Developing a Meaningful Digital Self-Archiving Model: Archival Theory vs. Natural Behavior in the Minds of Carolina Research Project

Megan A. Winget
University of Texas at Austin; 1 University Station, D7000, Austin, TX 78712-0390 megan@ischool.utexas.edu

Marisa Ramirez
Arizona State Library, Archives and Public Records, 1700 W. Washington, Ste 300, Phoenix, AZ 85007 mramirez@lib.az.us

This paper will review the findings from a research project concluded in 2004, which had the primary goal of learning more about the natural behaviors of people choosing materials for inclusion in a digital archive. Project participants, retiring faculty at the University of North Carolina at Chapel Hill had a number of tasks to perform. They had to 1) survey and choose the materials to include in the archive; 2) develop a "collection development plan" defining the nature of their materials, their intended audience, and the organizational scheme of their collection; 3) provide materials for digitization; 4) supply metadata for the digitized materials; and 5) write a "finding aid" for their collection. The research team digitized the materials and developed a demonstration repository website for the participants. This paper will focus on the disconnect between traditional archival theory and the users' natural behavior; and the challenges of developing an archival collection given current digital repository software.

Introduction

University faculty members are necessarily a prolific lot. Not only do they produce published work in the form of new scholarship, research reports, and pedagogical material for their classes; they also generate a plethora of unpublished materials in the form of correspondence with colleagues, publishers, and students; and drafts and research materials related to both their published and unpublished work. University archives have traditionally collected these unpublished materials in the form of personal faculty papers, but because there is neither time nor space process and process and store the materials from the majority of University faculty, only the most famous faculty members' papers will be collected and preserved {Tibbo, 2003 (Tibbo, 2003a)}. In addressing the issue that there are not enough archivists to effectively process and represent collections, the Minds of Carolina team identified the concept of self-archiving as a likely answer; to address the concern that there is not enough space to store the resulting collections; we identified digitization as a likely response.

Digitization, however, is not a cure-all. While digitization provides the hope of ubiquitous access (Tibbo, 2003b), the inherent vulnerability and ephemeral nature of digital media and files, the rapid obsolescence of software and hardware systems, and the need for well-constructed file systems and metadata offer significant obstacles in terms of longevity (Rothenberg, 1999) (Bearman, 1999). Seamus Ross has argued that materials already in digital form have little chance of being accessible as soon as 2013 if proactive preservation

measures are not undertaken now and continuously throughout their useful lifespan (Ross, 2000), and there are numerous examples of digitized materials being lost or otherwise non-functional (Higham, 2002) (Bury, 2003).

The Open Archival Information System (OAIS) was developed to address these challenges (Sawyer & Reich, 2003). OAIS is a conceptual framework for developing an archival system that is dedicated to preserving and maintaining access to digital information over the long term. An OAIS type archive must meet six minimum responsibilities: 1) negotiate and accept appropriate information from information producers; 2) obtain sufficient control of information to ensure long-term preservation; 3) determine the scope of the designated community; 4) ensure information is understandable by that community without the assistance of information producers; 5) follow the documented policies and procedures to ensure the information is preserved; 6) make the information available to the designated community (Lavoie, 2000). One publicly available instance of an OAIS repository is MIT's DSpace project (Tansley et al., 2003).

In 2003, due to growing interest and publicity regarding institutional digital repository initiatives like DSpace, retiring scholars began to contact Paul Jones, the director of the public's digital library, http://www.ibiblio.org , asking if that repository would accept digitized materials the faculty had accumulated throughout their careers. It quickly became evident that the project would be more complicated than simply digitizing materials and putting them in the repository. Faced with the complexity of organization and meaningful retrieval of these documents, we hoped that traditional archival theory on issues such as appraisal, arrangement and description, which had served the traditional print archive community so well, might help structure and design a robust, authentic, and evidentially significant contributor-run digital archive.

This discussion led to a pilot project called Minds of Carolina, with goals to develop tools, methodologies, and guidelines to help scholars prepare their materials for archiving within a trusted digital repository. One of the most interesting findings from early discussions and interactions with contributors was the realization that this self-archiving model would have to allow for different patterns of archival representation. The two contributors had vastly different types of materials, expectations, and intended audiences in mind when choosing, organizing and presenting their materials, and had almost entirely dissimilar narrative styles.

