Curation in the Curriculum: Equipping the Profession to Ensure the Preservation of Information

Ross Harvey
Graduate School of Library and Information Science
Simmons College
ross.harvey@simmons.edu

ABSTRACT
This paper proposes a new area of professional practice based on preservation, required in the LIS profession because of changes in the ways that libraries operate and of changes in education for librarianship, as exemplified by the iSchools paradigm. It notes the significant similarities between analog preservation and digital preservation, and proposes these as the basis for new curriculum for a curatorial stream.

Categories and Subject Descriptors
K.3.2 Computer and Information Science Education – *curriculum*

General Terms
Documentation, Human Factors.

Keywords
Curriculum, Curation, Preservation.

1. INTRODUCTION
This paper is speculative. It posits a question: if (or should it be when?) the iSchools paradigm becomes the dominant paradigm in schools that educate for library and information science (LIS), what happens to the preservation function of libraries, one of the traditional concerns of librarianship? This paper proposes curation as a useful model for considering this question. It therefore relates to the interests of the iConference 2010 in that it reflects on "the core activities of the iSchool community, including … engagement between the iSchools and wider constituencies" and is concerned with two of the conference's areas of interest, "Information management: … technologies of forgetting and remembering" and " Digital libraries: preserving digital information …".

My speculation was prompted by a comment by Daniel Greenstein reported in September 2009, that "The university library of the future will be sparsely staffed, highly decentralized, and have a physical plant consisting of little more than special collections and study areas."[1] This comment contains a concept that has significant implications for the concerns, research activities and curricula of iSchools: the de-emphasizing of library collections and their preservation.

2. THE CHANGING NATURE OF LIBRARIES
Is Daniel Greenstein right? As reported[1] he also suggests that "Within the decade …groups of universities will have shared print and digital repositories where they store books they no longer care to manage. … Under such a system, individual university libraries would no longer have to curate their own archives in order to ensure the long-term viability of old texts". Whether or not we agree with Greenstein's comments, there is no doubt that university libraries, like other kinds of libraries, are restructuring in response to the changing ways in which information is created, managed, and used. One of the many sources that address this realignment is Derek Law's discussion of challenges and changes facing university libraries. One of Law's conclusions is that "one glaring gap remains, the absence of any acceptable definition of trusted repositories"[2]. This point is noted again later in the paper.

Comments such as these suggest that there will be a reduced role for the preservation of physical collections, except in specialized centers[1]. There will also be an increased role for the preservation of digital collections (for which there is ample supporting evidence that is not noted in this paper), but there are at present "glaring gaps" in the infrastructure to accomplish this role[2].

3. THE PRESERVATION FUNCTION AND CHANGING CURRICULUM
The preservation function is one that has traditionally been considered as central to LIS practice. This is indicated in statements such as "The preservation function – the stewardship of the accumulated knowledge base – represents the central obligation of librarianship"[3] and "The archival functions of collecting and preserving are intrinsic parts of the research library's service"[4]. This central role has in the past been acknowledged in the traditional curriculum of LIS schools. However, it should be understood that the centrality of preservation is not universally acknowledged. An emphasis on preservation is still perceived by some librarians as "a step backwards to a world from which automation, new media, management science and those exciting possibilities of the new technology had rescued them"[5].

Curriculum is changing in LIS schools. The increasing reach and influence of the iSchools paradigm was recently characterized as reflecting "the extent to which LIS schools have engaged and embraced technological change" and as "signifying a paradigmatic shift in the educational and disciplinary philosophy of many schools that historically were the providers of library education"[6]. Does the iSchools paradigm (the set of practices that define a scientific discipline) and the curriculum associated with it give preservation a central role? It is neither explicitly noted in, nor excluded from, the general statement of purposes on the iSchools web site that "expertise in all forms of information is required for progress in science, business, education, and culture. This expertise must include understanding of the uses and users of

information, the nature of information itself, as well as information technologies and their applications"[7]. In fact some iSchools pay considerable attention to preservation in teaching and conducting research in digital curation, for example the School of Information and Library Science at the University of North Carolina at Chapel Hill though its DigCCurr project. The iSchools are not alone in not acknowledging a more central role for preservation; the ALISE Research Areas classification scheme[8] pays scant attention to preservation, which is represented mainly in category 91, "Preservation and Archiving", and implied in other categories, perhaps category 100 "Digital Archive Informatics" (whatever that means), and categories relating to special materials ("26. Archival Collections" and "27. Special Collections/Rare Books", for instance).

