

2013

If It's Not Powered Up, It's Rotting: Transforming Archival Collections from Obsolete Platforms and CD-Rs to an Institutional Repository and Storage Array

Todd Bruns, *Eastern Illinois University*

Stacey Knight-Davis, *Eastern Illinois University*

If It's Not Powered Up, It's Rotting: Transforming Archival Collections from Obsolete Platforms and CD-Rs to an Institutional Repository and Storage Array

Todd Bruns
Eastern Illinois University
600 Lincoln Avenue
Charleston, IL 61920
tabruns@eiu.edu

Stacey Knight-Davis
Eastern Illinois University
600 Lincoln Avenue
Charleston, IL 61920
slknight@eiu.edu

ABSTRACT

Digital collections projects began at Eastern Illinois University in the mid 1990s. These early projects had specialized interfaces and different storage strategies. With the inception of an institutional repository (IR) in 2011, these collections were brought under a common interface and storage practices. The benefits and challenges of utilizing an IR for the dissemination of legacy digital projects is discussed, along with data storage solutions appropriate for a mid-sized academic library.

Categories and Subject Descriptors

H.3.7 [Information Storage and Retrieval]: Digital Libraries – collection, dissemination, systems issues.

General Terms

Management, Performance, Reliability.

Keywords

Institutional repositories, bit-rot, digital preservation, legacy collections.

INTRODUCTION

Eastern Illinois University is a mid-sized public comprehensive university located in Charleston, Illinois. The University is home to approximately 10,000 students and offers bachelor's, master's, and specialist's degrees. Until the very beginning of the 1990s Eastern Illinois University (EIU) did not have an official University Archives. Upon arriving on campus, new Dean of Library Services Dr. Allen Lanham sought to rectify this oversight and established University Archives and Special Collections in Booth Library. Nearly a decade later, Dr. Lanham,

intent that EIU should refrain from playing "catch up" again, sought to establish an institutional repository for the university. The repository would not however be a repository in the typical sense – while faculty research and scholarship would be an important part of the repository, Dr. Lanham sought to utilize the repository as a new off-shoot of University Archives: A digital University Archives.

In 2011, with the hiring of an Institutional Repository Librarian, Booth Library partnered with Berkeley Electronic Press (bepress) and began the design and implementation of the EIU institutional repository, The Keep (<http://thekeep.eiu.edu>). The selection of the Digital Commons platform from bepress for the repository was predicated on a number of factors: the desire to focus more on content recruitment than internal technology support, the stability of the Digital Commons platform (a significant impact factor for discoverability – patrons can't find and cite your materials if your repository is often down for repairs), and both the commonality and successful track record of the use of the Digital Commons platform by mid-sized academic libraries. [1]

The rapid digitization of University Archives materials also required some investment in setting up a scanning center. Equipment and staff were procured for the creation of a Scanning Center in Booth Library, a new service point that would serve to digitize items for The Keep, whether they be archives materials or faculty print collections. Three high-speed scanners were purchased, and funding was provided for approximately seventy hours per week of student help. Two civil service staff were repurposed for half-time supervisory duties over the students, each of them reporting to the Institutional Repository Librarian, and digitization of box after box of University Archives materials began in earnest. Early collections in The Keep included meeting minutes of faculty and staff governance committees, heavily used books and bulletins about the history of EIU, and historical press releases.

"WE'RE NOT DOING THAT AGAIN"

The unfortunately late establishment of a print University Archives at EIU has created a reality of being behind in archives curation and collection. The Dean of Library Services, determined to avoid a similar issue in the future, focused intently on two goals: one, establish an institutional repository for EIU, in keeping with similar efforts that were already underway at peer institutions like Illinois Wesleyan and Southern Illinois University; and two, use the repository to get ahead of the digitization curve, recognizing that more and more documents of record and historical interest were being "born digital," blurring previously

established lines that separate a university archives collection from an institutional repository. [2] A platform was required not only for ease of access but also simply because more archival documents were now being produced digitally.