Methodology

Our primary concern in this project was to find out how users would choose to archive their own materials "in the wild." We firmly believe that the answer to the problem archival staff constraints is a self-archiving model, and much needs to be done in further studying this area, to explore the ways and means "normal" users go about choosing to archive their own materials (Tibbo, 2003a). There is very little research in this area, and the Minds of Carolina research team attempted to make the process of choosing, organizing and presenting materials as natural as possible, not wanting to push the idea of "right" or "wrong" materials and archival ideals. The first step of the process was to conduct personal interviews with the contributors to assess their project goals, their intended audiences, their materials, and their attitudes regarding archives in general. After the initial interviews, a student worker paired off with one contributor and assisted them with appraisal and organizational decisions. Then, the contributors were given some time to go through their papers and assess the scope of their project within the very minimal bounds set for them. After the contributors had made definite decisions as to which materials should be included in the system, those materials were transferred to the student workers, who digitized and/or submitted the materials into the repository. The final step was to develop an appropriate interface for each contributor.

Project Goals

This project had number of specific goals:

Survey the Nature of the Materials

Two faculty members expressed interest in digitally preserving and archiving work they had compiled over the course of their careers.

The first is a humanist; a translator of Chinese poetry, recently retired from the Curriculum in Asian Studies. The collection he submitted is based on the primary materials related to one book of Chinese poetry translations published in the latter half of his career, and consists of both paper and digital formats including a hard copy draft with evaluative comments from two well-known poets, the book pre-print, daily activity and writing logs, correspondence, unpublished poem translations and original poetry.

The second participant is a scientist, a Professor in the Department of Cell and Molecular Physiology. He took a different approach to his collection. Instead of providing us with traditionally archival material (lab notebooks, correspondence, unpublished manuscripts), he wrote a detailed thematic review of his career, with archival materials (photographs, lab write-ups, unpublished findings) acting as complementary material. This narrative essay, and its corresponding material consist of both physical and digital formats.

One of the noteworthy research questions discovered while interacting with contributors in this first phase of the project included the need for alternate patterns of archival representation. The two contributors had vastly different types of materials, expectations, and intended audiences in mind when choosing, organizing and presenting their materials. They had almost entirely dissimilar narrative styles. The locus of that dissimilarity may be cultural (humanism versus the hard sciences), materials based (discrete units of review - a book and all of its related materials, for example - versus thematic reflections on an entire career), or related to the digital archive's intended primary audience. We will need to conduct further research to 1) ascertain why these differences exist, through more extensive entrance interviews with a greater number of prospective contributors; 2) advance theories as to how these narrative dissimilarities can or should be incorporated into standard digital archival practice; and 3) develop archival-narrative templates to simplify the submission procedure for future contributors and enrich the archival experience for system users.

Develop a demonstration repository website

After interviewing the collection donors, Minds of Carolina project staff assessed the collection and evaluated metadata standards for the project. Some noteworthy questions have surfaced from this phase of the project, notably relating to workflow issues, system architecture, and interface / interaction design.

Workflow issues: (metadata creation, materials digitization, and collection submission) What are the expectations placed on prospective contributors? How much time can we realistically expect faculty members to spend on these archives? How do we ensure complete and accurate metadata when the responsibility of digitally depositing documents passes from contributor to assistant? What procedures can we install that will minimize contributor time and effort, while maximizing the quality of output?

System architecture: Do current digital repository systems - DSpace being the current favorite - utilize archival notions of hierarchy and inter-relatedness of documents in their system architecture - if so, how is it done, if not - should these principles be added to the system? How important are these archival principles, developed for print and analog media, in the digital age? What type of digital repository model will reconcile

the fundamental principles of the archival tradition with the current capabilities of existing digital repository structures? Is DSpace the most favorable system? If not, what functionalities are necessary in terms of our specific needs?

Interface / Interaction Design: In contrast to brick and mortar archives, digital archives can be represented in a myriad of ways - we no longer have the limitations placed on us by working with physical media. If the representation is rich enough, relationships and hierarchies in a digital archive can be dynamically rendered in the interface. However, it is difficult to determine user needs in digital archives because there have as yet been very few relevant studies of user needs for digital archives. This is an area in need of much further study.

