Where's the Archivist in Digital Curation? Exploring the Possibilities through a Matrix of Knowledge and Skills*

CHRISTOPHER A. LEE and HELEN TIBBO

RÉSUMÉ Depuis quelques années, le concept de la conservation numérique (« *digital curation* ») recouvre des activités englobant une vaste étendue de professions, d'institutions, de joueurs et de secteurs. Dans cet article, les auteurs font le résumé des intentions, des principes, des méthodologies et des données qui ont servi de base à la DigCCurr Matrix of Digital Curation Knowledge and Skills et élaborent ses six dimensions : 1) les mandats, valeurs et principes; 2) les fonctions et compétences; 3) le contexte professionnel, disciplinaire ou institutionnel/organisationnel; 4) le type de ressource; 5) les connaissances prérequises; 6) le point de transition dans le continuum de l'information. Ils expliquent comment la matrice peut appuyer l'identification et l'organisation du matériel dont il sera question dans les programmes d'éducation professionnelle. Cette énonciation détaillée des activités de la conservation numérique révèle de nombreuses occasions pour promouvoir, développer et parfaire davantage les compétences et habiletés des archivistes. Les auteurs présentent des aspects de la DigCCurr Matrix qui croisent les compétences archivistiques bien établies, tout en présentant des domaines de croissance et de collaboration potentielles. Enfin, ils expliquent comment l'élargissement du champ professionnel à partir de la préservation numérique (« *digital preservation* ») vers l'univers plus vaste de la conservation numérique ressemble et complète le mouvement vers des approches fondées sur le continuum des documents et le dépassement de la conservation (« *postcustodialism* »). Les auteurs élaborent enfin sur les domaines dans lesquels les archivistes pourraient prendre avantage de ce changement.

ABSTRACT In recent years, the concept of digital curation has served as an umbrella spanning activities across a diversity of professions, institutions, actors, and sectors. We summarize the motivation, principles, methodology, and data that have served as the basis for the DigCCurr Matrix of Digital Curation Knowledge and Skills, and elaborate its six dimensions: 1) mandates, values, and principles; 2) functions and skills; 3) professional, disciplinary, or institutional/organizational context; 4) type of resource; 5) prerequisite knowledge; and 6) transition point in the information continuum. We explain how the Matrix can support the identification and organization of material to

* This work has been supported through several grants including DALL (Andrew W. Mellon Foundation), DigCCurr (IMLS # RE-05-06-0044), DigCCurr II (IMLS #RE-05-08-0060-08), CDCG (IMLS #LG-05-09-0040), and ESOPI-21 (IMLS # RE-05-09-0085-09).

be addressed in professional education programs. This detailed articulation of digital curation activities reveals numerous opportunities to promote, advance, and further enhance the skills and capabilities of archivists. We present aspects of the DigCCurr Matrix that intersect with long-standing, archival competencies as well as discussing areas of potential growth and collaboration. Finally, we explain how expanding one's professional focus from digital preservation to the wider universe of digital curation is similar and complementary to the move toward postcustodial and continuum approaches within the archival profession. We elaborate areas in which archivists can take advantage of these connections to advance the archival enterprise.

Introduction

The past several decades have been witness to numerous advances in the management, preservation, and dissemination of digital resources. There have been many distinct streams of activity that address parts of the puzzle or particular types of materials, e.g., computer science, archival administration, records management, librarianship, data engineering, electronic discovery, digital forensics, museum curation, and the management of information systems (MIS). Within the past two decades, individuals engaged in many seemingly disparate streams of activity have increasingly come to recognize that they share a common set of challenges and opportunities. The term "digital curation" has recently come into use, reflecting the increasing confluence of previously distinct communities. Elizabeth Yakel offers the following definition of digital curation:

> The active involvement of information professionals in the management, including the preservation, of digital data for future use.[1]

Within this broad scope, two general questions have motivated a great deal of our work: What knowledge and competencies do professionals need in order to perform digital curation work? What can and should be the roles of archivists in digital curation?

Emergence of the Concept of Digital Curation

The term "curation" has a long history. It was used as early as the fourteenth century to denote healing or curing of the body or soul, and later as guardian-

1 Elizabeth Yakel, "Digital Curation," *OCLC Systems and Services*, vol. 23, no. 4 (2007), pp. 335–40. The Digital Curation Centre offers a somewhat different definition: "maintaining and adding value to a trusted body of digital research data for current and future use; it encompasses the active management of data throughout the research lifecycle." See *DCC Charter and Statement of Principles*, http://www.dcc.ac.uk/about-us/dcc-charter (accessed on 12 July 2011). The DCC definition emphasizes that digital curation extends beyond preservation activities. However, its focus on "the research lifecycle" does not clearly include phenomena such as families keeping their own histories or curation of government records.

ship or caretaking for one's affairs.[2] By the 1960s and 1970s, the term was used to designate systematic care for collections of scientific specimens.[3] In the past few decades, the term curation has often been used in the literature regarding the administration of museums, a functional linguistic shift from the much more long-standing term "curator," used since the seventeenth century to describe "the officer in charge of a museum, gallery of art, library, or the like."[4] Lewis R. Binford introduced the concept of "curation" to the field of archeology in 1973, as the transport for "anticipated performance of different activities" and reuse of tools by ancient people across locations[5]; the implication was that early humans were caring for artifacts in ways that were attentive to future use. Many anthropologists and archeologists have since taken up the concept of curation, and it can be construed as perpetuating the utility of tools "beyond the context of a single use."[6] In the 1980s and 1990s, use of the phrase "data curation" emerged in literature related to the management of scientific data.[7] Two reports in 2003 brought further attention to the idea of data curation within the context of e-science and cyber-infrastructure.[8]

2 "Curation." *Oxford English Dictionary*, 2nd ed. (Oxford, UK, 1989).

3 See for example, the 1967 colloquium entitled, "The Application of Electronic Data Processing in the Curation and Use of Biological Collections" discussed by Theodore J. Crovello in "Problems in the Use of Electronic Data Processing in Biological Collections," *Taxon*, vol. 16, no. 6 (1967), pp. 481–94; James Hicks and Pearl M. Hicks, "A Selected Bibliography of Plant Collection and Herbarium Curation," *Taxon*, vol. 27, no. 1 (1978), pp. 63–99; Michael G. Bassett, ed., *Curation of Palaeontological Collections: A Joint Colloquium of the Palaeontological Association and Geological Curators' Group* (London, 1979).

4 "Curator." *Oxford English Dictionary*, 2nd ed. (Oxford, UK, 1989).

5 L.R. Binford, "Interassemblage Variability: The Mousterian and the 'Functional' Argument," in *The Explanation of Culture Change: Models in Prehistory; Proceedings*, ed. Colin Renfrew (Pittsburgh, 1973), pp. 227–54.

6 Michael J. Shott, "An Exegesis of the Curation Concept," *Journal of Anthropological Research*, vol. 52, no. 3 (Autumn 1996), pp. 259–80.

7 See for example, Great Lakes Archaeological Research Center Inc., *Archaeological Investigations, Navigation Pool II, Upper Mississippi River Basin,* Vol. 2, "Reports of Investigation No. 151" (Waukesha, WI, 1984); Nancy Evans, "Trends in Planetary Data Analysis: Executive Summary of the Planetary Data Workshop" (Washington, DC, 1984); Space Physics Division, Office of Space Science, "Report of the Community-Wide Workshop on NASA's Space Physics Data System" (NASA, 1993); Judith A. Blake, Carol J. Bult, Michael J. Donoghue, Julian Humphries, and Chris Fields, "Interoperability of Biological Data Bases: A Meeting Report," *Systematic Biology*, vol. 43, no. 4 (1994), pp. 585–89; Robert J. Robbins, ed., "Report of the Invitational DOE Workshop on Genome Informatics, 26–27 April 1993, Baltimore, Maryland," *Journal of Computational Biology*, vol. 1, no. 3 (1994), pp. 173–90.

8 Daniel E. Atkins, Kelvin K. Droegemeier, Stuart I. Feldman, Hector Garcia-Molina, Michael L. Klein, Paul Messina, David G. Messerschmitt, Jeremiah P. Ostriker, and Margaret H. Wright, *Revolutionizing Science and Engineering through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure* (Washington, DC, 2003); Philip Lord and Alison Macdonald, *E-Science Curation Report:*

The phrase "digital curation" has more recent origins. In 2001, a seminar entitled "Digital Curation: Digital Archives, Libraries, and E-Science" was hosted by the Digital Preservation Coalition and the British National Space Centre.[9] The following year, the United Kingdom (UK) Joint Information Systems Committee (JISC) issued an "Invitation to tender: requirements and feasibility study on preservation of e-prints,"[10] which resulted in the formation of the Digital Curation Centre (DCC) in 2004. In 2006, the Institute for Museum and Library Services (IMLS) – building on the 21st Century Librarian Program that it began in 2003 – called for grant proposals to develop educational programs in digital curation; the IMLS has funded several programs as a result of this call.[11]

Digital curation has served as an umbrella concept spanning activities across a diversity of professions, institutions, actors, and sectors. "Digital curation" is less wedded to specific institution types than phrases such as "digital archives" or "digital libraries." It also reflects a tendency toward increasing convergence across different types of cultural institutions, driven in large part by technological changes and considerations.[12]

In 2004, Neil Beagrie noted that "digital curation" was "increasingly being used for the actions needed to maintain digital research data and other digital materials over their entire lifecycle and over time for current and future generations of users," which includes not only digital preservation but also "all the processes needed for good data creation and management, and the capacity to add value to data to generate new sources of information and knowledge."[13]

Efforts to further articulate and define the characteristics of digital curation have come at the same time that the archival profession faces unprecedented opportunities and challenges related to electronic records. Many archival institutions have been addressing electronic records issues for decades, and in our experience, the most successful ones have been those who actively seek

Data Curation for E-Science in the UK: An Audit to Establish Requirements for Future Curation and Provision (London, UK, 2003).

9 Neil Beagrie and Philip Pothen, "The Digital Curation: Digital Archives, Libraries and E-Science Seminar," *Ariadne* 30 (2001), http://www.ariadne.ac.uk/issue30/digital-curation/ (accessed on 12 July 2011).

10 See http://web.archive.org/web/*/http://www.jisc.ac.uk/index.cfm?name=funding_7_02 (accessed on 12 July 2011).

11 Joyce Ray, "Sharks, Digital Curation, and the Education of Information Professionals," *Museum Management and Curatorship*, vol. 24, no. 4 (2009), pp. 357–68.

12 Helen R. Tibbo and Christopher A. Lee, "Convergence through Capabilities: Digital Curation Education for Libraries, Archives and Museums," in *Archiving 2010: Final Program and Proceedings, June 1–4, 2010, Den Haag, The Netherlands* (Springfield, VA, 2010), pp. 53–57.