A common thread in early discussions about the establishment of the institutional repository was to avoid a Johnny-come-lately situation as had happened with the University Archives. “We’re not doing that again,” was the repeated mantra, reinforcing that the repository must be used to help University Archives begin to leap ahead, at least digitally. Whereas many institutional repositories were established primarily for the showcasing of an institution’s scholarship [3], The Keep began primarily as a platform for digitizing content from University Archives. Scholarship would of course be a focus of the repository, but unlike some repositories, the repository at EIU would focus more heavily on many different facets of EIU life: faculty and student scholarship, archival records, historical events, student life, exhibits and special collections, and also collections not produced directly by EIU faculty or students but to which permission had been given to collect, curate, digitize, and make accessible.

The potentially overlapping roles of an institutional repository, digital collections, and university archives can be confusing and may require an institutionally defined setting of boundaries. At EIU, a blurring of boundaries was deemed acceptable due to the specific benefits each platform afforded. For example, compiling an archive of photographs into digital collections onto the library website allows patrons browsing the site to view the collection. Yet, also housing the collection in the repository introduces a higher level of accessibility for Google and other search engine end users. If the content is not born digital, keeping the print in University Archives is another potential form of preservation. Indeed, some collections exist in all three formats.

Early collection activity involved pulling archived meeting minutes from several campus governance groups, digitizing them in the Scanning Center, and posting them in The Keep. Although these documents would seem to be primarily of internal interest, the IR Librarian noted with interest that particular meeting minutes actually saw significant download counts. [4] Another early collection activity was the cloning of the EIU digitized master’s theses collection.

CLONING FOR ACCESSIBILITY AND PRESERVATION

At EIU, master’s theses are regularly digitized when they are submitted to the library. Prior to the creation of the EIU institutional repository, master’s theses were already accessible digitally via the EIU library catalog and WorldCat. Collecting the master’s theses into the repository served three purposes.

First, it provided The Keep with a content “jump start,” pulling in born digital content that increased repository content and visibility by hundreds of documents within its first month of production. This proved to be crucial for early efforts at faculty recruitment. Faculty responded positively both to the repository as a “viable” platform – already having nearly 1,000 documents – and also to the download counts that were already being generated due to the master’s theses.

Second, cloning the master’s theses into the repository greatly improved the visibility and access to the collection. Although the

theses had been digitized and made available beyond EIU via WorldCat, much greater exposure was now possible thanks to the metadata fields and search engine optimization of the repository platform. An example of the increase in accessibility can be seen in the download statistics of *Improvements to Vertical Axis Wind Turbine Blades to Aid in Self-starting*, a Master of Science thesis in the EIU School of Technology by Joseph P. Tillman. In October 2011, under EIU and WorldCat catalog access, the thesis was downloaded thirty-five times, thirteen of those during a classroom demo. By comparison, in October 2012 the thesis was downloaded 729 times from the institutional repository. The original cataloging metadata proved invaluable in the ease of batch uploading these documents to The Keep, and provided additional access points for web crawlers.

Third, the repository master’s thesis collection results in second copies of the theses on servers outside of the EIU campus. A catastrophic server loss at EIU now no longer threatens the entire digital collection of master’s theses, as the same collection is cloned and kept on Berkeley Electronic Press servers. The master’s theses collection can be viewed at <http://thekeep.eiu.edu/theses/>.

The Eastern Illinois University student yearbook, the *Warbler*, was also cloned. Digitization of the first eleven volumes (1919-1927) was completed in 2006. This project was funded by a grant awarded by the Illinois State Library (ISL), a Division of the Office of the Illinois Secretary of State, using funds provided by the Institute of Museum and Library Services (IMLS), under the federal Library Services and Technology Act (LSTA). As a condition of the grant, the completed project files were uploaded to the Illinois Digital Archive (IDA). IDA is a CONTENTdm installation provided as a service of the Illinois State Library. CONTENTdm is a software package produced by OCLC. Similar to Digital Commons, it provides access and management utilities for digital collections. Inclusion in IDA allows users to easily search for content and view the material in the large context of Illinois history, but the material is separate from the larger body of digitized Eastern Illinois University materials. The files for these early *Warblers* were copied to The Keep, allowing them to be used and searched along with thousands of other documents related to the history of EIU. Newer *Warblers* were digitized and added to the institutional repository as well.