Develop Tools & Guidelines to Aid Contributors

After initially meeting with the two project participants, we noted that they had reservations and misunderstandings regarding the nature of archives. They had little understanding of the nature of an archive and its place within a university setting and did not appreciate the importance of depositing their materials in a digital archive. The two issues digital archivists should be able to address when assisting contributors of a digital repository:

What is an Archive?
Namely, what materials are generally deposited into an archive, and how are they most often used? Care should be taken here to not exclude any materials that a contributor might want to deposit - only to provide a general introduction to the idea of archives.

Why my papers?
Although both contributors initially approached team members with an interest in archiving their papers, once we contacted and began working with them, both became hesitant. This was not due to some sense of the work involved (they were very self-motivated), but out of a sense of modesty. Development of a persuasive document outlining the arguments in support of faculty self-archiving their work would be beneficial for further growth of this agenda.

There are three major issues related to contributor preparation of materials: appraisal, metadata generation, and digitization guidelines (if appropriate).

Appraisal: We were working with older faculty, whose papers were primarily paper-based. They had to make some digitization decisions at the outset. These decisions, we found, were based primarily their intended audience. The Minds of Carolina team identified a need to develop documentation to assist contributors with digitization decisions: Which, of the myriad of papers they've stored throughout their careers, is relevant in the realm of personal papers? How do contributors envision their papers' use? How do the papers in their possession, over which the contributor has intellectual control, work together to tell their story?

Metadata Generation: After surveying existing metadata standards, we developed an OAI compliant Dublin Core-based schema for creation of descriptive and accurate discovery-level preservation metadata Further research includes: the use of METS or FRBR standards to automate the inter-relation of documents. In development of the metadata schema we attempted to balance the need of providing context to the documents with the need for faculty ease of use.

Digitization Guidelines: While digitization decisions were largely left to the team, some documents were either born digital or digitized by the faculty. Issues of scan quality and future file migration emerged as items that required future attention, including the development of tutorials complete with examples and procedures for proper digital capture of information.

Develop Policy Statements

Based on analysis of existing policies within the DSpace member federation, there are several components involved in developing a comprehensive repository policy statement. These considerations include:

Repository Definition: This element is central to developing a statement that defines the nature and role of the repository in establishing and implementing an electronic document management system. This serves as a mission statement, and functions as a compass to guide the direction of current and future repository practices including the management, access, preservation, and distribution of digital collections.

Collection Content Guidelines: It is important that a digital archive strikes a balance between the general and specific in instituting collection development policies. The collection development policy should be a master plan that defines the nature of acceptable material within the repository. More specifically, this guideline should address the creation and origination of the material, the nature of the work, the form of the information, acceptable digital formats, rights retained by the repository, and rules and procedures regarding commercial use of repository information.

Communities & Collection Guidelines: Typical digital archives will host a variety of electronic documents within collections from an array of contributors. In anticipation of the myriad of needs that are associated with different digital media, Minds of Carolina must develop collection guidelines that describe the rights and responsibilities of individual contributors, communities of contributors as well as the repository host in regard to submission management. This includes traditional collection policy concepts such as appraisal, access and preservation but also addresses digital management of workflow, interface design, document migration and reorganization, file size limitations and submission withdrawal policies. Nonetheless, given the possible commonalities between collections and individual documents, several questions arise in trying to map the input into making these commonalities public. How do we make connections between contributors, collections, groups of materials and individual documents? What should be the extent of input contributors have into these connections?

Community Start-up Procedures: In starting up a new digital self-archiving community, the research team determined that the best approach is to develop a set of pre-defined steps, especially in regard the host and community. The digital archive is responsible for providing overall collection development policies, but contributors are encouraged to participate in this process by providing information on specific requirements and restrictions on the accessibility of their digital materials. Some questions that will need to be answered include: Can the electronic information be accessible to everyone or to a specific audience? If paper documents are restricted, should they be digitized? If digitized, should we upload the materials even though they will not be accessible? Can the metadata on restricted materials be made accessible?