13 Neil Beagrie, "The Digital Curation Centre," *Learned Publishing*, vol. 17, no. 1 (2004), pp. 7–9.

connections and collaborations with allied experts and professionals.[14] Digital curation is a recent (and we believe, useful) phrase to designate a set of opportunities for cross-institutional and cross-disciplinary engagement that have been evolving – with many valuable advances and contributions by archivists – for decades.

The Nature of Digital Curation

In contrast to caring for analogue materials, digital curation brings a wide array of opportunities and challenges. Opportunities include both wider and integrated access, representation of an increased range of human experience, persistence through redundant copying, economies of scale, and enrollment of collective expertise. Challenges include bit rot, obsolescence, social inertia, technology monitoring, intellectual control, access environments, and the ability to convey meaning over time. The technical complexities involved in digital preservation relate primarily to technological dependency. A document stored on an analogue medium such as ink on paper, can be read directly with the human eye. Accessing and using a document stored as a digital object, however, requires the coordinated operation of various hardware and software components (e.g., storage medium, peripheral devices, operating system, device drivers, application software). Because of innovations in the information and communication technology industries, these components quickly become obsolete and unavailable. Future access to, and use of, digital objects that depend on current technology raise issues of what Margaret Hedstrom has called "temporal interoperability."[15]

The challenges associated with digital curation are not solely (and often not even predominately) technical.[16] In order for digital collections to be sustainable over time, the actors responsible for the archives must continuously have appropriate expertise, resources, and a political/institutional mandate to carry out the work required. Given the cost and complexity of digital archives, as well as the potential to exploit the rich sets of relationships across individual collec-

14 We are certainly not the first authors to make this argument. See for example, Margaret Hedstrom, "Building Record-Keeping Systems: Archivists Are Not Alone on the Wild Frontier," *Archivaria* 44 (Fall 1997), pp. 44–71.

15 Margaret Hedstrom, "Exploring the Concept of Temporal Interoperability as a Framework for Digital Preservation," *Paper presented at the Third DELOS Network of Excellence Workshop on Interoperability and Mediation in Heterogeneous Digital Libraries, Darmstadt, Germany, 8–9 September 2001*, http://www.ercim.eu/publication/ws-proceedings/DelNoe03/10.pdf (accessed on 12 July 2011).

16 For example, of the "ten basic characteristics of digital preservation repositories" identified by the Center for Research Libraries, only one focuses specifically on the "technical infrastructure adequate to continuing maintenance and security of its digital objects" (2007), http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/core-re (accessed on 12 July 2011).

tions, coordination of work across social boundaries (institutional, regional, disciplinary, organizational, and professional) is also important.

When acquiring, managing, and providing access to materials, professionals in collecting institutions must consider various norms, laws, codes of ethics, policies, procedures, and personal values. As they address the curation of digital collections, they are increasingly discovering "policy vacuums,"[17] in which there is either no existing guidance on new issues, and/or "latent ambiguities"[18] in established policies and procedures. Digital curation professionals must be able to clearly articulate policies, procedures, and practices in ways that were not previously necessary.

Digital resources are composed of interacting components that can be considered and accessed at different levels of representation (e.g., bitstream, through a file system; files as rendered through specific applications; records composed of multiple files; abstract "works"; larger aggregations such as websites). To ensure integrity and future use, digital curation professionals must make decisions regarding treatment at multiple levels of representation. This will require new and enhanced mechanisms for professional development. Numerous programs across the globe are engaged in this collective educational endeavour.[19] We have been actively developing and administering one such program.

Digital Curation Education Projects at the University of North Carolina at Chapel Hill

The School of Information and Library Science at the University of North Carolina, Chapel Hill (UNC SILS) is engaged with digital curation education at the master's, doctoral, and professional levels. The DigCCurr (Digital Curation Curriculum) and DigCCurr II projects – both funded by the IMLS – have developed conceptual frameworks, educational offerings, professional field experiences, and research opportunities to prepare digital curation professionals.[20] DigCCurr I (2006–2009) focused on developing a curriculum and practicum

17 James H. Moor, "What Is Computer Ethics?" *Metaphilosophy*, vol. 16, no. 4 (1985), pp. 266–75.

18 Lawrence Lessig, *Code: Version 2.0* (New York, 2006), pp. 25–28, 155–277. Lessig borrows the term "latent ambiguity" from Samuel Williston, *A Treatise on the Law of Contracts*, 3rd ed., ed. by Walter H.E. Jaeger (Mount Kisco, NY, 1957), pp. 627, 898.

19 For information about collaboration across the digital curation education initiatives, see the International Digital Curation Education and Action (IDEA) Working Group, http://ideaworkgroup.org, and the International Curation Education (ICE) Forum, http://www.dcc.ac.uk/events/workshops/international-curation-education-ice-forum (both accessed on 12 July 2011).

20 IMLS (Institute for Museum and Library Services) Grant Award # RE-05-06-0044. For a summary of activities and the final report of DigCCurr I, see http://ils.unc.edu/digccurr/ (accessed on 12 July 2011).

experiences for master's students, while DigCCurr II (2008–2012) is supporting doctoral and professional education, including summer institutes for the continuing education of professionals.

A fundamental activity in both projects has been a detailed elaboration of the components and scope of digital curation, i.e., what it means to "do" digital curation. Based on data from numerous sources, we have developed a Matrix of Digital Curation Knowledge and Skills discussed below.

Three other IMLS-funded projects have both applied and provided further information for the DigCCurr Matrix. Educating Stewards of Public Information in the 21st Century (ESOPI-21) and Educating Stewards of the Public Information Infrastructure (ESOPI2) are collaborations between SILS and the UNC School of Government, focusing on the preparation of professionals who can navigate the intersection between digital curation, public policy, and public administration.[21] The program is funding students who are completing a dual degree between SILS and the School of Government, complemented by intensive work experiences with partner institutions. Closing the Digital Curation Gap (CDCG)[22] is a collaborative agreement among SILS, the IMLS, and the UK Joint Information Systems Committee (JISC). It is designed to serve as a locus of interaction between leading edge, digital curation research, development, teaching, and training in academic and practitioner communities in libraries, archives, and museums, addressing the issue of professional diffusion of innovations related to digital curation. The growing importance of digital curation activities in supporting the missions of all three types of institutions poses exciting new opportunities for collaboration in professional education.[23]

Developing a Matrix of Digital Curation Knowledge and Skills

As indicated above, one of the driving questions behind the above projects has been: What knowledge and competencies do professionals need in order to do digital curation work?

Methodology and Data

The development of the DigCCurr Matrix has been based on a grounded theory[24] research design, a powerful approach for "understanding of intermingled

21 ESOPI-21 is IMLS Grant Award # RE-05-09-0085-09 and ESOPI2 is IMLS Grant Award # RE-05-11-0076. For more information about the project, see http://ils.unc.edu/esopi21/ (accessed on 12 July 2011). As of the production of this article, ESOPI2 was just getting underway, with a 1 July 2011 start date.

22 IMLS Sponsor Award #LG-05-09-0040.

23 Tibbo and Lee, "Convergence through Capabilities," pp. 53–57.

24 Barney G. Glaser and Anselm L. Strauss, *The Discovery of Grounded Theory: Strategies for*

types of work."[25] In grounded theory, "data collection and analysis are interrelated processes."[26] We began our analysis with the fundamental assumption that the knowledge required to do digital curation work will depend upon characteristics of the specific work that needs to be performed. Because digital curation is an umbrella term that spans a diversity of professional, disciplinary, and institutional contexts – not only work within memory institutions such as museums, libraries, and archives, but also care for digital objects within numerous other types of organizations and the practices of individuals caring for their own materials – we believe that there will never be a single answer to the question "What does one need to know in order to do digital curation?" Instead, the answer depends upon several factors, which we have expressed as the dimensions of the DigCCurr Matrix.

Our work on the Matrix began by articulating a relatively simple set of dimensions, and elements within those dimensions. We then collected, analyzed, and incorporated a diversity of data sources. Our research design has been iterative, with numerous opportunities to gain feedback on and revise both our findings and curriculum materials. As we encountered new data (including frequent input from other scholars and interested professionals), we either incorporated the data within existing elements of the Matrix or revised the Matrix. The result has been a blending of models, guidance documents, and frameworks concerning digital curation, with insights from practitioners who are creating curation processes.

By exploring what digital curation professionals are doing, and discussing the pressing challenges with scholars in the field, we are developing a proposed body of knowledge, skills, and perspectives for educational programs to provide to their students. We recognize that a digital curation curriculum can never be considered finalized but must instead change in relationship to the world it serves. As highlighted below, traditional archival principles, skills, and knowledge assume a central and foundational role in the DigCCurr Matrix and Curriculum, but are by themselves not sufficient for expressing the full range of digital curation activities to be performed in a variety of contexts.

Qualitative Research (Chicago, 1967).

25 Susan Leigh Star, "Grounded Classification: Grounded Theory and Faceted Classification," *Library Trends*, vol. 47, no. 2 (1998), pp. 218–32.

26 Juliet M. Corbin and Anselm Strauss, "Grounded Theory Research: Procedures, Canons, and Evaluative Criteria," *Qualitative Sociology*, vol. 13, no. 1 (1990), pp. 3–21.

Table 1: Sources of Data for the Development of the DigCCurr Matrix

Documentary Sources	We have qualitatively coded syllabi of courses being offered at UNC SILS, using the Matrix of Topics for a Digital Curation Curriculum (described below) as the basis for our coding categories. We have also examined materials from many existing educational offerings outside of UNC SILS.
Interviews with Domain Experts	DigCCurr brought together a seventeen-member Advisory Board of experts from seven countries. We conducted semi-structured interviews with all Board members, gaining their insights on: how to define digital curation; the main functions or activities necessary for digital curation; the topics to cover in a digital curation curriculum; the distinct roles that are needed in digital curation; the types of environments in which graduates of a digital curation program should be able to work; what they would look for in a job candidate; the important skills or knowledge required for digital curation that may be lacking within the professions engaged in the work; and the characteristics of a good, practical field experience.
Participation in Educational Workshops	Between 2006 and 2011, members of DigCCurr and DigCCurr II project teams have participated in ten professional development offerings (workshops and courses) in order to gain insights about scope, content, pedagogical design, and implementation.
Survey	On 20 April 2007, we administered a questionnaire to participants at the DigCCurr 2007 conference. The instrument surveyed the perceptions of respondents in three areas: 1) digital curation challenges, needs, and deficiencies at their local institutions; 2) necessary digital curation curriculum components; and 3) essential digital curation professional competencies. The responses provided a rich set of perspectives from which we identified several themes to serve as a basis for further inquiry. We used the results of the April 2007 questionnaire to design a wider survey of digital curation professionals in March–April 2008, distributed to 221 individuals and completed by fifty-five (25 percent), eliciting their perspectives on 1) barriers to digital curation; 2) core curriculum competencies and functions; and 3) professional competencies and hiring practices.
Ongoing Feedback from Professionals and Students	We have been disseminating and gaining feedback on project deliverables at a diversity of professional conferences. This has included numerous conference papers, presentations, tutorials, and workshops. We have also gained experience in implementing the curriculum at UNC SILS, which has provided both informal day-to-day feedback from students, as well as more formal input through periodic structured evaluation forms completed by students as part of courses and Fellowship experiences.

