts since graduate school and I was curious to see what both the Library and the medieval studies faculty wanted out of such a project. We have a small collection of medieval manuscripts in our Special Collections & University Archives Department and from time to time medieval studies faculty have asked about the possibility of digitizing them to use in teaching and research. A graduate student had previously created an online catalogue of our Burgess Collection that included descriptions and a few images with an emphasis on bindings, decoration, and scripts but not entire pages or manuscripts. **This is an example of some of the images from that project. [Slide #3]** No one had yet undertaken a large-scale digitization effort of our medieval manuscript holdings. I began to speak to potential

stakeholders in the library and around campus about related projects. We don't always know about every project undertaken by the different departments and programs and I wanted to be sure to talk to the people who would be most interested in such a project and who would have relevant skills and knowledge that we could draw upon.

In my fact-finding mission I learned that UO Feminist Humanities Project (FHP) team members had visited Oxford's Bodleian Library and the Wellcome Trust Library in London in the summer of 2001. There they selected manuscripts that could be digitized, added to their database and eventually incorporated into digital teaching units in their specific subject areas.[i] **Here you see the home screen for one of these units that uses a manuscript from the Bodleian to illustrate "the Medicine of Gender." [Slide #4]** I also found that in 2002 the Head of UO Special Collections & University Archives had traveled to Mexico to work with FHP faculty on a project creating digital interactive versions of historic Latin American maps.[ii] **This slide shows how once a quadrant of the map has been blown up, text can be selected to show a transcription and the buttons at the bottom allow scholars to add their own comments. [Slide #5]** Although the UO Libraries' manuscript collection would require a different research approach than previous projects due to its relatively small size and scattered coverage, FHP scholars were very interested in digitizing our own collection, as well. This would be an excellent chance to pursue a project of long-time faculty interest and to tie it to an already established, exciting and well-respected program.

[Slide #6] During this time, the libraries had been pursuing digital projects of their own, as well. In August 1997 the UO Libraries established a Digital Library Initiative (DLI). Members selected hardware and software and established guidelines for projects

throughout the libraries. In 2002, Carol Hixson, the Head of Metadata and Digital Library

Services, authored DLI's *Access to Digital Collections: Best Practices*, a document

covering such issues as the definition of access, metadata standards and storage, and

software tools.[iii] In the introduction, Hixson emphasizes some of the primary concerns of

the Library:

> The digital context raises some access issues that had not been a concern for print
> collections and the traditional means of providing physical and bibliographic
> access to them. Among these are the issues of persistent access and
> interoperability. There is also the issue of authorized use that is handled in print
> collections by limiting physical access to resources or restricting circulation
> privileges.

[Slide #7] Hixson identifies the metadata provided in order "to facilitate the discovery,

identification, and retrieval of the digital object" as another key component of an online

collection. Metadata for a digitized manuscript might include a title, subject headings,

information about its origin, how it was produced both physically and digitally, the

standards used in its virtual creation, and other descriptive details about the manuscript

and its location.

So far, the UO Libraries have applied these standards to projects focusing mainly

on photographs, such as "Picturing the Cayuse, Walla Walla, & Umatilla Tribes," images

of Native Americans from the late 19th and early 20th centuries, and the "Western Waters

Digital Library: The Columbia River Basin in Oregon," digitized maps, texts and

photographs of regional history. Both projects were in collaboration with other cultural

and educational organizations.[iv] Digitizing medieval manuscripts provided us with the

opportunity to branch out and learn about materials in a different medium with a different

set of needs.

The Library has a record of collaborating with organizations both inside and out of the University of Oregon. It has partnered with the Feminist Humanities Project and its sponsoring organization, UO's Center for the Study of Women in Society, from the beginning. So, it was a logical arrangement to bring our mutual interests and experience together to work on the digitization of medieval manuscripts.

Here I'll put in my disclaimer. For clarification, I am not the person doing the physical work of digitization or making the decisions regarding resolution and other technical details. So, when I provide you with detailed specs, this is the work of our Image Services Center staff. I am part of the group that created the guidelines for such projects and I may help later with the creation of metadata content. We have experts on *standards* in Metadata and Digital Library Services who are far more knowledgeable than I about this area. I do have a continuing interest in studying scholars' use of this and other manuscript digitization projects which I hope to pursue further once we have established a more robust online collection. My main role so far has been that of liaison between the groups involved, and now chronicler for the project.