4. INCREASING ROLE FOR PRESERVATION

If Greenstein, Law and others are correct, there will an increasing role for preservation in the practice of librarianship, although it will (according to Greenstein at least) be concentrated in the hands of a small number of storage facilities rather than, as is the case now, in most libraries. This leads to a consideration of the skills and other requirements for running large repositories housing both digital and analog (non-digital) materials. These skills and other requirements are, I suggest, different from the technology focus represented by the iSchools paradigm These differences support an argument for new thinking about professional roles and, therefore, of the curriculum of LIS schools and iSchools – thinking that is based on the coexistence of the new technologies, which bring considerable potential benefits, and the traditional services, practices and values as represented by preservation. This rethinking could result in the development of two streams in the profession, the first principally concerned with *adopting, implementing and using new technologies to serve the user* (the "understanding of the uses and users of information, the nature of information itself, as well as information technologies and their applications" of the iSchools agenda), and the second a *curatorial stream*, primarily concerned with maintaining the sources of information, rather than with the means to access and exploit it. This latter role will be one of preserving and ensuring the availability of the sources that contain the information. The curatorial role has a long tradition in the library profession, but has been de-emphasized from the latter part of the twentieth century to the present.

But the term *preservation* doesn't quite cut it to describe what is required. It is redolent of old books, pest control, and an obsession with climate control – remnants of "a world from which automation, new media, management science and those exciting possibilities of the new technology had rescued" librarians[5]. (Ratcliffe's comment ignores many of the facts– Which "new" technology has as exciting a history as that of mass deacidification, with its space capsules and explosions?) In the world of digital information the term *preservation* has associations that limit its applicability. We need to redefine the term to free it from its associations with solid objects so that we can accommodate the preservation challenges of digital information.

5. CURATION AND STEWARDSHIP

The terms *curation* and *stewardship* are useful to consider in this context. These terms are, relatively speaking, free from the associations with physical objects (especially printed books) that the term *preservation* has. This freedom allows us to develop new ways of working that focus on both the physical objects that store information (analog preservation) and the information contained in bit streams whose physical location, if they have one at all, is likely to change frequently. It also suggests a useful way of considering the curatorial stream proposed above, which would have as its primary concern maintaining the sources of information, regardless of their form, to ensure their availability, currently and over time.

What, then, is *curation*? My definition is based on the many life-cycle models that have developed to describe the requirements for managing of digital objects over time. These are plentiful. I find the Digital Curation Centre's Curation Lifecycle Model[9] to be the most useful. This considers information in digital forms (in terms of data, digital objects, and databases) and recognizes the centrality of metadata, planning, and collaboration in managing such information over time. Its key "sequential actions" consider the conceptualizing and planning of projects that generate data, through its ingest into a repository and the actions involved with managing, storing and preserving them over time, to the requirements for accessing and using them, and re-using them. This encompassing view is also apparent in other definitions, such as that of the Institute of Museum and Library Science (IMLS): "digital curation (creation, authentication, archiving, preservation, retrieval, and representation of high-quality data for use and reuse over time)"[10].

For data and information sources in digital form, the term *digital curation* is a more inclusive concept than either digital *archiving* or digital *preservation*. It addresses the whole range of processes applied to data over their life-cycle. Digital curation begins before data are created by setting standards for planning data collection that results in "curation-ready" data – data that are in the best possible condition to ensure they can be maintained and used in the future. Digital curation emphasizes adding value to data sets, through things such as additional metadata or annotations, so they can be re-used. Digital curation involves a wide range of stakeholders cutting across disciplinary boundaries; as well as cultural heritage organizations such as libraries, archives and museums, it also involves funding agencies, government bodies, national data centers, institutional repositories and learned societies.