We have obtained and analyzed a variety of documentary sources to build the Matrix and the subsequent curricular framework. We have carried out – and continue to supplement – a detailed review of existing literature relevant to digital curation and digital curation education. By drawing literature from a diversity of disciplinary and professional contexts, we have been able to elaborate and combine many considerations in more detail than one could previously find in any one place.

A good example of our approach to incorporating and synthesizing existing sources is our use of the *Reference Model for an Open Archival Information System* (OAIS), developed from 1994 to 2002 and adopted as an International Standard in 2003 (ISO 14721:2003). The OAIS describes components and services required to develop and maintain archives, in order to support long-term access to, and understanding of, the information in the archives.[27] Many of the requirements for an OAIS are based on the needs of its Designated Community, which is the set of one or more "user communities" that the OAIS is serving. An OAIS is responsible for digital information over the "long term," which is "long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community."[28] An important insight in the Reference Model is that the "Content Information" to be preserved by an archive is composed not only of a "set of bit sequences" (the "data object") but also of sufficient, associated "Representation Information" to allow the bits to be rendered, used, and understood. The three main roles played by the external entities with which an OAIS interacts are *Producer*, *Consumer*, and *Management*. *Producers* are "persons, or client systems, who provide the information to be preserved." *Consumers* are "persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail." *Management* is "the role played by those who set overall OAIS policy as one component in a broader policy domain."[29]

The OAIS document elaborates both a functional model and an information model. Roughly speaking, the former indicates what an OAIS must do, and the latter indicates what the OAIS must have in its collections. The functional model is composed of seven main functional entities and the interfaces between them: Access, Administration, Archival Storage, Common Services, Data Manage-

27 For a more detailed discussion of the *Open Archival Information System* (OAIS), see Christopher A. Lee, "*Open Archival Information System* (OAIS) Reference Model," in *Encyclopedia of Library and Information Sciences*, 3rd ed., eds. Marcia J. Bates and Mary Niles Maack (Boca Raton, FL, 2009), pp. 4020–30.

28 Consultative Committee for Space Data Systems, *Reference Model for an Open Archival Information System (OAIS),* CCSDS 650.0-B-1 (Washington, DC, 2002), p. 1-1 [hereinafter OAIS].

29 Ibid., p. 1-11.

ment, Ingest, and Preservation Planning. The figure from the Reference Model that has received the most attention in the digital curation literature is a representation of six of the functional entities (see Figure 1). The functional model distinguishes between Submission Information Packages (SIPs), Archival Information Packages (AIPs), and Dissemination Information Packages (DIPs). SIPs are what the OAIS receives from Producers, AIPs are what the OAIS manages and preserves, and DIPs are "derived from one or more AIPs, [and are] received by the Consumer in response to a request to the OAIS."[30] Ingest is the entity that receives SIPs, performs quality assurance on the SIPs, generates AIPs, extracts Descriptive Information from AIPs, and coordinates updates to Archival Storage and Data Management. Archival Storage is responsible for "receiving AIPs from Ingest and adding them to permanent storage, managing the storage hierarchy, refreshing the media on which archive holdings are stored, performing routine and special error checking, providing disaster recovery capabilities, and providing AIPs to Access to fulfill orders."[31] Data Management supports the populating, maintenance, and accessing of both Descriptive Information and administrative data associated with the OAIS holdings; this includes database administration, database updates, performing queries on the data, and producing reports that result from the queries. Administration is responsible for "the overall operation of the archive system," which includes soliciting and negotiating submission agreements, auditing submissions, configuration management, system engineering, activating stored requests, and the establishment and maintenance of standards and policies.[32] Preservation Planning monitors the environment for changes in technology or the needs of the Designated Community; evaluates the implications of those changes to the archive's holdings; designs Information Package templates; "provides design assistance and review to specialize these templates into SIPs and AIPs for specific submissions"; "develops detailed Migration plans, software prototypes, and test plans," and provides periodic recommendations for "archival information updates," standards, and policies.[33] Access both provides and appropriately restricts Consumers' ability to discover, request, and receive information from the archive, including DIPs, "result sets," and reports.[34]

30 Ibid., p. 1-10.
31 Ibid., pp. 4-1, 4-2.
32 Ibid., p. 4-2.
33 Ibid.
34 Ibid.

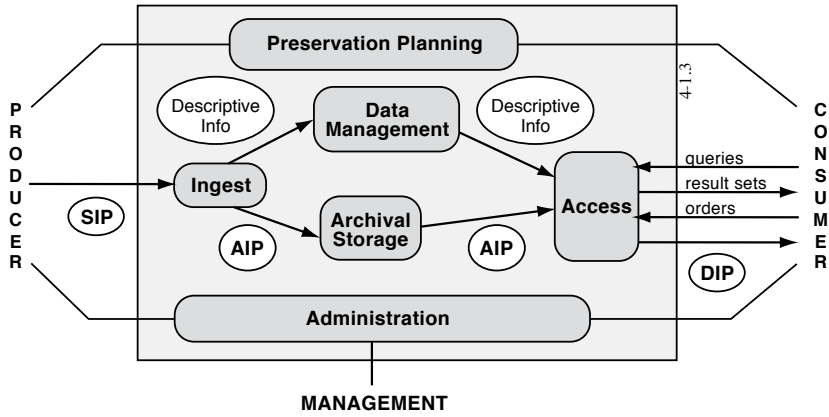

Figure 1: OAIS Functional Entities. Credit: Consultative Committee for Space Data Systems, *Reference Model for an Open Archival Information System* (OAIS), (Washington, DC, 2002), p. 4-1.
Note: SIP = Submission Information Package; AIP = Archival Information Package; DIP = Dissemination Information Package

The information model defines and describes "the types of information that are exchanged and managed within the OAIS."[35] It is based on the recognition that long-term preservation of digital information requires an archive to "store significantly more than the contents of the object it is expected to preserve."[36] An "Information Package" is a logical unit that includes both a digital object and the other types of information that should be associated with the digital object in order to preserve and provide meaningful access to it over time. AIPs are the information packages that are managed internally by the OAIS. An AIP can be either an Archival Information Collection (AIC), "whose Content Information is an aggregation of other Archival Information Packages,"[37] or Archival Information Unit (AIU), "whose Content Information is not further broken down into other Content Information components."[38] Figure 2 presents the main types of information that constitute and are associated with an AIP. The Package Description is information about an Information Package, which is used by Access Aids. An Access Aid is "a software program or document that allow[s] Consumers to locate, analyze, and order Archival Information Packages of interest."[39] Packaging Information, in contrast, is not intended for direct use by Consum-

35 Ibid., p. 4-18.
36 Ibid., p. 4-19.
37 Ibid., p. 1-7.
38 Ibid., p. 1-8.
39 Ibid., p. 1-7.

ers but is instead "used to bind and identify the components of an Information Package" (e.g., volumes and directory information for the components).[40]

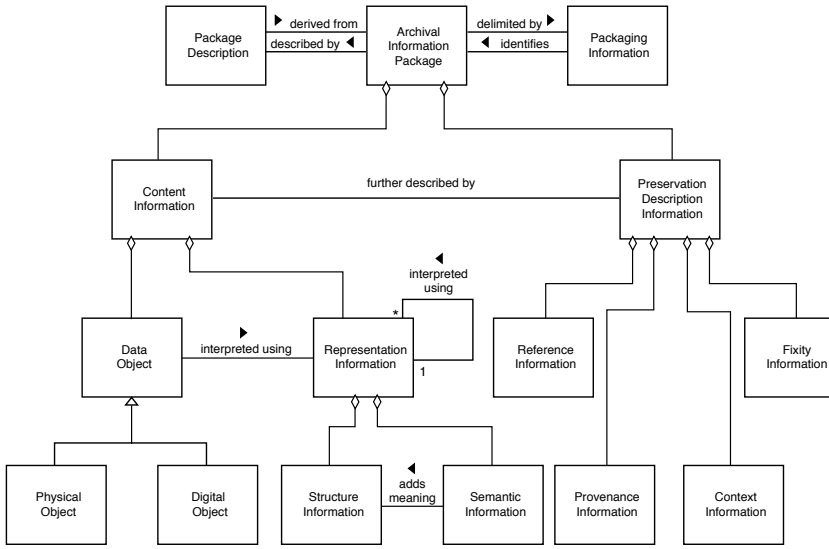

Figure 2: Archival Information Package (Detailed View). Credit: Consultative Committee for Space Data Systems, *Reference Model for an Open Archival Information System* (OAIS), (Washington, DC, 2002), p. 4-37.

The Content Information is "the original target of preservation."[41] It is composed of both the Data Object and Representation Information, which "allows for the full interpretation of the data into meaningful information."[42] Representation Information can be either Structure Information, which "imparts meaning about how other information is organized,"[43] or Semantic Information, which indicates the meaning of language used in either the Structure Information or other parts of the Content Information. Representation Information is often conveyed using Digital Objects, which then require their own Representation Information, resulting in what is called a Representation Network. Content Information is "further described by" Preservation Description Information (PDI), which is composed of:

- Reference Information – "identifies, and if necessary describes, one or more mechanisms used to provide assigned identifiers for the Content

40 Ibid., p. 1-12.
41 Ibid., p. 1-8.
42 Ibid., p. 4-19.
43 Ibid., p. 1-13.

Information" and "provides identifiers that allow outside systems to refer, unambiguously, to a particular Content Information."[44]

- Provenance Information – "origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated."[45]
- Context Information – "documents the relationships of the Content Information to its environment" including "why the Content Information was created and how it relates to other Content Information objects."[46]
- Fixity Information – "documents the authentication mechanisms and provides authentication keys to ensure that the Content Information object has not been altered in an undocumented manner."[47]

PDI is an extension and elaboration of what the Task Force on Archiving of Digital Information called the "features that determine information integrity and deserve attention for archival purposes": "content, fixity, reference, provenance, and context."[48]

Over the past several years, the OAIS has become a widely accepted foundation for research and development on digital archives. The OAIS has become "the reference model of choice of those involved in digital preservation worldwide,"[49] serving as a "galvanizing force"[50] and a "major factor in the advancement of digital archiving efforts."[51] It has contributed "a common language and concepts for different professional groups involved in digital preservation and developing archiving systems"[52] and represents "common ground upon which to consolidate understanding of the needs and requirements of digital preservation."[53] Many types of institutions – including archives – were involved in its