After meetings with various interested groups, we began actively discussing the project that had been considered in a variety of ways over the years. In the end, the group chose one manuscript to use as a prototype—book IV of the 1490 *Geste de Garin de Monglane*, also known as the Cheltenham mss. It is one of the more unique texts that the Library owns and would provide the chance to work out a system for digitizing other manuscripts in their various incarnations. **These are some images from the test files that I saved as JPEGs. You can see that the text is readable even as a JPEG. The original TIFFs are of higher quality, however. [Slide #8]** The content doesn't

necessarily touch on feminist topics, which are of particular interest to the FHP, but could be used more broadly for medieval research and teaching. [Slide #9] One professor has subsequently chosen to tie the mss to a sabbatical project. This gives the group even more incentive to move ahead.

The library had reached the right moment in our collective skills, knowledge and technological resources that we were ready to undertake such a project. We needed to research issues specific to the digitization of medieval manuscripts, but had a strong base to build on. Stakeholders agreed on a plan for the larger project once we got into production mode. Library staff would digitize selected manuscripts and would create records within ContentDM. This would provide access to images of different sizes and resolution and metadata describing the images. We would provide separate files to the Feminist Humanities Project to be manipulated and incorporated into digital teaching units. Library and FHP staff would work together on the database interface and metadata with the idea of allowing widely scattered scholars to work interactively with each other and with the material—possibly like the Mapas Project.

Special Collections and Image Services Center staff considered a number of factors when deciding on a means of digitizing the manuscript. Staff members decided to use cameras and copystands for the project. They worked out a system and a timeline based on experience, experimentation, budget constraints, other projects on their docket, and the physical state of the manuscript. Because the current budget did not allow for the new equipment needed for the preferred method, Lesli Larson, Image Services Center Supervisor, created a plan for the preliminary project using current resources. Once the

prototype is complete, we hope to find the resources to purchase new equipment for the next phase of the project.

During that next phase, Image Services Center staff hope to create 4"x5" transparencies that can be subsequently digitized. These transparencies would have a longer life and higher resolution than images created with the digital camera. They would have greater preservation value due to their analog form and could actually be consulted directly with the use of a low-tech light box. So, even if the digital copy degraded, we wouldn't have to cause more wear on the original manuscript in order to make another image. However, with our current resources, the estimated price of $17 per page is prohibitive.[v]

[Slide #10] The digitization plan for the prototype employs a Nikon D100 6.3 megapixel digital camera and a two light 250-watt halogen lamp copystand with an approximate focal length of 75mm using polarizing filters. The high-resolution TIFF files will be adjusted in Photoshop on a Mac G4 for color, density, contrast and sharpness. The end files of approximately 51MB each can be resized or resampled if needed.[vi] These images will be of high enough quality to be legible in order to facilitate paleographic work. JPG files will be created from the TIFFs to display in the ContentDM database for faster loading and the TIFF files will be kept as archival copies with more limited access. The drawbacks of this plan are that it provides no analog surrogate and therefore has a more limited preservation value.

As with any interdisciplinary/interdepartmental project, there can be challenges as well. Good communication is extremely important to keep the project on track. Responsibilities should be clearly delineated, so everyone knows where they should focus

their energies. Awareness of and consideration for each other's time and budget constraints can help things go more smoothly. Be honest about limitations when planning, so everyone will have reasonable expectations. Try creating a prototype to help estimate staff and funding commitments to a larger project. This will also help ensure that the final product will be worth the effort—long lasting, relatively low maintenance, and useful to its intended audience. Be open to new ideas and ways of doing things. Respect each other's knowledge and experience. In the end, when we create a product that provides long-term benefits to a wide audience, it's worth the effort.