The benefits of moving the *Warbler* collection to the repository were similar to those of moving the master’s theses. In the repository, the *Warbler* collection benefited from increased visibility and access, embedded social media utilities, and preservation via multiple access copies. Additionally, the EIU repository allows assertion of intellectual property rights, with an automatic assignment of the Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 license. Adding the collection to the repository also served to increase the gravitas of The Keep as an archive of historical documents of importance to the institution. The *Warbler* collection can be viewed at <http://thekeep.eiu.edu/warbler/>.

A related digital collection was the EIU Laboratory School Yearbooks (http://thekeep.eiu.edu/labschool_yearbooks/). From 1899 to 1974 EIU operated a laboratory school on campus, so that student teachers could gain experience working in a real classroom. The repository platform provides a “Paper of the Day” link, which the Institutional Repository Librarian used to promote the Laboratory School Yearbook collection during a 2012 on-

campus reunion of graduates of the Laboratory School. Providing easy digital access to the yearbooks helped to make the reunion a success and was a vital component to establishing important institution-community connections.

The cloning of the master's theses and *Warbler* collections to the repository was a relatively easy process and an early success. Other collection efforts proved to be more difficult but also more vital – whereas the master's theses collection was already in a stable digital format and platform, other collections existed on platforms that had been built on rapidly outdated coding or stored in formats prone to data loss.

MIGRATING FROM UNSTABLE PLATFORMS

An unseen benefit of the repository was that it gave us a platform for migrating content that existed on unwieldy and difficult to update legacy web applications. Two collections required immediate migration. Both of them were built using Microsoft Access and ASP Classic. As early pilot projects, these applications were intended to provide a test bed to gauge interest in library-provided digital content. The success of the projects exposed the limits of the underlying technology. Although Access was familiar to support staff working with the database, it is not designed to support large numbers of users. The end of support for ASP Classic and the increasing difficulty of keeping legacy applications running on newer servers was a strong incentive to migrate the content in The EIU University Archives Photo Collection and the *Journal of Collective Bargaining in the Academy*.

The EIU University Photo Archives Photo Collection began in 2003. With only paper finding aids and the memory of the University Archivist available, there was no automated retrieval system or even a complete paper index for locating photographs relevant to EIU's history. End users, including other University departments, were interested in the ability to search and retrieve these images. In order to provide a search system for Archives staff and an access method for patrons, the Photo Archive collection was created.

The creation of the collection depended on changes in technology and library facilities. From 1999-2002, Library personnel and collections were relocated in temporary facilities while the library building underwent renovation. With new facilities, University Archives gained new workspaces and network connectivity. This change in facilities happened as the price of flatbed scanners dropped and their speed increased. With the availability of space and equipment, digitizing key portions of the unique Archives photo collection became feasible. File name conventions based on the date the item was digitized were developed, which provided a unique identifier for each image, and an Access database held image metadata. Since the system was developed to assist Archives staff, metadata information entered mirrored the needs of the staff. Named buildings, places, and people, as well as key words were entered, as well as the date the photo was taken if it could be determined. A web interface to the Access database was also developed to provide a public access point.

The migration of the Photo Archive to the repository required generating a content report from the current database and transforming that information into a batch upload spreadsheet for uploading into the repository. This process allowed for the

checking and correcting of erroneous information on the original files, resulting in a stronger collection. Along with the benefits that the master's theses collection enjoyed – increased visibility, access, and preservation via multiple access copies – the Photo Archive in the repository was easier to browse. In The Keep, photos are arranged by broad topical area and can be displayed as a slide show or as a gallery with thumb nails. This is a great improvement over the original which required a search to display results. Another advantage of the Image Gallery in The Keep is that it offers the ability to post images from the collection directly to several popular social networking sites. The embedded social linking utilities make it easier for users to share the content, further increasing its visibility. The Archives Image Collections can be viewed at http://thekeep.eiu.edu/archives_images/.