44 Ibid., p. 1-12.
45 Ibid.
46 Ibid., p. 1-8.
47 Ibid., p. 1-10.
48 Task Force on Archiving of Digital Information, *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information* (Washington, DC, 1996), p. 12.
49 Daniel Greenstein and Abby Smith, "Digital Preservation in the United States: Survey of Current Research, Practice, and Common Understandings," in *New-Model Scholarship: How Will It Survive?* ed. Abby Smith (Washington, DC, 2003), p. 43.
50 Donald Waters, "Good Archives Make Good Scholars: Reflections on Recent Steps Toward the Archiving of Digital Information," in *The State of Digital Preservation: An International Perspective*, ed. Council on Library and Information Resources (Washington, DC, 2002), p. 80.
51 Gail M. Hodge, "Digital Preservation: Overview of Current Developments," *Information Services & Use*, vol. 22, nos. 2/3 (2002), pp. 73–82.
52 Neil Beagrie, *National Digital Preservation Initiatives: An Overview of Developments in Australia, France, the Netherlands, and the United Kingdom and of Related International Activity* (Washington, DC, 2003), p. 45.
53 Brian F. Lavoie, *The Open Archival Information System Reference Model: Introductory Guide* (Dublin, OH, 2004), p. 2.

development and have used it as a basis for their work.[54]

Although the OAIS is an excellent resource for the development of the Matrix, there are two reasons why it is not a sufficient one for the elaboration of the full range and depth of digital curation knowledge, skills, and capabilities. First, as a reference model, the OAIS has been purposely designed to express concepts and entities at a high level of abstraction, rather than elaborating details about any individual component or how specific tasks will be performed. For purposes of digital curation education, it is often important to drill down into activities in more detail. For example, we have included functions in the Matrix called "collaboration, coordination, and contracting with external actors" and "systems engineering and development," both of which are arguably represented in various parts of the OAIS but are not explicitly elaborated as a set of professional activities. Second, the scope of the OAIS is not inclusive of the full range of professional activities associated with digital curation. It focuses on an archives acquiring, managing, and providing access to digital resources, but it does not address many associated services and decision-making processes. Examples of core archival functions that we have added in order to supplement the OAIS include: advocacy and outreach; analysis and evaluation of the producer information environment; destruction and removal; reference and user support services; and selection, appraisal, and disposition.

We have benefited from a significant body of archival literature that addresses all of the above functions. The emergence of electronic records has been a catalyst for archivists to translate existing professional principles and heuristics into more formally detailed abstractions. Two research projects in the 1990s – one at the University of Pittsburgh[55] and the other at the University of British Columbia[56] – developed formalizations of established abstractions in the archival profession, particularly evidence, record, and authenticity. The former project produced functional requirements, production rules, and a metadata specification; the latter generated an entity model. The InterPARES (International Research on Permanent Authentic Records in Electronic Systems) 1

54 Christopher A. Lee, "Defining Digital Preservation Work: A Case Study of the Development of the *Reference Model for an Open Archival Information System*" (PhD diss., University of Michigan, 2005), pp. 142–46, 165–67, 180–82, 257–64, 287–94.

55 Wendy Duff, "Ensuring the Preservation of Reliable Evidence: A Research Project Funded by the NHPRC," *Archivaria* 42 (Fall 1996), pp. 28–45; Wendy Duff, "Harnessing the Power of Warrant," *American Archivist* 61 (Spring 1998), pp. 88–105.

56 Luciana Duranti, Terry Eastwood, and Heather MacNeil, *The Preservation of the Integrity of Electronic Records* (Vancouver, 1997); Luciana Duranti and Heather MacNeil, "The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project," *Archivaria* 42 (Fall 1996), pp. 46–67.

(1999–2001),[57] InterPARES 2 (2002–2007),[58] and InterPARES 3 (2007–2012)[59] projects have extended the work that began at the University of British Columbia, by continuing to investigate digital preservation issues, with a strong focus on the authenticity of electronic records. Several standards have also been developed to facilitate the design and management of "record-keeping systems," which ensure the authenticity of electronic records as evidence. One of the most visible standardization efforts in this area was a metadata scheme for the Commonwealth of Australia.[60] The Electronic Records Management Software Applications Design Criteria Standard (DOD 5015.02 – STD) provides a set of requirements for the design and certification of applications used to manage electronic records.[61] The first records management standard to have progressed through a major international standards development organization is ISO 15489 – "Information and Documentation – Records Management."[62]

A common theme of both the OAIS Reference Model and the archival literature on electronic records is the importance of attending not only to individual information objects, but also to the wider, socio-technical systems in which the objects are embedded. Within an archival context, this has meant a shift from focusing primarily on aggregations of records to placing much more emphasis on characteristics of, and requirements for, entire recordkeeping systems[63]; this is an important move, because care for digital records usually requires successful interactions with a much wider and more diverse set of system components – people, hardware, and software – than care for analogue records. Outside of the archival context, many other individuals responsible for supporting long-term access to digital information have recognized the importance of "system

57 Luciana Duranti and Kenneth Thibodeau, "The InterPARES International Research Project," *Information Management Journal*, vol. 35, no. 1 (2001), pp. 44–46, 48–50.

58 Luciana Duranti and Randy Preston, eds., *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records* (Rome, 2008).

59 Luciana Duranti, "An Overview of InterPARES 3 (2007–2012)," *Archives & Social Studies: A Journal of Interdisciplinary Research*, vol. 1, no. 1 (2007), pp. 577–603.

60 Glenda Acland, "The Australian Recordkeeping Metadata Schema – Version 1.0: Note from the Research Team," *Archivaria* 49 (Spring 2000), pp. 241–47; Sue McKemmish, Glenda Acland, and Barbara Reed, "Towards a Framework for Standardising Recordkeeping Metadata: The Australian Recordkeeping Metadata Schema," *Records Management Journal* 9 (1999), pp. 177–202; Sue McKemmish, Glenda Acland, Nigel Ward, and Barbara Reed, "Describing Records in Context in the Continuum: The Australian Recordkeeping Metadata Schema," *Archivaria* 48 (Fall 1999), pp. 3–43.

61 Assistant Secretary of Defense for Networks and Information Integration, *Electronic Records Management Software Applications Design Criteria Standard, DOD 5015.02 – STD* (Washington, DC, 2007).

62 International Organization for Standardization, *Information and Documentation – Records Management* (Geneva, 2001).

63 David Bearman, "Record-Keeping Systems," *Archivaria* 36 (Fall 1993), pp. 16–36.

thinking" to their work.[64]

As with our use of other sources, we have incorporated concepts, ideas, and terms from the archival literature in ways that are intended to be applicable across a wide diversity of digital curation contexts. We have avoided the assumption that digital curation will always 1) treat archival structures – records, series, fonds, recordkeeping systems – as fundamental entities or, 2) be driven by recordkeeping requirements. A great deal of digital curation activity takes place outside the context of the "archival enterprise."[65] We have developed the DigCCurr Matrix as a way of formally expressing the diversity of contexts in which digital curation can be carried out.

Matrix of Topics for a Digital Curation Curriculum

As a cornerstone of the DigCCurr work, we have developed a six-dimensional matrix for identifying and organizing the material to be covered in a digital curation curriculum. Development of both the Matrix and digital curation curriculum has been founded on several main principles, which we have elaborated in more detail elsewhere[66]:

- Build on an installed base.[67]
- Digital curation activities address the entire lifespan[68] of digital resources.[69]

64 See for example, Michael L. Brodie and Michael Stonebraker, *Migrating Legacy Systems: Gateways, Interfaces & the Incremental Approach* (San Francisco, CA, 1995); Willem-Jan van den Heuvel, *Aligning Modern Business Processes and Legacy Systems: A Component-Based Perspective* (Cambridge, MA, 2007).

65 David B. Gracy, II, "Our Future is Now," *American Archivist*, vol. 48, no. 1 (1985), pp. 12–21.

66 Christopher A. Lee, Helen R. Tibbo, and John C. Schaefer, "DigCCurr: Building an International Digital Curation Curriculum & the Carolina Digital Curation Fellowship Program," in *Archiving 2007: Final Program and Proceedings, May 21–24, 2007, Arlington, VA*, ed. Scott A. Stovall (Springfield, VA, 2007), pp. 105–109.

67 Susan Leigh Star and Karen Ruhleder, "Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces," *Information Systems Research*, vol. 7, no. 1 (1996), pp. 111–34.

68 A variety of disciplines use the term "life cycle" to designate all points in the "life" of information. The Matrix is not tied to a specific "records life cycle" (as opposed to a records continuum) orientation toward records activities. Following the lead of Frank Upward ("Modelling the Continuum as Paradigm Shift in Recordkeeping and Archiving Processes, and Beyond – a Personal Reflection," *Records Management Journal*, vol. 10, no. 3 [2000], pp. 115–39, see specifically 129–30), we have instead used the term "lifespan."

69 Chris Rusbridge, Peter Burnhill, Seamus Ross, Peter Buneman, David Giaretta, Liz Lyon, and Malcolm Atkinson, "The Digital Curation Centre: A Vision for Digital Curation," paper presented at From Local to Global: Data Interoperability – Challenges and Technologies, Mass Storage and Systems Technology Committee of the IEEE Computer Society, Sardinia, Italy, 20–24 June 2005, http://eprints.erpanet.org/82/01/DCC_Vision.pdf (accessed on 12 July 2011).

- Keep the lifespan stages simple, and move complexity into the functions.
- Build from modules, rather than entire courses.
- Emphasize core, generalizable modules.

Table 2 summarizes the six dimensions of the Matrix. A given curriculum unit can focus on a dimension in general or specifically as it intersects with one or more other dimensions. For example, one could teach a general unit on digital preservation (main considerations and practices), but one might also want to teach a unit specifically on the preservation of video, the preservation measures to be applied at the time of digital object creation, the preservation in a corporate recordkeeping context, or some combination thereof.

Table 2: Six Dimensions of the Matrix

Dimension	Explanation or Elaboration
1. Mandates, Values, and Principles	Core reasons why the digital curation functions and skills should be carried out and should serve as the basis for criteria to evaluate whether the digital curation activities have been carried out responsibly and appropriately.
2. Functions and Skills	"Know how," as opposed to the conceptual, attitudinal, or declarative knowledge.
3. Professional, Disciplinary, Institutional, Organizational, or Cultural Context	Understanding of challenges, opportunities, and characteristics of particular disciplines or institutions (e.g., social science data archive in a university, commercial data warehouse, state archives, serving a population with specific cultural norms).
4. Type of Resource	Types of resources that are the target of digital curation activities.
5. Instrumental Knowledge	Elements of knowledge that are instrumental to understanding and applying other aspects of the curriculum, including specialized terminology and characteristics of technologies.
6. Transition Points in the Information Continuum	Points of transition that span from pre-creation design and planning to secondary use environments.