[Slide #11] My advice to other faculty members thinking about such a project is to **consider the expertise already existing on your campus.** Ask if someone in the library has experience with digitization or creating metadata. Maybe someone in computer science or another department has already done something similar. Special Collections staff could point you to other related projects even if they haven't created their own. Look at other digital projects—who created them? What methods did they use? Set up a time to discuss technical and funding issues with the experts. Think about subject expertise among library and teaching faculty—who else might be interested in cooperating on this project? Are there particularly good graduate students who would be interested in working for credit or experience? Consider who might be willing to fund such an undertaking—talk to campus grant offices. Check out library resources on grants. If your institution subscribes to it, sign up on the Community of Science, a database of interdisciplinary grant funding resources, to see what other funding might be available.[vii] Check out other projects outside of your university—don't limit your search to just medieval manuscripts, but consider digital libraries broadly. Some, such as the Digital

Scriptorium, provide detailed information on their process that you could adapt to your needs. Don't reinvent the wheel. Contact the creators of these other projects to ask for updated information. Equipment and technology change so rapidly that something on the cutting edge only a year ago may have become nearly obsolete today.

When deciding on the appropriate hardware and software, consider the equipment that your own institution or other regional institutions are using. Your campus may already have a license to use certain software saving your project money. If you have worked together on collaborative projects in the past or have the potential to do so, using the same software will allow you to use the same formats, so you won't have to worry as much about interoperability issues. If another institution has already vetted a product out, they can offer you troubleshooting advice and warn you of mistakes they may have made early on. Look for products that are easily upgradeable, so that you will be able to maintain your images and continue to provide access to them.

[Slide #12] **Remember that the time and money used to purchase equipment** and to create digital images are not the only expenses related to such a project. Metadata needs to be created at the same time as, or soon after, the creation of the images in order to provide strong access to your collection. It may seem like creating a digital surrogate for a manuscript and placing it on the web is easier than providing access in-house, but digital images actually require more detailed indexing rather than less. An in-house finding aid or catalogue can provide a description of the overall manuscript, but online collections require details for each individual image to provide the best access to them.

A digital surrogate may help to preserve the physical manuscript by providing online access. If this is your intention, you need to create images of high enough

resolution that they meet the needs of most researchers and you need to plan for the necessary amount of virtual storage space for larger images. These projects may also draw more people in to see the collection in-person and thereby actually cause more wear on the originals. Survey researchers early on to find out how they plan to use such a collection. Think about how you want to maintain control over the images. If your institution depends on income from the use of its images, you may want to consider using digital watermarks or providing lower resolution images openly and higher ones for a fee or in-house. Another vital issue is to consider how you will maintain this digital collection over time and who will be responsible for this maintenance. This is a large investment and you want to be sure it's going to last.

The different views and knowledge that each of us brings to such a project provide context for something that could otherwise be just a bunch of pretty pictures. It allows professors to highlight their scholarship and enriches the content that the library provides to other researchers. Although we work in trying times, collaboration helps us to create a digital collection we can all be proud of.

[Slide #13] The web sites listed here are some of those mentioned in my paper. I have also included one from the National Information Standards Organization (NISO), created "for people who…want to develop good digital collections…and…for funding organizations and agencies that want to encourage the creation of good digital collections." *Le Medieviste et L'ordinateur* **is an online publication of the Institut de Recherche et d'Histoire des Textes and has a number of articles of interest to medievalists using different technologies.**

i Judith Musick, "CSWS and Knight Library Collaborations Build Virtual Collections," From the Center: CSWS Newsletter Spring 2003: 6-7. http://csws.uoregon.edu/newsletter/pdf/spring_2003.pdf

[ii] Judith Musick, "The Feminist Humanities Project Does England," From the Center: CSWS Newsletter Autumn 2001: 6-7 http://csws.uoregon.edu/newsletter/pdf/autumn_2001.pdf
[iii] http://libweb.uoregon.edu/diglib/AccesstoDigitalCollections_Rev.html
[iv] http://libweb.uoregon.edu/catdept/digcol/index.html
[v] Discussion with Lesli Larson and e-mail messages among group members.
[vi] Image specs 2000x3308 pixels, 300 dpi, 6.667"x10.027"
[vii] http://www.cos.com/