Distribution License: Minds of Carolina will require the completion of a distribution license before making electronic materials accessible. This license, which requests permission from the contributor to reproduce, preserve and distribute materials submitted to the repository, will help avoid future accessibility obstacles. While this will enable us to preserve digital information through migration and backing up of information, it also serves to limit the type and extent of change that is allowed on a particular document and ensures that copyright laws are observed.

Privacy Policy: Contributor privacy is an important element of the policy statement. A comprehensive policy statement should define the type of personal information collected within the document submission process and should detail the extent to which efforts will be made to disclose or withhold that information. A privacy policy should answer questions like: What is considered “personal data?” Should we “scrub” documents containing such information? Under what conditions can personal information be utilized?

Metadata: DSpace, the host system for this project, collects various types of metadata according to the qualified version of the Dublin Core standard. While DSpace lists the various possible types of metadata that can be collected and recommends basic metadata elements, there are several questions that remain. What if the DC standard is not adequate for a particular document? Can you collect metadata based on metadata standards other than Dublin Core? What types of metadata will contributors want to collect on their materials? What is the feasibility of the standardization of metadata between contributors? What core pieces of metadata should we collect for each submitted document?

Format Support: This final piece of the policy statement will describe the supported and unsupported file

formats within the repository and explain the role of preservation in long-term file use. While this may take the form of a “laundry list”, detailing the recognized format extensions within the repository, this will describe the current support capabilities of the repository and ensure that the materials that are submitted to the repository can be easily migrated and updated for future accessibility.

Discussion and Challenges

Material Selection / Appraisal: The Scientist

The scientist contributor has a three-tiered hierarchy of anticipated audiences: First and foremost, he wants to provide a simple narrative for laypeople who are interested in ALS, more commonly known as “Lou Gherig’s Disease,” a disease that attacks the motor neurons of the host patient. This first group of users, typically family members of people stricken with this illness, are self-learners, up-to-date on the current research, and in need of a straightforward discussion of the complexities of the science. He attempted to do that with his narrative. The second group he identified as a likely audience would be people somehow related to the university as a place of study; they’re either considering coming to the Medical School - he provided notes and slides of some of the more relevant classes he’s taught in his 35 years at as a professor - or these users are simply interested in the university’s institutional history. The final group of people identified as a likely audience would be his peers - other neuroscientists. Instead of providing details on his career, he believes this archive might have value for his peers because he decided to include discussions detailing “failed” experiments and known roadblocks for any future work on motor neurons, which might prove valuable for anyone who wants to perform experiments similar to the ones detailed in his narrative.

This contributor recognizes that although he had a solid, stable career, he never made any “historically significant” findings, and his work would probably not be valuable from a historical perspective. These ideas of “importance” and “historicity” are fundamental to the development of self-archiving systems because in the digital future, creators themselves will have significantly more responsibility for preserving their materials in the early stages of creation, organization and maintenance. The scientist, for example, refused to recognize the relative importance of his lab notebooks to the archive he was building. Lab notebooks are often recognized as the “place to do science,” where scientists record primary lab data and experiment observations. Lab notebooks are also used to record reference to external data and literature sources, and lab notebooks are where interpretations, initial data analyses, and experimental conclusions are first defined (Lysakowski, 1997). The presence of lab notebooks in an archive make it possible for future researchers to reconstruct the real path followed by scientific research work, with its trials and errors, its question marks, and its hesitations (Welfel, 1998). Traditionally, lab notebooks have constituted the bulwark of scientific archives, both public and private. In the public sphere, these notebooks represent a way for scholars to recreate the scientific process. In the private sphere, lab notebooks often represent intellectual property of the scientist, and are valuable from the company’s Research & Development perspective. In many scientific collections, the notebooks are one of the primary foci of the archival collection (c.f., for example, Western Union Telegraph Company Records at the Smithsonian, series 25 (Harding & Oswald, 2001); and the Carl W. Gottschalk Papers at the University of North Carolina at Chapel Hill, series 2.6 (Winget & Andrew, 1999)). This scientific contributor, however, was adamant that lab notebooks are less than useful under even the best of situations: that situation being when the researcher himself is looking at the notebooks, and they were created by someone who was interested in providing useful, well-written information. They are absolutely incomprehensible, he argued, for someone from outside the lab, under any circumstances. He thought that the nominally more useful information would be in the “lab write-up” which was written by the notebook “keeper,” at the conclusion of the experiment. This is not a published report, but is, generally the basis for any further publications. Those he would consider including in