The Matrix is a tool for thinking about, planning for, identifying, and organizing the digital curation curriculum. It is also helping us to address the issue of core versus specialized (optional) educational elements.

Dimension 1: Mandates, Values, and Principles

This is the first and most fundamental of the DigCCurr Matrix dimensions. The mandates, values, and principles are the core reasons why the digital curation functions and skills should be carried out, and should serve as the basis for criteria to evaluate whether the digital curation activities have been carried out responsibly and appropriately. Table 3 summarizes digital curation mandates, values, and principles. The items on the table are often made explicit through professional codes of ethics; industry and professional standards; laws and policies; and design principles.

Table 3: Dimension 1: Mandates, Values, and Principles

Mandate, Value, or Principlex	Explanation
Abstraction	Recognizing the value of both promoting and taking advantage of hiding complexity from view for particular purposes, while still acknowledging that the complexity exists and may need to be a focus of attention when engaging in other activities.
Accountability	Recognizing the unique and significant role that persistent information sources play in holding actors to account (both to themselves and to others) for their actions over time.
Adaptability	Recognizing and cultivating the ability of individuals to respond creatively to new and unexpected situations.
Authenticity	Understanding the importance of being able to assure users or other stakeholders that given digital resources are what they purport to be, and the circumstances under which one can and should appropriately provide such assurances.
Automation	Recognizing the value of automating tasks that can be automated, in order to focus human energy on those that cannot.

Chain of Custody	Recognizing the important roles that documenting and proving the movement and transformations of digital resources over time play in the assessment of trustworthiness, meaning, and evidential value of the resources.
Collection	Appreciating and understanding the value that digital resources can have when treated as aggregate units over time, rather than simply managing them as sets of discrete data elements.
Context	Paying serious attention to the context in which digital objects are created, managed, and used.[70] This includes: an appreciation for unexpected consequences; the importance of social context; the limits of technological determinism; and the ways in which values are always embedded in technology.
Continuum Approach	Holistic attention to the full span of design, creation, management, use, and reuse of digital objects, rather than fixating too heavily on only one part.
Critical Inquiry	Digging below the surface level of professional activities to understand why they are undertaken, and systematically determining whether and how they might be done better.
Diversity	Appreciating the types of situations in which heterogeneity (of perspectives, materials, systems) can be beneficial and desirable.
Encapsulation	Recognizing the value of both promoting and taking advantage of the loose coupling between two systems or sub-systems when data or control is always passed as discrete messages through well-defined interfaces.[71]
Evidence	Recognizing the importance of curating digital resources in ways that allow the resources to serve as proof over time.

70 Christopher A. Lee, "A Framework for Contextual Information in Digital Collections," *Journal of Documentation*, vol. 67, no. 1 (2011), pp. 95–143.

71 Abstraction is a very closely related concept to encapsulation, but the former is more conceptual and strategic, while the latter is more system-focused and tactical.

"Informating"	Attending to opportunities for changing and differently documenting activities in cases when carrying out those activities involves the exchange of symbolically encoded information through computer systems.[72]
Interoperability	Recognizing the value of both promoting and taking advantage of "the capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units."[73]
Long Term	Understanding and appreciation of the unique design considerations associated with systems and resources that need to persist for "a period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing user community, on the information being held in a repository."[74]
Modularity	Recognizing the value of both promoting and taking advantage of the decomposition of systems into components that are relatively self-contained and loosely coupled, particularly as these relate to reuse and implementation independence.
Open Architecture	Recognizing the value of both promoting and taking advantage of architectures that have been purposely designed and maintained to avoid dependencies on particular proprietary technologies or services.
Organizational Learning	Recognizing and cultivating the ability of organizations to respond creatively to new and unexpected situations, and to take further advantage of successful innovations.
Provenance	Recognizing the increasingly important role played within many contexts, by documenting various aspects of the origin of information.

72 Shoshana Zuboff, *In the Age of the Smart Machine: The Future of Work and Power* (New York, 1988), p. 9.

73 International Organization for Standardization, *ISO/IEC 2382-1:1993. Information technology – Vocabulary – Part 1: Fundamental terms* (Geneva, 1993).

74 *OAIS, p. 1-11.*

Robustness	Appreciating the importance of the essential combination of diversity plus redundancy plus multiple locations.[75]
Scale and Scalability	Attending to challenges and opportunities related to scaling approaches and systems, both up and down.
Significant Properties	Understanding that appropriate digital curation strategies depend upon identifying the properties of digital objects that relevant stakeholders or methods have identified as important to reproduce over time, because they impact "quality, usability, rendering, and behaviour"[76] in important ways.
Stakeholders	Identification and consideration of, planning for, and attending to, the interests of those who have a legitimate stake in one's digital curation decisions. Common types of potential stakeholders are: current direct users; current indirect users; currently underserved populations; likely future users; those whose lives are documented by digital materials; information producers; donors of content; current and potential funding entities (includes citizens); parents/guardians; other digital curation professionals.
Standardization	Recognizing the value of both promoting and taking advantage of formal, voluntary consensus processes for developing standards.

75 Andrew B. Hargadon, and Douglas Yellowlees, "When Innovations Meet Institutions: Edison and the Design of the Electric Light," *Administrative Science Quarterly*, vol. 46, no. 3 (2001), pp. 476–501; Stijn Hoorens, Jeff Rothenberg, Constantijn van Orange, Martijn van der Mandele, and Ruth Levitt, *Addressing the Uncertain Future of Preserving the Past: Towards a Robust Strategy for Digital Archiving and Preservation* (Santa Monica, 2007); Eric Matheson Leifers, *Actors as Observers: A Theory of Skill in Social Relationships* (New York, 1991); Urs von Burg, *The Triumph of Ethernet: Technological Communities and the Battle for the LAN Standard* (Stanford, 2001).

76 Margaret Hedstrom and Christopher A. Lee, "Significant Properties of Digital Objects: Definitions, Applications, Implications," in *Proceedings of the DLM-Forum 2002, Barcelona, 6–8 May 2002: @ccess and Preservation of Electronic Information: Best Practices and Solutions* (Luxembourg, 2002), pp. 218–27.

Sustainability	Appreciating the importance of, and understanding, the dynamics of "the set of business, social, technological, and policy mechanisms that encourage the gathering of important information assets into digital preservation systems, and support the indefinite persistence of digital preservation systems, enabling access to, and use of, the information assets into the long-term future."[77]
Trust	Appreciation for the ways in which professional authority depends upon legitimately gaining trust of those being served; general ways in which trust is gained and lost; and importance of attending not only to technical competence but also to "softer" considerations such as social capital, personal integrity, credibility, honesty, and care for the well-being of individuals.

There are several important points to note about this dimension. First, no such list will ever be exhaustive. The items come from a variety of sources and are the result of multiple iterations based on feedback from others; therefore, we believe they address many important considerations. However, readers are still likely to identify mandates, values, and principles that they believe should be added, or further distinctions that should be reflected in this dimension of the Matrix. We welcome suggestions for such revisions.

Second, we have not attempted – nor do we intend to attempt – a single hierarchical structure to organize the items on the list. There are numerous ways in which one could potentially nest items within other items. We have instead provided them as a single flat list.

A third point to emphasize is that we have not attempted to rank the items. Prioritization and degree of emphasis will depend on where one is working along the other Matrix dimensions. Likewise, different digital curation educational offerings and programs will have many reasons to place heavier emphasis on some items than others. In the curriculum developed in the DigCCurr project, we do address all of the items on the list to some degree.

Fourth, the line between mandates, values, and principles is often amorphous and difficult to define. Rather than trying to draw crisp boundaries between those three labels, we suggest a working definition for this dimension that

77 Brian Lavoie, Lorraine Eakin, Amy Friedlander, Francine Berman, Paul Courant, Clifford Lynch, and Daniel Rubinfeld, *Sustaining the Digital Investment: Issues and Challenges of Economically Sustainable Digital Preservation* (La Jolla, 2010), p. 19.

is roughly "that which digital curation professionals aspire to advance."

Finally, "within social structures, purposes are often formally enacted through functions; one often addresses and pursues the function itself, rather than directly referencing the purposes [for which] it is enacting (i.e., within a given social structure, the function effectively acts as the purpose)."[78] Understanding and appreciating the motivation behind specific, professional activities can be extremely important, particularly in changing environments, because the details of the activities are likely to change over time. However, there are always practical limitations on the degree to which any curriculum can 1) elaborate the reasons behind doing things in a certain way, or 2) elaborate the potential activities one might be called upon to carry out in order to advance a given mandate/value/principle. A sound pedagogical approach will combine these two factors in ways that make the mandates/values/principles tangible but also allow students to carry more "generalizable" lessons into other types of tasks.

Dimension 2: Functions and Skills

The elaboration of digital curation functions and skills was a major focus of the DigCCurr Matrix and curriculum development activities. This dimension addresses digital curation "know how," as opposed to the conceptual, attitudinal, or declarative knowledge that dominates several of the other matrix dimensions. Functions and skills are essential – though often quite challenging – for educators to address. We have identified twenty-four high-level functions or function categories. Each is then composed of many sub-functions. The main functions and first-level sub-functions are presented in Table 4.[79]

Table 4: Dimension 2: Functions and Skills

Function or Function Category	Definition/Explanation
Access	Making digital resources available to Consumers.
Administration	Control, coordination, and oversight of day-to-day digital curation operations.

78 Lee, "A Framework for Contextual Information in Digital Collections," p. 110.

79 A version of the table that lists sources of items is available from the DigCCurr project site, http://www.ils.unc.edu/digccurr/digccurr-matrix.html (accessed on 12 July 2011). The full matrix developed by the DigCCurr project further elaborates many sub-functions, drilling down as far as six levels in some cases and constituting several hundred total functions at various levels; Table 4, however, conveys only the top twenty-four functions.

Advocacy and Outreach	Activities aimed at influencing systems or behaviour outside the Archive.[80]
Analysis and Characterization of Digital Objects/ Packages	Identifying and documenting the properties of digital objects/packages that are relevant to the ongoing curtion and use of the objects/packages. This includes the identification of significant properties, which are "properties of digital objects that affect their quality, usability, rendering, and behaviour."[81]
Analysis and Evaluation of Producer[82] Information Environment[83]	This is often done in relation to known benchmarks or standards. It includes the assessments of record-keeping systems and the authenticity of documents within those systems. It can also include the analysis of work practices within the producer environment. Focus can be at the level of organization/institution, information system (e.g., recordkeeping system), collection, or individual items.
Archival Storage	"Services and functions used for the storage and retrieval of Archival Information Packages."[84]
Common Services	"Services such as inter-process communication, name services, temporary storage allocation, exception handling, security, and directory services necessary to support" digital curation.[85]

80 As discussed earlier, we have adopted many terms and definitions from the OAIS model. It defines "Archive" as "an organization that intends to preserve information for access and use by a Designated Community" (p. 1-8).