The library became involved in electronic journal publishing with the hosting of a platform for the *Journal of Collective Bargaining and the Academy*. This included hosting the proceedings of the annual conference of the publishing organization, the National Center for the Study of Collective Bargaining in Higher Education and the Professions. Although the website provided some visibility for the *Journal*, overall use and discoverability was not optimal. Maintaining the site required a fair level of technical skill and *Journal* editors were dependent on library staff for adding and editing content. Journal editors and reviewers used email to exchange manuscripts and accept submissions.

Migrating the *Journal of Collective Bargaining in the Academy* and the proceedings proved to be a painstaking process of collecting the full text of varied content (articles, presentations, handouts, manuscripts, and schedules) from server files while matching them up to the appropriate author(s). In the case of the proceedings, this included replicating the conference schedule while correctly matching materials to presentations.

Hosting the *Journal of Collective Bargaining in the Academy* (<http://thekeep.eiu.edu/jcba/>) in The Keep opened up an entirely new section of the repository. Under the current bepress agreement, The Keep is able to host five e-journals, with each journal able to have its own interface design and with built-in editor functions that make editing the journal a more streamlined process. These features provided additional functionality that wasn't possible with the original platform, giving added value to the journal editors. Similarly, the search engine optimization provided article authors with increased visibility and downloads. Based on reports from Google Analytics and The Keep statistics, downloads increased by a factor of ten, from 336 in 2011 (under the previous platform) to 3,309 in 2012 when hosted in The Keep.

Utilizing the repository as a replacement for the ASP-coded website introduced new functionality for the editors and a much more stable platform for the journal. Accessibility and preservation were both improved due to the migration. Hosting the *Journal of Collective Bargaining in the Academy* also allowed EIU to shine as an example of partnering with the broader community, putting one of the key purposes of the repository – functioning as a showcase for marketing and fundraising purposes – front and center.

The materials collected for the *Journal of Collective Bargaining in the Academy* were all born digital, thus the main challenge was in collecting and correctly identifying the files. In other cases, collections migrated to the repository came from stored media where bit-rot had already caused damage.

ROTTING AWAY ON DISC

A serious challenge to digital libraries collections is the rapid loss of data and knowledge due to faulty formats and bit-rot. People from all walks of life have dealt with the fading of magnetized images on VCR tape, unreadable floppy discs, and the loss of data on CD-R, not to mention the obsolescence of formerly “hot” hardware. Laser discs are now nearly as quaint as 78-RPM vinyl records.

The Keep is intended to be the digital archives of the university, and as such, materials collections coming out of departments are just as important as the collected scholarly works of faculty and students. One of the major projects of The Keep’s first year was the creation of the EIU Department of Theatre Art’s theatrical productions collection.

The Theatre Arts department maintains a webpage called “Memory Lane” that lists theatrical productions going back to the 1940s. While the page includes some photos, the collection is largely incomplete, lacking programs, reviews, and additional photos. The incompleteness comes from the fact that many of the photographs and other materials associated with the productions exist in scattered locations and formats. Some exist on servers in Booth Library, some on the hard drive of the EIU photographer’s desk top computer, many are in print in file folders in Theatre Arts department file cabinets and still more are stored as images in drawers full of CD-Rs.

A clear benefit for the creation of the Theatre Arts Productions collection (http://thekeep.eiu.edu/theatre_productions/) in The Keep was the compiling of materials from all these disparate sources. Print materials were digitized in the Scanning Center and uploaded to Booth Library servers. Image files on other parts of Booth Library servers were accessed, as well as image files on the EIU photographer’s desktop computer. Most of this compiling was fairly straight forward and a matter of simply getting access. The problem was in the discs.