his repository. He would not consider including the notebooks, as he judged their digitization a “waste of time.” He insisted that, except for those experiments that were not published because they were either unsuccessful or had negative findings (and hence difficult to prove and even more difficult to publish) - the relevant and appropriate information was in the published research. For those unsuccessful or negative experiments, he would include his own lab write-up and instructions for dealing with future obstacles within the narrative. When asked if he didn’t consider the lab notebooks of Einstein, for example, to be important, his reply was that Einstein or Galileo or Newton were fundamentally different from “normal” people and those notebooks would be objects that might be significant as objects primarily - something akin, maybe to Leonardo’s notebooks.

Finally, other than the narrative, which could be thought of as a very elaborate collection description, the only other materials traditional archivists would consider “appropriate” for archival deposit would be this contributor’s unpublished photographs of motor neurons. Everything else in his collection is a reference to his published materials. For his published materials, he provided links to the PubMed abstracts, and anchors within the narrative, to contextualize those published findings.

Material Selection / Appraisal: The Humanist

In contrast to the scientist, the Chinese literature professor’s archival selections were guided by the perceived impact of his work from a niche group of literary translators. His selection of materials focused on the documents that he thought other literature professors would find useful or interesting. And students could gain insight into the iterative process involved in the translation of poetry, especially those poems whose meaning depended upon deep understanding of specific historical and cultural nuances.

Unlike the scientist, the translator selected materials that centered on the creation of one publication, a book of translated poetry. The group of documents that were provided to the Minds of Carolina staff included documents that recorded the emergence of the idea for the book, the decision process leading up to the selection and translation of poems, and a pre-print of the completed publication.

Contextualization and documentation of the creative process guided the selection of the materials. While some documents ostensibly had little relevance to the publication, they provided a more complete landscape of the work that was underway during the creation of the publication. He saw tremendous value in brainstorming sessions that not only included relevant literary ideas but also video game scores and personal reminders. Original poetry and other unpublished translations were included in the selected materials because they either developed concurrently with the publication, or evolved from ideas in the publication.

Documenting the creative process of Chinese translation was another driving force behind the selection of materials. Unlike the scientist who chose to provide materials that spanned his career, the humanist’s selections focused on one published document to not only show the iterative process involved in developing a literary work, but also to show the collaborative scholarly process. The humanist provided a pre-print with comments by other well-regarded colleagues in his field, something that he considered of tremendous importance because of the nature and length of the remarks. This provided some fascinating glimpses into what Diane Crane termed an “invisible college”, or rather, the loosely knit scholarly circles to which the humanist belonged (Crane, 1972).

Availability of Materials

Both project contributors, through no prior planning on our part, are nearing their career’s end. This

characteristic had two positive effects. First, they both mentioned that they would be much less likely to submit to this sort of repository at an earlier stage in their careers. In the case of the scientist, the reason was relatively theoretical - this whole process would not have been possible earlier in his career because this entire exercise was a fairly cathartic end-of-career accounting of his lifetime accomplishments. Not only did he want to provide a useful scientific guide for the layperson interested in motor neurons, he also wanted to contextualize research themes and threads that run through different stages in his career. This desire to "look back" would either not exist in a younger scholar, or would at best be incomplete. In the case of the humanist, he noted that early translation attempts could be construed as foolish mistakes ("people would have been able to tell I didn't know what I was doing,"), and he, earlier in his career would understandably not have wanted colleagues or peers to see any such mistakes. At the end of a successful career, it's much easier to accept that one makes mistakes, and those mistakes are not definitive. It's not so easy to accept that truth when one is still striving for tenure and promotion.

