

When Just Doing It Isn't Enough: The University of Oregon Takes Stock

by Carol Hixson

Background

The University of Oregon Libraries, like many libraries and other cultural heritage institutions, began to create and provide access to a variety of digital content about a decade ago. Some of the earliest efforts began in Special Collections & University Archives with the digitization of portions of the library's holdings of more than 400,000 rare photographs. Scanning and metadata collection were done with little awareness of, much less adherence to, emerging standards. Digital files were stored on a variety of media, with no systematic backups, no checks for file integrity or media degradation, and little internal consistency in file naming, creation of digital images, or collection of metadata. In the space of a few years, some files were no longer usable, due to degradation of storage media or inadequate collection and tracking of metadata. Other efforts, such as the E-Asia Digital Library [1] or digital collections of aerial photographs, were developed with an awareness of and attempt to follow standards but with little institutional buy-in or support.

In 2000-2002, the Libraries took steps to address some of these problems by pulling together a group of stakeholders to study digital library issues and needs.

A number of results came out of this Digital Library Initiative [2], including:

- Guidelines for the design and use of a mass storage unit for digital content.
- Standards for providing access to digital collections.
- A decision to purchase the CONTENTdm® digital asset management system.
- Formation of a Metadata Implementation Group [3] (MIG) and a Software Implementation Group (later combined into one group) to develop

metadata standards for digital collections and to implement one or more collections using the CONTENTdm® system.

- The assignment of a group of functional experts, named the Digital Collections Gatekeepers, to monitor digital collection activity, make policy decisions and recommendations, and keep Library Administration apprised of needs and developments.

In late 2002, the MIG began to meet and hammer out agreements regarding descriptive, technical, and administrative metadata for a collection of photographs that had previously been digitized. We all became conversant in Dublin Core, we reviewed standards and practices being developed and utilized worldwide, and we tested the software. We planned to migrate the digital content into the CONTENTdm® system, rescan some content to ensure it met our standards for digital images, and apply the newly approved metadata standards to this collection.

In March 2003, our plans took a very different direction. In what was to become a characteristic abandonment of fledgling plans in favor of a quick response to a critical need, the library identified a completely different collection of materials for digitization and mounting in CONTENTdm®. We created and made this collection publicly available in record time (four months) to fulfill our obligations to a granting agency. We achieved this by calling on the expertise of four librarians in what was then the Catalog Department, working closely with staff in Special Collections & University Archives (especially the Image Services Center), and recruiting staff of the Catalog Department to learn how to scan glass plate negatives, apply descriptive and technical metadata, and load the digital objects and accompanying metadata into the CONTENTdm® system. While this was underway, the library also established an institutional repository, called Scholars' Bank [4], using DSpace software. I was involved in both efforts as the chair or co-chair of the groups working to implement the systems.

The die was cast. In December 2003, the Catalog Department was rechristened Metadata and Digital Library Services (MDLS) [5] in recognition of its expanded role of implementing and maintaining digital collections –in addition to cataloging and preserving analog materials. We are responsible for the digitization of items, collection of metadata, upkeep of the software, coordination with and marketing to internal and external partners, and developing and maintaining the public interfaces to the collections. In the summer of 2004, we absorbed the staff of the Image Services Center, with whom we had been working very closely, into MDLS. However, within the library, no new classified or professional positions have been added –all work on digital collections is an add-on to other ongoing work.

I have been concerned for some time about the sustainability of our model for building and preserving digital collections. In the spring of 2004, one of my colleagues and I requested and received funding from our library to attend the weeklong Digital Preservation Management workshop [6] presented by Cornell University Library. In July 2004, he and I were part of a group of twenty-four people taking part in the Cornell workshop. Attending the workshop served three main purposes: first, it confirmed our fears that our model of digital collection development and maintenance was not sustainable; secondly, it gave us a clear understanding of why it wasn't sustainable; and thirdly, it provided practical tools for developing a sustainable program and becoming a trusted digital repository. [7]

The Cornell workshop presented digital preservation needs as the three legs of a three-legged stool, all of which are needed to keep the stool balanced: Organizational Framework (policies and planning), Technological Framework (hardware, software, technical skills), and Resources Framework (staffing, monetary, space)[8]. The Cornell model identified key indicators of the stages of

development for each leg, ranging from stage one to stage five. During the course of the workshop, my colleague and I independently completed an exercise to assess our institutional readiness for digital preservation. We both identified the University of Oregon Libraries as being at stage one regarding the Organizational Framework, stage two on the Technological Framework, and stage one on the Resources Framework.