Although the Theatre Arts department had only been storing images on CD-R for a little over ten years, there was already significant data loss. Some of the discs could not be read at all. Some required older equipment in order for the images to be retrieved. During the migration effort (still underway), Booth Library staff did what they could to retrieve images and combine them with the other materials sources.

The new Theatre Arts Productions community in The Keep thus combines photographs, programs, newsletter, and posters, all compiled from divergent sources, into concise collections of documents per each theatrical production. As with other collections in the repository, the items are more accessible than before and are far more secure in terms of preservation. Images that previous resided only on a desktop computer can now be Googled and discovered from anywhere in the world. Photographs that were slowly decaying in a file drawer on CD-R are now preserved via an access copy in The Keep as well as in master file TIFF format on library servers.

The preservation of the master files remains a challenge. While The Keep provides preservation of access copies (and introduces, in a very limited sense, the philosophy of Lots Of Copies Keeps Stuff Safe), the digital preservation of master files relies on the data redundancy within a storage array and off-site backups.

MAINTAINING DIGITAL MASTERS

Up to this point, we have discussed only web-ready JPEG and PDF files associated with the digital collections that were incorporated into The Keep. In addition to updating and improving user interfaces and discoverability, the challenge of finding a way to store the TIFF master images generated by previous and ongoing library digitization projects had to be addressed. Master TIFF files were saved during the scanning process for each image in the *Warbler* collection, Photo Archive collection, and all other archival materials scanned for The Keep.

Digitization of the *Warbler* was carried out following the Illinois Digital Archives Guidelines for Images, which specify resolution and file format. Collections created after the *Warbler* follow the standards of the Booth Library Digital Collections Manual produced by the Booth Library Collection Development Committee Subcommittee on Digital Resources. The Both Library Manual is in compliance with the Creation of Digital Collections: Digitization Best Practices for Images and Text developed by the Consortium of Academic and Research Libraries in Illinois (CARLI) Digital Collections Users’ Group. [5]

The Booth Library Digital Collections Manual specifies file formats for use copies and digital master images. Following these guidelines from the outset of digitization projects results in data normalization practices being in place as the images are produced. In compliance with the CARLI Digitization Best Practices, digital masters are saved in TIFF format. TIFF provides very good support for high-resolution images and the format continues to be one of the preferred formats for bitmapped images. [6]

For the first few years of the Photo Archive project, as space on library servers was consumed, the large TIFF files were burned to CDROM in the hope that affordable and stable long-term storage options would become available in the future. Storing large master files on CD was not an uncommon practice, and institutions such as the British Library also stored TIFF images on CD. [7] However, the drawbacks of this storage system quickly became apparent. Retrieving the TIFF images from disc was cumbersome when these files were needed to fill patron requests for print-quality images. Re-burning the discs periodically to mitigate decay and manually checking the discs for integrity were time-consuming tasks that quickly proved burdensome to add to existing staff workloads. While the Optical Storage Technology Association estimated the life span of recordable compact discs at 50-100 years, [8] it was also noted that the life span of a recorded disc would be extremely difficult to estimate reliably. As was seen in the CD-Rs used by Theatre Arts, the life span for CD-Rs proved to be much shorter than 50 years. Even under controlled conditions in University Archives, damage has been discovered on some of the discs used for Photo Archive masters. Images display with discolored areas or cannot be opened at all. Since the photographs from the early years of the Photo Archive project are still available in the Archives, these images can be re-digitized.

With the beginning of the *Warbler* project, the decision was made to maintain digital masters on a server instead of on disc. Space was allotted on existing machines for digital master image storage. While allocating space on existing file servers allowed for automatic error detection within files, more space, and more drives, were needed. Adding drives requires a system for distributing information across the disks in the system. Management and distribution is accomplished with a RAID array. RAID refers to a storage system composed of multiple disks with

data copied and divided over several disks. By design, a RAID array holds at least two copies of each file.