These types of statements make one fundamentally reconsider the prospect of digital institutional repositories in general. It is understandable that individuals will not want to ruin their chances of gaining tenure by depositing their materials into a system that might subvert the publishing / promotion / tenure process. But from an archival standpoint, this hesitancy is itself subversive. If faculty are unwilling to deposit early versions of their work (preprints) into an institutional repository- and they're unable or unwilling to preserve these documents themselves either because they don't deem the documents "important enough," or simply due to the complexity or time consuming nature of digital preservation - the kinds of materials available in the future will be haphazard at best. Furthermore, if the faculty don't want to make their work available immediately because of tenure or pride, if the majority of materials are therefore "dark" within a digital institutional repository, what will be the incentive to submit papers?

Currently people give their papers to manuscript collections because those papers will be preserved, they'll be available for future research, and it's a mark of honor to be included in an institution's repository. If, for preservation reasons, these materials must be submitted essentially at the time of creation, and because of copyright concerns might never be viewed, and perhaps not even discoverable, how will the honor of being included in an institutional repository be bestowed? The answer, we believe, after working with these two contributors, is end-of-career faculty - which in itself isn't an enormously groundbreaking idea - people have always submitted work to a repository after the "active" phase of their careers. In terms of digital repositories, in which many institutions are looking for ways to entice participation, however, it might be more "cost-effective" to focus on senior or even retiring faculty - those who have essentially nothing to lose from participation. Focusing on retiring faculty will give the repository time to build momentum along with an aura of impressiveness, which will make future collection development that much easier and will perhaps buy time for the archivists to educate faculty on the relative worth of their seemingly inconsequential papers.

Technical Challenges

Metadata Generation

The second positive effect from using "end of career" contributors is related to their willingness to do a lot of work on the project. They were highly motivated. This quality in a contributor is absolutely necessary, unless the institution wants and/or has the funding to create a department devoted to digitizing collections, submitting materials to the system, and generating metadata. As it is, even with highly motivated faculty, those who sincerely desire inclusion in the system; user-defined metadata is not analogous to professionally assigned metadata, which, although still problematic, tends to be much more controlled, inclusive, and appropriate.

Even with highly motivated contributors who were willing and able to provide robust metadata, an institution would still want professional catalogers or processing archivists to provide quality control.

The Repository Software: DSpace

We chose to use DSpace as the Repository Software for this project, mostly because it was the open source alternative most fully developed (Fedora being the other major option, and it was deemed not complete enough at the beginning of the project). While the development of DSpace is an incredible accomplishment, and should not be diminished, there are some significant obstacles for using DSpace specifically for this sort of archival repository:

Workflow Issues: We identified the need to submit materials in a batch process, with common metadata among all of the objects in the batch, but additionally wanting the ability to assign distinct metadata elements to individual objects within a batch. As far as our programmers have been able to tell, this is not possible in DSpace, which will either allow batches with identical metadata or piece-by-piece submission. When there are hundreds and hundreds of pages within a translation notebook, for example, the contributor would rather be able to submit the whole object with a given set of metadata, and identify individual pages that would garner additional information, than submit the book page by page. In another example, if a contributor has a box full of letters all relating to the same subject or time period, which must be individually digitized, why not allow the batch upload of all those letters, with identical metadata for those identifiers which are identical (subject, period, date, whatever) and add unique information (author, for example) to individual objects?

Archival Needs: Contextualizing, or providing the provenance of a given object, through hierarchical relationships to other objects within a collection is one of the definitive characteristics of archives. Even if digital archivists were willing to question traditional notions of appropriate archival content; and traditional processes of appraisal, acquisition, and processing - the archival principles of provenance and contextualization are fundamental to archival repositories. To call something a digital archive, we must to be able to define relationships between objects within a collection. Again, as far as our programmers can tell, hierarchical organization more than one level deep is not a DSpace functionality.

We have identified METS as a likely solution to both of the above challenges. How that will play out within the DSpace repository system remains to be seen.

Summary and Conclusion

Because the onus of “good practice” is rapidly transferring to users rather than trained professionals, we wanted to explore the choices that invested users would make regarding the appraisal, organization, and representation of their materials, given little, if any archival guidance. We found that users’ conceptions of historicity and importance were at odds with archival best practice, as were their organizational methods and aspirations for their collections. Whether going against “best practice” is positive or negative in this situation is a question worth exploring, however. The development of digital archives is so recent, and the landscape in the digital realm is so different from the analog world, that it might be worth exploring natural user behavior in this environment a bit more in depth before either trying to change user behavior or force them into a mold they would not naturally fit into.