Stage one development in terms of organizational infrastructure is often characterized by nonexistent or implicit policies. Key indicators of stage two technological infrastructure are being project-specific and reactive. Being at stage one in the resources framework indicates that an organization's resources devoted to digital preservation are generally low, finite, and ad-hoc. In addition to helping us identify problem areas, the workshop also provided a number of tools to enable us to move forward to the next stages of development.

What have we done with our knowledge in the intervening fifteen months?

Organizational and technological frameworks

Recognizing that we could have little short-term impact on the resources devoted to preserving digital collections, we decided to focus our immediate attention on improving the organizational and technological frameworks as much as possible. Immediately after returning from the workshop, I arranged to meet with our Library Administration to outline the challenges and propose some strategies for addressing them. I prepared a two-page document for them: laying out a statement of need, defining digital preservation, listing the attributes of a trusted digital repository, and proposing a series of next steps. [9]

In order to get administrative buy-in to this approach, I suggested that we could undertake this effort with no additional outlay of staffing or financial resources –

and with no reduction in our existing workloads. Naïve as this may sound, this is actually a defining characteristic of the way we do things at Oregon: we don't let our lack of resources stand in the way of moving forward on issues that we identify as priorities; we just do it.

I began by calling together our Digital Collections Gatekeepers group. We met on August 20, 2004 –the first meeting of the group since it was named in the fall of 2002. Prior to that meeting, my colleague, Corey Harper, and I prepared some background documents and suggested some reading for group members. We also worked with a member of the Systems Department and revised the guidelines for the mass storage unit (MSU) [10] that had been drafted initially as part of the Digital Library Initiative in 2000-2002.

At the first meeting of the group, we discussed the overall framework for preserving digital content, using the Digital Preservation Snapshot document that Corey and I had prepared. The group agreed on what the next steps needed to be. We also discussed what we meant by digital content and agreed to define it broadly to include any digital content that we acquire, license, or create and that we want to retain or preserve. There was agreement that we wanted to look at including digital preservation strategies (such as LOCKSS)[11] in our licensing agreements and to use the OAIS Reference Model [12] to help us evaluate vendors from whom we license or buy content, as well as using it for self-evaluation.

We also discussed the role of the Gatekeepers and other individuals and units within the library. We agreed to hold open meetings once a month, as well as to conduct business between meetings via an archived discussion list. At this first meeting we also looked at digital collections that were already in progress. This review provided a clear indication of how quickly collections can proliferate and how important it was for this group to become involved in helping to set a realistic

framework for new projects. We renamed the group the Digital Content Coordinators (DCC). The Digital Content Coordinators were subsequently charged by Library Administration to 1) develop a robust digital preservation strategy for the UO Libraries and work with Library Administration and digital content providers to implement it; 2) discuss, review, and make decisions about policies and guidelines; 3) make recommendations to Library Administration about the resources needed to support digital collections (e.g., labor, equipment, software, etc.); 4) serve as advisers to other members of the library or campus community in the planning and creation of digital collections; and 5) make recommendations to Library Administration about which proposals it strongly supports, supports or considers not ready.

In subsequent meetings, we began to discuss the policy framework, using the Cornell *Action Plan for Developing a Digital Preservation Program* [13] provided during the workshop. We managed to work our way through the first several sections of the *Action Plan* before our attention was diverted to other matters.

Since beginning our work in August 2004, we have [14]:

- Developed a checklist of information needed for a digital content survey.
- Developed a digital content survey questionnaire and a contextual document for it, laid out the costs and benefits, and tested the survey instrument on an existing digital collection.
- Reviewed and made preliminary recommendations for file naming conventions.
- Revised and approved the mass storage unit (MSU) guidelines and migrated existing projects into the MSU structure.
- Developed a preservation policy statement for digital resources.
- Drafted a mission statement that is now awaiting administrative approval

- Developed a standardized workflow for acquired digital content, outlining roles and responsibilities, and attempted to handle several newly licensed collections using the workflow.
- Reviewed and begun to document our backup procedures for digital content.
- Begun to explore the possibility of setting up offsite storage of a copy of our backups with a sister institution in the region.
- Kept apprised of new developments in digital preservation.

In June 2005, I attended the fourth DELOS International Summer School on Digital Preservation in Digital Libraries [15] in Sophia Antipolis, France. Anne Kenney, one of the trainers and creators of the Cornell workshop, spoke for half a day on the principles of digital preservation. It was extremely useful for me to be able to review Cornell's model again and take stock of the progress we had made (or failed to make) at the University of Oregon in the eleven months that had passed since I attended the Cornell workshop. When I filled out the survey on our institutional readiness this time, it seemed to me that we had made some progress. On the Organizational Framework, I rated us at stage two, rather than stage one. There was even some evidence that we might be moving on to stage three. The key indicators of stage two development in the Organizational Framework are general policies and planning with increased evidence of institutional commitment. At stage three, there are basic and essential policies in place. Our progress in this area has been the result of the work of the Libraries' Digital Content Coordinators. Regarding the technological framework, I rated us between stages two and three, the improvement being due to the fact that we have made progress away from being purely project-specific and reactive to doing more assessment of our technology needs and becoming more proactive. This is due, again, to the efforts of the Digital Content Coordinators.