81 Hedstrom and Lee, pp. 218–27.

82 Recall that the OAIS defines Producer as "the role played by those persons, or client systems, who provide the information to be preserved. This can include other OAISs or internal OAIS persons or systems" (p. 1-12).

83 This function addresses analysis and evaluating of the current Producer information environment. This function and its sub-functions are strongly influenced by the DIRKS (Designing and Implementing Record Keeping Systems) methodology. Efforts to change or influence that environment are addressed in Advocacy and Outreach above.

84 OAIS, p. 1-8.

85 OAIS, p. 1-8. Common Services plays a unique role within both the OAIS and this taxonomy of functions. It includes a set of underlying technical services that are necessary for successful digital curation but are carried out by someone else (not directly by a digital curation professional). It is very important for digital curation professionals to be aware – at least at a basic level – of what the Common Services are, how they relate to each other, and some of the most viable ways for them to be provided (e.g., in order to evaluate, supervise, and contract for the services). However, Common Services are assumed, by definition, to fall outside the direct purview of digital curation. Common Services are not directly represented in Figure 1. Within the OAIS, they are "supporting services" that must be in place for computer systems to operate and perform properly. Although Common Services are

Collaboration, Coordination, and Contracting with External Actors	Initiation, management, and cultivation of relationships between the Archive and other entities in the environment (including other Archives).
Data Management	Design and maintenance of the intermediate data structures that are used to manage and provide basic access to digital data. Many of these activities have traditionally been the responsibility of database administrators, with the intermediate data structures being tables in relational databases. However, intermediate data structures in other data management layers/environments can also play a similar role in digital curation and require responsible management, e.g., file systems, Extensible Markup Language (XML) data elements, and catalogue data within data grids.[86]
Description, Organization, and Intellectual Control[87]	Development, capture, and management of descriptive information (DI), preservation description information (PDI), and packaging information (PI) associated with Archival Information Packages (AIPs). This is at a higher level of abstraction than both Data Management and Archival Storage. It ensures that the data associated with Content Information that is addressed in Data Management, Archival Storage, and Access is sufficiently detailed, complete, and accurate.[88]
Destruction and Removal	"The process of eliminating or deleting records beyond any possible reconstruction."[89]

necessary for an OAIS, they are not a major focus of the Reference Model, because they are "assumed to be available" (p. 4-2). Whenever it would be important for a digital curation professional to know how to actually carry out or provide a service him/herself, it should be elaborated in one of the other functions within this taxonomy. Note that all of the common services are addressed in detail in Characteristics of Technologies, which is part of the Instrumental Knowledge dimension of the DigCCurr Matrix.

86 Reagan W. Moore, "Building Preservation Environments with Data Grid Technology," *American Archivist*, vol. 69, no. 1 (2006), pp. 139–58.

87 The OAIS places "Add Descriptive Information" under Ingest. However, we have broken out description as its own function, in order to recognize that it can be done at many points in the lifespan of information. It is not only a part of the ingest process.

88 This is at a higher level of abstraction than both Data Management and Archival Storage. It ensures that the data associated with Content Information that is addressed in Data Management, Archival Storage and Access is sufficiently detailed, complete, and accurate. Note that Representation Information is considered part of the Content Information, and it is addressed primarily in Preservation Planning and Implementations.

89 International Organization for Standardization, *ISO 15489-1:2001:Information and docu-*

Identifying, Locating, and Harvesting	Identification, locating, and harvesting (i.e., "gathering up")[90] aggregates of resources, for purposes other than direct and immediate use of the resources.
Ingest[91]	"Services and functions that accept Submission Information Packages from Producers, prepares Archival Information Packages for storage, and ensures that Archival Information Packages and their supporting Descriptive Information become established within" an Archive.[92]
Management	Activities of the actor(s) who sets overall Archive mandate, policy, and resources "as one component in a broader domain of activity."[93]
Preservation Planning and Implementation	"Services and functions for monitoring the environment" and designing, recommending, and initiating strategies "to ensure that the information stored in the OAIS remains accessible to the Designated User Community over the long term, even if the original computing environment becomes obsolete."[94]
Production	Appropriate creation of digital objects/packages, either directly (i.e., born digital) or through the digitization of analogue materials.
Purchasing and Managing Licenses to Resources	Activities that ensure the appropriate and timely expenditure of financial resources for the software or data required for the curation of digital collections.
Reference and User Support Services	Direct engagement with Consumers[95] in order to help them find, make use of, make sense of, answer questions related to, or perform tasks that rely upon, curated information.

mentation – Records management – Part 1: General (Geneva, 2001), p. 2.

90 "Harvesting." *Oxford English Dictionary*, 2nd ed. (Oxford, UK, 1989).

91 The main conceptual boundary between Transfer and Ingest is: getting an object into the archives environment generally, which can include a staging area (Transfer), and the formal incorporation of the object as part of an AIP into the Archive (Ingest).

92 OAIS, p. 1-11.

93 Ibid.

94 Ibid., p. 4-2.

95 Recall that a Consumer is defined by the OAIS model as "the role played by those persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail. This can include other OAISs, as well as internal OAIS persons or systems" (p. 1-8).

Selection, Appraisal, and Disposition	Processes associated with determining what subsets of all possible digital information should be kept, how long they should be kept, and where they should be kept. This includes disposition, which is the determination that, at a particular time or upon the occurrence of a particular event, a digital object or set of digital objects should be either 1) removed out of an operational system and into another one, or 2) destroyed.
Systems Engineering and Development[96]	"Systems analysis and development work necessary for IT infrastructure development. It also lends technical assistance to … activities surrounding the acquisition, development, and deployment of advanced IT and communications systems."[97]
Transfer	Moving data from one environment into another.
Transformation of Digital Objects/ Packages	Activities that result in a "change of state information"[98] that is considered to be part of a digital object or package. For the purposes of digital curation, it is important to attend to 1) the ways in which and the extent to which transformations violate the integrity of state information; 2) whether or not a given transformation is reversible; 3) what transformations are most appropriate to apply at given points in a digital curation workflow; and 4) how to document the nature and rationale behind transformations.

96 The sub-functions are often presented as a serialized set of stages (as in the waterfall development model). While some sub-functions naturally follow from others, we are not assuming any particular development methodology. One could iterate through the sub-functions in whatever order and as many times as the organization deems necessary.

97 US National Archives and Records Administration (NARA), *ERA Program Management Information*, http://wayback.archive.org/web/*/http://archives.gov/era/program-mgmt.html (accessed on 13 July 2011).

98 Reagan Moore, "Towards a Theory of Digital Preservation," *International Journal of Digital Curation*, vol. 3, no. 1 (2008), pp. 63–75.

Use, Reuse, and Adding Value to Accessed Information[99]	Users acting upon information objects or packages (including after they have received DIPs). The Archive may provide support for use, such as tools that allow client-side visualization of data sets. Users may also provide value-added information (e.g., annotations or tagging), which the Archives then ingests to ensure persistent access to the information.
Validation and Quality Control of Digital Objects/ Packages	Identify component parts and ensure everything expected is present (e.g., compare to included definition file, "packing list," negotiated agreement, selection criteria).

In order for professionals to both perpetuate and continuously improve their work, they must engage in various second-order activities. In other words, they cannot simply "do digital curation" but must also reflect upon, question, document, investigate, and reconsider what they are doing. We have identified four meta-level functions, each of which can be applied to any of the functions listed above:

- Analysis and documentation of curation functions.
- Education and sharing of expertise or guidance on curation functions.
- Evaluation and audit of curation functions.
- Research and development to support curation functions.

We believe that it is important to explicitly reflect the above meta-level functions in order to ensure that they are not overlooked in a curriculum. Attention to the meta-level functions will help to instill the importance of life-long learning, professional development, professional service, organizational memory, and assessment.

99 Much of the value added by a user can be incorporated into the Description function (e.g., user annotations, user tagging of content). However, users can also add services that are incorporated into other functions.

Dimension 3: Professional, Disciplinary, Institutional, Organizational, or Cultural Context

Successful digital curation can require a strong understanding of the challenges, opportunities, and characteristics of particular types of work contexts (e.g., a social science data archive in a university, commercial data warehouse, state archives, serving a population with specific cultural norms). Table 5 summarizes professional, disciplinary, institutional, organizational, or cultural context.

Table 5: Dimension 3: Professional, Disciplinary, Institutional, Organizational, or Cultural Context

Work Context Categories and Elements	Explanation or Elaboration
Professional Contexts	Relatively bounded (but evolving) arenas of work activities that are widely recognized as falling within the responsibility of institutional structures called professions (e.g., librarianship, archival administration) – each having its own history, expectations, and mechanisms for engagement (e.g., professional associations, conferences, continuing education).
Disciplinary Contexts	Areas of study supported or represented by collections, e.g., history, physics.
Institutional or Organizational Contexts	Types of institutions or organizations, e.g., private vs. public sector, collecting repository vs. institutional archives.
Characteristics of Information and Record Creating Environments	Types of producers, e.g., paying corporate customers, private donors, scientific researchers transferring their data.
Cultural Contexts	"The distinctive ideas, customs, social behaviour, products, or way of life of a particular society, people, or period."[100]

100 "Culture." *Oxford English Dictionary*, 3rd ed. (Oxford, UK, March 2008), online version, June 2011, http://www.oed.com/view/Entry/45746 (accessed on 13 July 2011).

Dimension 4: Type of Resource

Table 6 summarizes various ways of characterizing types of digital resources. This is the fourth dimension of the DigCCurr Matrix. The categories are neither mutually exclusive nor exhaustive. They are intended to highlight "differences that make a difference" for digital curation professional education.

Table 6: Dimension 4: Type of Resource

Resource Type Categories and Elements	Explanation
Level of Aggregation	The span of materials upon which one is focused, which can range from components of digital objects (e.g., specific data elements or fields in databases) to files, objects, packages, or entire collections.
Level of Abstraction	The conceptual level upon which one is focused, e.g., work, expression, manifestation, or item.[101]
Medium	The physical carrier of the bits that constitute the digital resources.
Format	A particular way of encoding digital information (often at the file level), which conveys an associated set of behaviours and means of interpreting the content.
Genre	"Socially recognized types of communicative actions … that are habitually enacted by members of a community to realize particular social purposes."[102]

101 These four levels of abstraction are based on the Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records: Final Report*, Vol. 19, ed. Marie-France Plassard (München, 1998), http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records (accessed on 13 July 2011).

102 JoAnne Yates and Wanda Orlikowski, "Genres of Organizational Communication: A Structurational Approach to Studying Communication and Media," *Academy of Management Review*, vol. 17, no. 2 (1992), pp. 299–326.