Instead of forcing users to comply with archival best practice, we believe that more research is called for. The research could be traditional; a larger study like this one looking at the appraisal, description, and organization

of materials by uninitiated users would be very valuable for the development of robust, evidentially valuable, and functional user-developed digital repositories.

There are also more theoretical subjects worth consideration: the role of narrative in self-archiving situations; the meaningful representation of provenance and context in digital systems, and what those concepts actually mean to invested users. While the archival concept of contextualization emerged as a natural behavior by both contributors, the manifestation of the contextualization calls for study of various supporting models/templates. New questions about the possible application of archival methods in a digital realm emerge, especially as regards the types and applications developed for digital archiving. More extensive study on the descriptive processes for metadata creation by non-archivists is also needed.

Finally, there are societal and legal issues that will ultimately determine the success or failure of these systems: user definitions of achievement and the relative importance of their work; and concerns about intellectual property, promotion and tenure. We found that even though users approached us to have their work digitized and placed in a digital repository, they were ultimately hesitant to define certain types of their work as “important,” or “significant” enough for digitization. They also stated that they would have been even more hesitant to include their work in such a repository at any earlier stage of their career for fear of retaliation or “being found out.” Digital repository developers will need to be cognizant of these challenges and hesitancies when developing, populating, and advertising their collections.

In this paper we have discussed the need for a more stable self-archiving model, and some of the challenges and questions inherent in the existing model. We have provided a discussion of the spontaneous behaviors of users who were interested in digitizing and archiving their own work. Finally, we have provided recommendations for system development and outreach initiatives and opened the door for much needed future research in this area.

References

Bearman, D. (1999) Reality and chimeras in the preservation of electronic records *D-Lib Magazine* 5(4)

Bury, C. (2003) *Nightline: The loss of the Columbia* http://www.globalsecurity.org/org/news/2003/030203-shuttle10.htm

Crane, D. (1972) *Invisible colleges: Diffusion of knowledge in scientific communities* Chicago, IL: University of Chicago Press

Harding, R. S., & Oswald, A. (2001) *Western Union Telegraph Company Records, 1820 - 1995* http://invention.smithsonian.org/resources/fa_wu_container25.aspx

Higham, R. (2002, Monday, 2 December) *Digital doomesday book unlocked* http://news.bbc.co.uk/1/hi/technology/2534391.stm

Lavoie, B. (2000) Meeting the challenges of digital preservation: The OAIS reference model *OCLC Newsletter* (January / February), 26-30

Lysakowski, R. (1997) *Comparing Paper and Electronic Laboratory Notebooks - Part 1* http://www.censa.org/html/Publications/Comparing_Paper_and_Electronic_Lab_NotebooksP1_SC&A_May-1997.htm

Ross, S. (2000) Changing trains at Wigan: Digital preservation and the future of scholarship *NPO Preservation Guidance Occassional Papers*

Rothenberg, J. (1999) *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation* Washington, D.C.: Council on Library and Information Resources (CLIR)

Sawyer, D., & Reich, L. (2003) *ISO Reference Model for an Open Archival Information System (OAIS)* Tutorial Presentation: National Aeronautics and Space Administration (NASA)

Tansley, R., Bass, M. J., Stuve, D., Branschofsky, M., Chudnov, D., McClellan, G., et al. (2003) The DSpace institutional digital repository system: Current functionality *Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries, Houston, TX*

Tibbo, H. (2003a) Minds of Carolina - Introduction: Situation and Challenge Retrieved June 22, 2006, from http://www.ibiblio.org/minds/

Tibbo, H. (2003b) On the nature and importance of archiving in the digital age *Advances in Computers* 57

Welfel, O. (1998) On the written trail of science: The laboratory notebook *Center for the History of Physics Newsletter* (Vol. XXX): American Institute of Physics

Winget, M., & Andrew, A. (1999) *Carl W. Gottschalk Papers* http://www.lib.unc.edu/mss/inv/htm/04935.html#N3399