In 2007 an Apple Xserve RAID was purchased for storing digital master images. The RAID was in service until 2012 when it was replaced by HP array system. While the storage array does not eliminate the possibility of data loss, it greatly decreases the odds of non-recoverable damage. The manufacturer specified mean time before failure for individual drives in the system is 1,200,000 hours, or 136 years. Obviously, this far exceeds the expected life span of other components in the system. Calculated probability non-recoverable of data loss in 7 years is 0.1342%. [9] Having digital masters on the storage array provides for continual error checking and multiple copies within the RAID. For our system, two copies of the mater files are spread across 22 disks. The RAID controller continually checks the data for errors. On a weekly basis, empty space in the system is checked for errors and if bad sectors are discovered on the disk they are excluded from future use. For further security, files are backed up to another location outside of the library.

Having multiple copies in multiple locations follows the digital preservation strategy of data redundancy. This strategy has become widely accepted. This security comes with added costs. In his discussion of the limits of the practical limits of digital preservation, Mike Kastellec states “multiple copies necessarily increases expenses, therefore—given equal levels of funding—less information can be preserved redundantly than can be preserved without such measures.” [10] Thus factors of cost, security, and the amount of data to be stored must be balanced. Cost has become the overriding factor limiting the growth of digital collections. Maintaining multiple copies and off-site backup requires not only equipment funds, but also skilled IT support. As the amount of information requiring long-term storage increases the complexity of the hardware required to do so also increases.

One cost containment strategy is to limit the amount of material to be stored. The Booth Library Digital Collections Manual specifies evaluation procedures that are to be carried out before beginning a digitization project. The significance, audience, uniqueness, and the scope and size of the collection to be digitized must be established before work begins. All legal and technical issues, including storage costs, must also be considered. These steps, taken before any new digital collection is begun, ensure that the collection will be of value to users and of sufficient worth to justify resources allocated to it.

Another cost containment strategy is collaboration. Networks currently exist for the shared storage of use copies. Stanford University’s LOCKSS network offers private networks for maintaining content in a distributed environment. [12] While this system works very well for preserving web-ready content, the massive amount of data involved with digital masters comes with a very high cost in bandwidth usage to store in a network system. Persuading stakeholders to invest heavily in equipment that will be used by non-stakeholders, even if it is in the best interest of all parties, can also be problematic.

Although technology provides continually improving options for data security, demand for digital data also increases while funding support remains stagnant. Rosenthal aptly describes the problem: “Like almost all engineering problems, bit preservation is fundamentally a question of budgets.” [13] Working within this

constraint, we strive to provide good security and room for growth.

FUTURE PROJECTS

In 1998 Booth Library took on hosting responsibilities for the campus newspaper, *The Daily Eastern News*. This collection was delivered by the Journalism department and consisted of PDF files beginning with editions from 1993. Documents could be retrieved by date and page number when the collection was first launched. In 2005, a Google Mini Search Appliance was purchased to allow keyword searching of the content. While keyword searching was an improvement, there are still some usability issues. Each page was saved as a separate PDF file, so articles that continue on a second page are difficult to navigate. When users search the *Daily Eastern News*, they are generally interested in finding all information available on a particular event or person. There is currently no way to search the *Daily Eastern News* in conjunction with other sources.

As it stands, the *Daily Eastern News* receives about 300 hits per month. The collection is currently about 15 GB of material in PDF format. Since the *Daily Eastern News* is the record of daily happenings at the University, it is of great interest to students, faculty, alumni, and community residents. Due to its value to the community, and the usability issues noted above, the *Daily Eastern News* will be the next collection migrated to The Keep. This will require combining all pages of individual issues into a single file. Once this material is in The Keep, users will have a more complete picture of the history of Eastern Illinois University.

The migration of *The Daily Eastern News* will require some data massaging. This important collection was handed off to Booth Library staff in fragmented condition – some documents in pdf, some in HTML, some as one document, some as separate documents per page. Since newspaper functions as a paper of record for historical events at the university, preserving the content in a stable form is essential.