Dimension 5: Instrumental Knowledge

Instrumental knowledge includes elements of knowledge that are instrumental to understanding and applying other aspects of the curriculum, including specialized terminology and characteristics of technologies. Table 7 summarizes the two main instrumental knowledge categories.

Table 7: Dimension 5: Instrumental Knowledge

Instrumental Knowledge Categories	Explanation
Terminology	All areas of the curriculum will introduce new terminology to students. This item in the table of topics is intended to call out fundamental terminology that might not be addressed elsewhere (for example, glossaries issued by professional associations such as the Society of American Archivists [SAA]).
Characteristics of Technologies	In order to successfully apply digital curation functions and skills, individuals must understand the characteristics, dynamics, limitations, and capabilities of the information and communication technologies (ICTs) associated with their work. This includes an understanding of the history and evolution of ICTs (including lessons about the previous trajectories, adoption, and influences of technologies over time); recognition of the ways in which technology is socially embedded and often unpredictable; essential characteristics and elements of the current and emerging ICT landscape; and the roles and types of ICT standards.

The elements of the Mandates, Values, and Principles dimension can help to answer the question: "Why was Alice's digital curation decision appropriate?" [answer: because it was based on principle X], and elements of the Functions and Skills dimension can help to answer the question: "How did Alice successfully carry out her decision?" [answer: by performing function Y]. However, it is elements of the Prerequisite Knowledge dimension that can help to answer the question: "What did Alice need to know before she could perform function Y?" [answer: definition of the term Z and characteristics of technology A]. The boundaries between the dimensions are somewhat permeable and subject to interpretation, but the basic idea is that knowing a given element of Prerequisite Knowledge is necessary "in support of" rather than "as a core and intrinsic of"

digital curation.[103] For example, it is often important for professionals working in the current technological environment to understand the basic characteristics of XML markup; but we would not argue that XML is in itself a defining element of what it means to do digital curation work.

Dimension 6: Transition Points in the Information Continuum

There are many potential points of transition for digital resources, from pre-creation design and planning, to secondary use environments. Table 8 elaborates seven major transition points of digital objects.

Table 8: Explanations of Transition Points

Transition Point	Explanation
Archival Custody (Preservation Environment)	Residing within an environment devoted to long-term preservation, such as a government archives, manuscript repository, or scientific data centre.
Creation	Point of generating a persistent and reproducible digital representation of information (either "born digital" or digitized from an analogue artifact).
Primary Use Environment (Active Use)	Environment that enables use of digital objects within a context that is substantially similar to the context within which they were created.
Pre-Creation Design and Planning	The arena of conceptual, strategic, and technical activities that together result in a system used to create digital objects.
Secondary Use Environment	A persistent environment where digital objects are managed outside of their primary use context, e.g., records centre, scientific data aggregator.
Transfer to Archives (Preservation Environment)	Point of movement across the "archival threshold"[104] can be relatively direct and instantaneous, or involve extensive operations and one or more "staging areas."

103 A reviewer of this article pointed out quite correctly that there are many areas of knowledge that are important to digital curation, such as "knowledge of the uses of information, the value of cultural resources, societal expectations of memory, ideas about persistence, concepts of archives and 'archiving', awareness of user needs, and user behaviour." These are all areas that we have addressed in other dimensions of the matrix. In other words, we have elevated all of them above the level of "merely instrumental" knowledge.

104 Luciana Duranti, "Archives as a Place," *Archives and Manuscripts* 24 (1996), pp. 242–55.

Transfer of Copies or Surrogates to Secondary Use Environment	Point of movement from the preservation environment to a secondary environment where the objects are managed; as with transfer to archives this can be relatively direct and instantaneous or involve extensive operations and one or more "staging areas."

It is important to note that there is no single assumed order in which transition points will occur; Table 8 simply lists them alphabetically. Figure 3 is a graphic representation of the major transition points. The arrows illustrate many – but not all – of the potential paths of information flowing through the transition points. Information flow can take the form of information packages that contain digital objects, the digital objects themselves, parts of digital objects, surrogates, and information about digital objects. Transition points are often where professional and ethical decisions are necessary. For example, at a given point: Does one need to capture a copy of the information to reflect the information's state (in archival terms, this often amounts to determining whether a new record has been created)? Are there distinct metadata elements that are or should be generated?; Does the transition point constitute a significant stage or event in the chain of custody that should be documented?; Who has permission to access the information?; Should use be logged? It is also often important for policies to determine whether movement of information constitutes a formal transfer of custody (and associated responsibility) or instead constitutes simply a "copy" of information, with primary curatorial responsibility remaining unchanged.

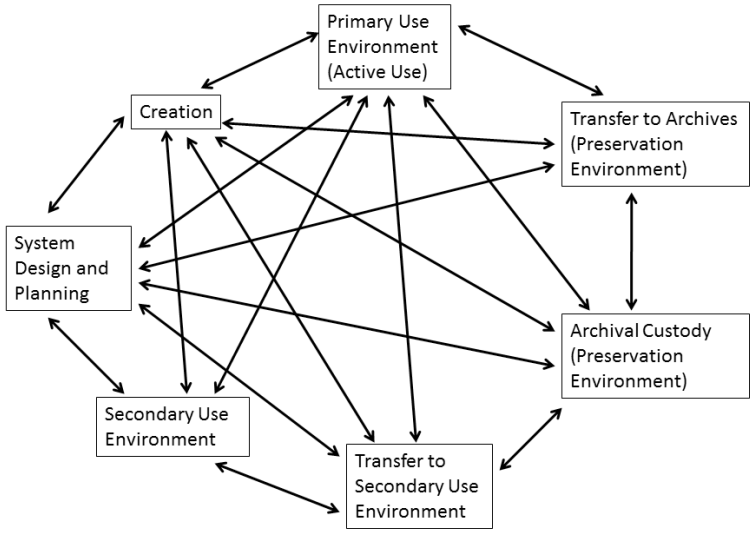

Figure 3: Dimension 6: Transition Points in the Information Continuum.

There are two transition points specifically devoted to transfer (to archives and to secondary use environment), because these are both transition points that often require specific "staging areas" for the purposes of storage and (often extensive) treatment of the digital objects. When the path into the archives or secondary use environment is relatively immediate and unmediated, then the "transfer" transition point is effectively not addressed as a distinct transition point (e.g., creator self-submission into an archive or secondary use environment for immediate access by others).

Not all digital objects will pass through all transition points (e.g., not all digital objects that are subject to digital curation requirements will be transferred to a separate archives). Transition points can also occur more than once (e.g., the same set of digital objects could be moved into many different secondary use environments). There is no single, linear order to the transition points; recursion, iteration, and parallel execution are common and often acceptable. Responsibility for specific transition points will often be shared, and responsibility can shift over time.

We have characterized archival custody as just one possible (and not necessary) transition point. There are two important reasons for this. First, as we discuss later in this article, our characterization of digital curation – including the digital curation work of records professionals – is postcustodial. Successful curation of digital resources involves a full range of activities that do not necessarily occur within the technical, institutional, or juridical boundaries of archival repositories. Custody is neither a necessary nor a sufficient condition for successful digital curation. Second, archival custody is only valuable to the extent that it can support the needs of stakeholders in the wider environment. By characterizing a formal preservation environment (which archivists characterize as custody) as one transition point within a larger network of transition points, we are highlighting the importance of its connections to other points. For example, acquisition of digital objects should be informed by their role within their creation environment; preservation and description should be attentive to specific user needs and secondary use environments; and characteristics of preservation environments should inform and reflect the design of the systems used to create digital objects.

Building the DigCCurr Matrix Functions into the Curriculum at UNC SILS

In addition to developing products to be used by digital curation educators, the DigCCurr project also developed and set the groundwork for a digital curation curriculum at UNC SILS. Our analysis of course syllabi at UNC SILS identified a large number of valuable elements relevant to digital curation that were already being taught in our school. We were also fortunate to have a complementary Digital Library Curriculum Project: a collaboration between Barbara Wilde-

muth and Jeffrey Pomerantz at UNC SILS, and Edward Fox at Virginia Tech.[105]

We continue to analyze the existing UNC SILS curriculum to determine opportunities and potential areas of revision or expansion. We also look to other course offerings at UNC and nearby institutions. For example, the ESOPI-21 project has provided us with significant exposure to both the public policy and administration offerings at UNC. In the past several years, we have translated our conceptual products into more specific decisions about course expectations for the Digital Curation Fellows, ESOPI Fellows, and other UNC SILS students interested in pursuing careers in digital curation. Based on the DigCCurr work, we have now instituted a Certificate in Digital Curation, which UNC graduate students can earn by taking an additional five courses beyond those required for their graduate degrees.

We are developing modules on specific topics; many are based on content developed by UNC SILS faculty, and several others are based on content developed by outside experts who have served as guest lecturers in specialized digital curation seminars that we have organized. We are beginning to share the modules through the Web. We have also developed entirely new classes to address areas of recognized need, based on the DigCCurr Matrix. We developed an introductory seminar for the Digital Curation Fellows (Fall 2008), two digital curation special-topics seminars (Spring 2007, Spring 2008), and a new course entitled "Understanding Information Technology for Managing Digital Collections" (Fall 2008). In collaboration with the Data Intensive Cyber Environments (DICE) group, which recently joined the UNC SILS faculty, we introduced a course on iRODS Rule Construction (Spring 2009), which has provided students and area professionals with hands-on experience with the development of rules-based data grid environments.

The DigCCurr grant funded five Carolina Digital Curation Fellows, who built on what they learned in the curriculum by participating in practical field experiences designed by a set of campus project partners. The field experiences provided Fellows with the opportunity to contribute to the management of a wide range of digital objects including public records, cultural heritage assets, teaching materials, and research data. The Fellowships helped us to integrate the curriculum and experiential components, advertise the existence of the program at UNC SILS, draw attention to the need for digital curation, and provide an essential empirical testing ground for the viability and appropriateness of the curriculum content in specific contexts.

105 Seungwon Yang, Barbara M. Wildemuth, Seonho Kim, Uma Murthy, Jeffrey P. Pomerantz, Sanghee Oh, and Edward A. Fox, "Further Development of a Digital Library Curriculum: Evaluation Approaches and New Tools," in *Asian Digital Libraries: Looking Back 10 Years and Forging New Frontiers: 10th International Conference on Asian Digital Libraries, ICADL 2007, Hanoi, Vietnam, December 10–13, 2007: Proceedings*, eds. Dion Hoe Lian Goh, Tru Hoang Cao, Ingeborg Sølvberg, and Edie Rasmussen (Berlin, 2007), pp. 434–43.