The term *stewardship* is also a contender. In the context of data, the terms *curation* and *stewardship* "both focus on the data but have different views about the nature of data, their life cycles and relations with their environments"[11]. Curation is principally interested in "organizing and overseeing data holdings" and "deals with guidelines and procedures for data ingestion, archive and delivery". Data stewardship is a larger concept, which "provides a large conceptual framework, an overarching process occurring now but attending to the past and taking into account and influencing the future, stretching from data planning to sampling, from data archive to use and reuse"[11]. Both are concerned with a wider view than just preservation considered only as a technical process isolated from services, policies and stakeholders[12].

6. CONCEPTUALIZING PRESERVATION MORE BROADLY

Curation concepts and the curation lifecycle provide a way of conceptualizing preservation more broadly so that it encompasses both analog preservation and digital preservation. Table 1 summarizes the results of a more detailed comparison of the principles and practices of analog preservation and digital preservation carried out as the basis of a presentation to Harvard Library Staff in 2009. Although this table has limitations, such as being a crude content analysis of various public statements about preservation, it assists in developing a broader understanding of curation.

Table 1. Analog and digital preservation principles and practices compared

OBVIOUS SIMILARITIES	
Analog Preservation	**Digital Preservation**
Obsolescence and degradation of artefacts are always with us	Obsolescence and degradation of artefacts are always with us
Ensuring the longevity of artefacts	Protect data
Ensuring the longevity of the information content stored in artefacts	Maintain ongoing access to digital materials despite technological change
Creation of 'preservation-friendly' artefacts	Negotiate with the creators of material to use open, well-supported standard formats for which access tools may remain available; Conceptualize; Create or receive
Redundancy – multiple copies are also a good thing	Provide adequate data backups and create multiple copies; Multiple copies/redundancy
Security and emergency management	Have disaster recovery contingencies in place
Improving storage environment and maintaining it at controlled levels; Prolonging the life of the artefact through preventive action	Provide stable, secure media storage conditions and proper handling
Reformatting (converting the information to a more stable form); Replacing deteriorated artefacts	Copy data to new media well within the expected media life, and check the accuracy of copying
Careful documentation of the condition of the artifact and of procedures and materials used in treatment	Gather sufficient metadata about the material's technical characteristics and requirements to support its preservation and management; Description and representation information; Enhance the metadata
Ongoing policy and procedures review	Monitor the technological environment for signs that formats etc are becoming obsolete; Monitor for evolving solutions; Preservation planning
Protecting artefacts	Maintain adequate data security and protection from viruses, system attack and unauthorized modification of data
Stabilization of artefacts	Limit the range of formats to be managed
Appraisal	Appraise/Select
Collaboration	Work with or seek help from others to develop solutions; Community watch and participation; Interoperability: 'you are not alone'
Keep the original – we keep the original after we reformat it (for example, retain the artefact after digitizing)	Keep the original (bit-stream, analog after digitising)
Encapsulation – we can enclose artefacts in protective material	Encapsulation (digital files – XML wrappers)
All copying introduces change which needs to be accommodated (for example, in reformatting we emphasize checking and validating of the copy)	Constantly check and validate, because all copying of data (such as migration) introduces change
Authenticity – we strive to maintain the authenticity of the artefact (although we acknowledge this isn't always possible) is a good thing	Decode to uncompressed and save as uncompressed (in addition to keeping the original)

This listing of principles and practices suggests significant similarities. For analog preservation, most of the list is encompassed by an emphasis on the artefact – the physical object – and especially on its characteristic of staying reasonably stable over time. This is expressed particularly in the concept of benign neglect: the idea that most artefacts do not deteriorate rapidly if ignored, thus buying time before preservation treatments need to be applied. It is also apparent in practices and procedures such as those that aim to stabilize the artefact, for instance by using stable materials. The integrity of the artefact, its original state if you like, is maintained as far as possible by practices such as limiting intervention in treatment so that its role as an object of material culture is not detracted from. Storing lesser-used materials off-site in an optimally controlled environment is also based on keeping the artefact for as long as possible, with the implication that in doing so its information content will not become unreadable.