The *Daily Eastern News* will be collected into a unified master file format, preserved using the RAID method, and presented as an access collection via The Keep. Part of the collection process for The Keep will be Adobe full text recognition on the content, which will allow scholars to search for news articles without having to browse individual editions looking for content. Analysis of The Keep using Google Analytics has shown that the repository page created for the *Daily Eastern News* is already receiving significant traffic, even though there isn’t yet content there. This demonstrates a strong community interest in the collection, making the migration of the *Daily Eastern News* a priority.

CONCLUSION

For EIU, the implementation of an institutional repository has been more about leaping ahead in digital archives preservation than in collecting scholarly output, although it must be said that the archives content has served to stoke, not stall, faculty participation. In 2012, its first full year of operation, The Keep grew in content by over 10,000 documents, compared to the average first-year growth among repositories of 731 documents. [Berkeley Electronic Press, personal communication] The repository provided a significantly improved platform in terms of stability and preservation for a number of important collections at

EIU, and in some cases, provided a timely migration of data and content from failing digital storage environments.

The utilization of The Keep as a digital archives will continue into the foreseeable future, along with continued growth in faculty participation and digital publishing initiatives. As noted by Mahyar Izadi, the EIU Dean of the College of Business, The Keep represents “building a second [digital] library for campus.” With dedicated attention to improvements in digital preservation methods, it is hoped that the EIU digital archives will be preserved well into the future.

REFERENCES

- [1] Lynch, C. A. and Lippincott, J. K. Institutional repository deployment in the United States as of early 2005. *D-Lib Magazine*, 11(9):1.
- [2] Johnson, R. K. Institutional repositories: Partnering with faculty to enhance scholarly communication. *D-Lib Magazine*, 8(11). Retrieved February 1, 2013, from *D-Lib*: <http://www.dlib.org/dlib/november02/johnson/11johnson.html>
- [3] Shearer, K. Institutional repositories: Towards the identification of critical success factors. *Canadian Journal of Information and Library Science*, 27(3):89-108.
- [4] Bruns, T. One Year In – Top Downloads of Documents from The Keep in the First Year of the Repository. Retrieved February 1, 2013, from The Keep: http://works.bepress.com/todd_bruns/13/
- [5] Corrigan, E. K. Booth Library Digital Collections Manual. Booth Library Collection Development Committee Subcommittee on Digital Resources. Retrieved February 1, 2013, from http://works.bepress.com/ellen_corrigan/9/
- [6] Consortium of Academic and Research Libraries in Illinois (CARLI). Digital Collections Users' Group. (Rev. 2009). Retrieved February 1, 2013, from The Keep: http://www.carli.illinois.edu/mem-prod/contentdm/guidelines_for_digitization.pdf
- [7] National Digital Information Infrastructure and Preservation Program. TIFF, Revision 6.0. Sustainability of Digital Formats Planning for Library of Congress Collections. Retrieved February 1, 2013, from NDIIPP: <http://www.digitalpreservation.gov/formats/fdd/fdd000022.shtml>
- [8] IS10 Potential bit rot in image files that were stored on CD. Retrieved February 1, 2013, from SCAPE: <http://wiki.opf-labs.org/display/SP/IS10+Potential+bit+rot+in+image+files+that+were+stored+on+CD>
- [9] Bennett, Hugh. Understanding CD-R & CD-RW. Optical Storage Technology Association. Retrieved February 1, 2013, from CARLI: http://www.osta.org/technology/pdf/cdr_cdrw.pdf
- [10] Whitehead, J. Calculating mean time to data loss (and probability of silent data corruption). Retrieved February 1, 2013, from Zetta Scalabytes Blog: <http://www.zetta.net/blog/calculating-mean-time-to-data-loss-and-probability-of-silent-data-corruption>
- [11] Kestellec, M. Practical Limits to the Scope of Digital Preservation. *Information Technology And Libraries* 31(2):63-71.
- [12] Global & Private LOCKSS Networks. Retrieved February 1, 2013, from Stanford University: <http://www.lockss.org/community/networks/>
- [13] Rosenthal, D. S. H. Bit preservation: A solved problem? *International Journal of Digital Curation*, 5(1), 134-148.