Resources Framework

In June 2005, when I rated the UO Libraries on the third leg of the digital preservation stool, the Resources Framework, I concluded that we were still firmly entrenched at stage one: generally low, finite, and ad-hoc financial commitment. This is true with regard to staffing, monetary, and space resources. The Oregon model of “just do it” is about to do us in. Since the DCC began its work in August 2004, the digital content managed through CONTENTdm® has increased 1452% and the digital content managed through Dspace has increased 1119%. There is no record of how much digital content created and managed by other groups or individuals in the library using other software systems has increased in this same time period. There are, as yet, no formal business plans for any of these collections. Some small amounts of outside funding have been secured, enabling us to hire a few temporary staff for short periods to assist with some of this growth. However, the additional staffing does not approach what would be needed to plan for, create, and manage the substantial increase in digital content that we have experienced. Requests for new hardware and software to keep pace with the needs for new types of digital content to be delivered ever more quickly go into the same pool for consideration as the needs to upgrade staff and public workstations, to purchase new system packages for controlling licensed electronic resources, or for acquiring new servers for all of the library’s computing needs. There is never enough money to meet all of our needs and we frequently find ourselves robbing Peter to pay Paul. Our ability to make steady progress on the goals and objectives that we identified as being necessary to become a trusted digital repository is continually hampered by a resources framework that is insufficient for the task. We are also victims of our own success. Our colleagues’ and our users’ appetites for more digital content, with ever more sophisticated user interfaces, have outstripped our ability to create and manage the content well. In such an environment, we are hard pressed to plan adequately and to manage the content appropriately. It has

become increasingly difficult just to follow the standards and policies that we have already developed.

Clearly, the next step must be to develop business plans and impose project planning on our digital collections work. Reflecting on our efforts and slight progress since attending the Cornell workshop, I am reminded of a slide that Anne Kenney and Nancy McGovern presented when they first addressed the five organizational stages of digital preservation. In discussing the need to acknowledge that digital preservation is a local concern, they showed a slide that said, "Hello, my name is Cornell and I'm a digital imaging junkie." Paraphrasing the slide, the University of Oregon Libraries must acknowledge our need to devote adequate resources to the task at hand and say, "Hello, my name is UO and we can no longer 'just do it'."

Citations

[1] E-Asia Digital Library is available at: <http://e-asia.uoregon.edu/>

[2] A history of the University of Oregon Libraries' Digital Library Initiative can be found at: <http://libweb.uoregon.edu/diglib/>

[3] <http://libweb.uoregon.edu/catdept/meta/metahome.html>

[4] <https://scholarsbank.uoregon.edu/dspace/index.jsp>

[5] The department's web site is available at:

<http://libweb.uoregon.edu/catdept/home/>

[6] <http://www.library.cornell.edu/iris/dpworkshop/>

[7] RLG/OCLC Working Group on Digital Archive Attributes, 2002, [*Trusted Digital Repositories: Attributes and Responsibilities*](#)

[8] For more information on Cornell's model, see: Anne R. Kenney and Ellie Buckley, "Developing Digital Preservation Programs: the Cornell Survey of Institutional Readiness, 2003-2005," *RLG DigiNews*, August 15, 2005
http://www.rlg.org/en/page.php?Page_ID=20744#article0

[9] A slightly revised version of this document is available as: Carol Hixson and Corey Harper, 2004, *Digital Preservation Snapshot*
<http://darkwing.uoregon.edu/~jqj/diglib/archive/msg00551.html>

[10] University of Oregon Libraries, 2004, *Mass Storage Guidelines*
<http://libweb.uoregon.edu/diglib/mass-storage-guidelines.html>

[11] <http://lockss.stanford.edu/>

[12] Consultative Committee for Space Data Systems, 2002, *Reference Model for an Open Archival Information System (OAIS) – ISO 14721*
<http://www.ccsds.org/documents/650x0b1.pdf>

[13] Cornell University Libraries, 2004, *Action Plan for Developing a Digital Preservation Program*.

[14] All policy and working documents are available on the group's web site at:
<http://libweb.uoregon.edu/diglib/digcon.html>

[15] <http://www.dpc.delos.info/>