For digital preservation, there is an emphasis on the ability to use (and re-use) the digital object that is not apparent in statements about traditional preservation, presumably because in the preservation of an artefact, its information content is considered to be understandable without modification of the artefact. This emphasis is expressed in actions such as retaining old hardware and software to allow access to obsolete media and data, and those in the 'Access, use and reuse' action of the Digital Curation Lifecycle.

My approach has not been very scientific and my conclusions may not withstand too heavy a scrutiny. They need to be tested more

rigorously. One possibility is to apply a framework for thinking about how archival science and techniques translate to the digital environment, articulated by Ken Thibodeau in a 2008 presentation and used here with his kind permission[13]. Thibodeau's framework has four parts:

- *Keep:* apply established archival science or techniques when the knowledge or technique is valid independent of the context in which it is applied
- *Cut:* don't apply established archival science or techniques when the knowledge or technique is not independent of the context in which it is applied
- *Craft:* adapt or modify archival science or techniques that is fundamentally sound, but has not been articulated appropriately for cyberspace
- *Create:* develop new concepts and techniques needed in cyberspace.

What is now required is the application of this framework to the analog preservation principles and practices defined above, testing them to see if and how they need to be modified to be valid also for the principles of digital preservation.

7. CONCLUSION

The take-home message from this speculation can be stated as follows. Changes in the ways that libraries operate suggest a need for the development of a curatorial stream in the LIS profession. The primary concern of this stream is maintaining the sources of information, regardless of their form, to ensure their availability, now and in the future. Changes in education for librarianship are exemplified by the iSchools paradigm, which does not articulate the requirements of this new focus. The significant similarities between analog preservation and digital preservation, combined with new ways of thinking about curation (especially in the digital context) present a strong basis for new curriculum for a curatorial stream to be developed.

8. REFERENCES

[1] Kolowich, S. 2009. Libraries of the future. Inside Higher Ed (September 24, 2009) http://www.insidehighered.com/news/2009/09/24/libraries

[2] Law, D. 2009. Academic digital libraries of the future: an environment scan. New Review of Academic Librarianship 15, 1 (April 2009), 53-67 (p.64).

[3] Commission on Preservation and Access 1990. Annual Report 1989-90, p.1.

[4] Shoemaker, S. (ed.) 1989. Collection Management Issues. Neal-Schuman, p.66.

[5] Ratcliffe, F. W. 1991. In A Reading Guide to the Preservation of Library Collections (ed. G. Kenny). Library Association, p.4.

[6] Bonnici, L. J., Subramaniam, M. M., and Burnett, K. 2009. Everything old is new again: The evolution of library and information science education from LIS to iField. J. Educ. Libr. Inf. Sci. 50, 4 (Fall 2009), 263-274 (p.273).

[7] http://www.ischools.org/site/charter/

[8] http://www.alise.org/mc/page.do?sitePageId=55727&orgId=ali

[9] http://www.dcc.ac.uk/lifecycle-model/

[10] IMLS 21st Century Librarian grants: 5. Programs to Build Institutional Capacity http://www.imls.gov/applicants/grants/pdf/L21_2010.pdf (p.12).

[11] Karasti, H., Baker, K. S., and Halkola, E. 2006. Enriching the notion of data curation in e-science. Computer Supported Cooperative Work 15, 4 (Aug 2006), 321-358. http://portal.acm.org/citation.cfm?id=1166839 (p.352).

[12] Lavoie, B. F., and Dempsey, L. 2004. Thirteen ways of looking at ... digital preservation. D-Lib Magazine 10, 7/8 (July/August 2004) http://www.dlib.org/dlib/july04/lavoie/07lavoie.html.

[13] Thibodeau, K. 2008. Archival Curation in Cyberspace. Presentation to the Mid Atlantic Archives Conference, November 7, 2008.