Highlighting Archival Capabilities

Introduction of the term digital curation led some to wonder if this was not just "digital archiving" given a new name and veneer. Elizabeth Yakel's definition of digital curation as "the active involvement of information professionals in the management, including the preservation, of digital data for future use,"[106] makes clear that the roles and professions involved extend well beyond the archival domain; yet the detailed articulation of digital curation activities reveals numerous opportunities to promote, advance, and further enhance the skills and capabilities of archivists.

In his keynote address to the 2007 European Conference on Digital Libraries (ECDL), Seamus Ross argued that, "as digital libraries are more akin to archives than they are to traditional libraries we need to seek their theoretical foundations in the domain of archival science and their practices in archival and records management environments."[107] He goes on to observe that, "[a]rchival science, with its principles of uniqueness, provenance, arrangement and description, authenticity, appraisal, and its tool sets such as diplomatics and palaeography, may offer us a framework for a theoretical foundation for digital libraries."[108] We extend that framework beyond digital libraries to the broader construct and set of functions that make up digital curation.

Archival Strengths within the Matrix

Many aspects of the DigCCurr Matrix intersect with long-standing archival principles, knowledge, and competencies. The Matrix also reflects areas of potential growth and collaboration.

Dimension 1

"Mandates, Values, and Principles," includes concepts that are central to archival science: authenticity, chain of custody, collection, context, continuum approach, evidence, long term, and provenance. It also includes computer, and information science and technology concerns such as automation, encapsulation, interoperability, modularity, open architecture, robustness, and scale. For digital curation efforts to be successful, regardless of the setting – library, archives, museum, or business – staff must have facility with a wide range of

106 Yakel, pp. 335–40.

107 Seamus Ross, "Digital Preservation, Archival Science and Methodological Foundations for Digital Libraries," Keynote Address at the 11th European Conference on Digital Libraries (ECDL), Budapest (17 September 2007), p. 19, http://eprints.erpanet.org/131/ (accessed on 13 July 2011).

108 Ibid.

terms, concepts and skills, and be able to traverse back and forth across the archival and curatorial perspective to the information technology domain. Both the ability to understand information technology colleagues, and to explain the importance of archival concepts to those individuals stand as some of the most important – and often challenging – aspects of the digital curation professional's job. Educational grounding and practical experience in applying traditional archival concepts are fundamental elements of this equation, but there is also a need for the digital archivist to proficiently navigate (and often actively channel) new technological waters.

Dimension 2

"Functions and Skills" presents several activities that readers are likely to recognize as being core to the archival enterprise. These include administration and management; appraising and selecting materials (identifying, locating, and harvesting; and selection and appraisal); acquisition of materials (ingest); providing access, reference, and user support; advocacy and outreach; storing materials; description, organization, and intellectual control. New skills that stretch most archivists from their traditional roles include purchasing and managing licenses to resources, systems engineering and development, transformation of digital objects/packages, and validation and quality control of digital objects/packages.

Digital Curation is to Digital Preservation as Records Continuum is to Records Life Cycle

We believe that expanding one's professional focus from digital preservation to the wider universe of digital curation is similar and complementary to the moves within the archival profession toward both postcustodial and continuum approaches. More than thirty years ago, Gerald Ham argued that archivists need to think of their professional role much more broadly than simply managing physical artifacts.[109] In contrast to an anti-custodial position, which asserts that archivists should not take custody of electronic records, the postcustodial orientation advocated by Ham and others[110] is an attempt to place custodial activities within a larger context of professional services and commitments. Fixing the archival gaze solely on what happens within the walls of repositories runs the risk of neglecting other essential aspects of the archival enterprise, including

109 F. Gerald Ham, "Archival Strategies for the Post Custodial Era," *American Archivist* 44 (1981), pp. 207–16.

110 See for example, Terry Cook, "The Concept of the Archival Fonds: Theory, Description, and Provenance in the Post-Custodial Era," in *The Archival Fonds: From Theory to Practice*, ed. Terry Eastwood (Ottawa, 1992), pp. 31–85.

engagement with creators and users of records, advocating for archival priorities, influencing policies, and exploring connections across a diversity of collections. The DigCCurr Matrix has been consciously designed to reflect a postcustodial perspective. The functions of the OAIS Reference Model, for example, provide a valuable foundation for many of the digital curation functions represented in the Matrix; the Matrix, however, goes beyond those functions in order to reflect various forms of active engagement with producers, users, and other related stakeholders.

Contemporary scholarship about electronic records has also been strongly influenced by the Australian continuum model of recordkeeping,[111] which has both advanced and enhanced postcustodialism by expressing how recordkeeping should not be seen as a set of activities that begin at the moment that records are transferred to an archival repository and end at the moment that records enter the hands of users. Instead, recordkeeping involves ongoing care based on continuing value that can and should occur at many points throughout (and even before) the existence of records. Terry Cook has lauded the Australian orientation toward recordkeeping, because it "sanctions a potentially powerful strategy to get archival issues addressed by record creators at the front end of the records continuum, which is essential if an archival record is to survive in the electronic era."[112] Transfer to an archive (crossing the "archival threshold") is an important moment, for many professional and institutional reasons, but – in terms of the DigCCurr Matrix – it is not the only transition point that matters.

A practical implication of postcustodial and continuum perspectives has been a strong emphasis in both the archival literature and professional activities of archivists on understanding and engaging records creating environments, rather than waiting passively for records to eventually cross the archival threshold. A similar approach is reflected by numerous digital curation education and research initiatives – in North America, the UK, and elsewhere – that focus on active engagement with the producers of digital materials, whether they are scientific researchers, artists, government officials, or others whose activities are important to document over time.

111 Frank Upward, "Structuring the Records Continuum Part One: Post-Custodial Principles and Properties," *Archives and Manuscripts*, vol. 24, no. 2 (November 1996), pp. 268–85; Jay Atherton, "From Life Cycle to Continuum: Some Thoughts on the Records Management–Archives Relationship," *Archivaria* 21 (Winter 1985–86), pp. 43–51; Sue McKemmish, "Placing Records Continuum Theory and Practice," *Archival Science*, vol. 1, no. 4 (2001), pp. 333–59.

112 Terry Cook, "What Is Past Is Prologue: A History of Archival Ideas since 1898, and the Future Paradigm Shift," *Archivaria* 43 (Spring 1997), p. 40.

Opportunities to Advance the Archival Enterprise

Digital curation is an area of work that has been emerging for several decades, through the distinct and largely uncoordinated efforts of actors within many existing streams of activity. Each stream has come to recognize elements of this work as distinct and legitimate, but the elements have usually not been considered part of the core of any existing professions, institutions, or disciplines. During the mid- to late-1990s, actors involved in the separate streams of activity increasingly began to recognize that there were many places where the streams intersected. The efforts of archivists have taken place within this context.

Several major factors have defined the particular archival niche within this landscape, including a heavy focus on the definition of "the record" as distinct from other types of digital resources; dual origins of the profession in both librarianship and history; and ambiguity on the issue of whether archivists should take custody of electronic records.

If archivists – who are increasingly called upon to address electronic records – are to understand the history of their own profession, it is important to recognize many complementary efforts. According to the SAA's *Guidelines for a Graduate Program in Archival Studies*:

> Archivists, like all professionals, must rely on knowledge, methods, and perspectives from beyond their own discipline. The interdisciplinary nature of archival studies arises from the complexity of the records and papers, the contexts of their creation, the multiplicity of their potential uses, and from the many roles that archivists fill. Archivists need to be knowledgeable about significant theories, methods, and practices of some or all of these fields.[113]

If archivists are to act on the frequent calls to collaborate with allied professionals, it is also essential to understand what those other professionals have to offer archivists and what distinct value archivists are likely to bring to the relationships. In short, archivists' exploration of their own profession should reflect one of their most cherished values: understanding and reflection of context. The "allied professions" listed in the SAA Guidelines are "library and information science, museology, oral history, historic preservation, and historical editing." We would contend that archivists responsible for electronic records must cast this net much differently, if they are to understand and appropriately navigate this space.

The development of professional strategies can benefit from the drawing together of many pre-existing elements of the surrounding environment. Lawrence Lessig explains how even the most creative acts involve elements of

113 Society of American Archivists, *Guidelines for a Graduate Program in Archival Studies* (Chicago, 2002).

reuse.[114] Other authors have emphasized the value of innovations that combine and recombine existing artifacts and social structures.[115] Such activities are not a simple matter of mindless copying, but instead require the actors involved to attend to various barriers and facilitators to the reuse of artifacts, concepts, and actors from the environment.[116] Several studies of knowledge transfer suggest that groups are receptive to learning from others only during specific (often short) periods in the groups' life cycle, with the beginning of that life cycle being a particularly receptive time.[117] The time is ripe for archivists to both benefit from the expertise of others and clearly demonstrate what they have to offer to others.

There are many ways in which archivists could decide to navigate the digital curation landscape. Technologies, requirements, and opportunities will continue to evolve, so no strategy is immune to revision. Two factors that can ground these decisions are a sense of the professional values that archivists hope to advance[118] and the particular benefits that they can offer. Archivists have long-standing experience with managing materials as aggregates, conveying contextual information, documenting provenance, engaging with creators, and attending to projected secondary use needs. These are all considerations that have become widely recognized in the interdisciplinary space of digital curation. The opportunities for archivists to establish fruitful partnerships and collaborations are immense. Seizing the opportunities will require respectful engagement with experts and sources from numerous activity streams, many of which fall outside of archivists' traditional lists of allies. The result will be a

114 Lawrence Lessig, *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity* (New York, 2004), pp. 21–30.

115 Joseph Alois Schumpeter, *Capitalism, Socialism, and Democracy* (New York, 1942); Richard R. Nelson and Sidney G. Winter, *An Evolutionary Theory of Economic Change* (Cambridge, MA, 1982); Thomas Parke Hughes, *Networks of Power: Electrification in Western Society, 1880–1930* (Baltimore, 1983).

116 Linda Argote, "Organizational Memory," in *Organizational Learning: Creating, Retaining, and Transferring Knowledge* (Dordrecht, 1999), pp. 67–97; Linda Argote, Sara L. Beckman, and Dennis Epple, "The Persistence and Transfer of Learning in Industrial Settings," *Management Science*, vol. 36, no. 2 (1990), pp. 140–54; Linda Argote, Paul Ingram, John M. Levine, and Richard L. Moreland, "Knowledge Transfer in Organizations: Learning from the Experience of Others," *Organizational Behavior and Human Decision Processes*, vol. 82, no. 1 (2000), pp. 1–8; Gabriel Szulanski, "Exploring Internal Stickiness: Impediments to the Transfer of Best Practice within the Firm," *Strategic Management Journal*, vol. 17, Winter Special Issue (1996), pp. 27–43; Gabriel Szulanski, "The Process of Knowledge Transfer: A Diachronic Analysis of Stickiness," *Organizational Behavior and Human Decision Processes*, vol. 82, no. 1 (2000), pp. 9–27.

117 Argote, Beckman, and Epple, pp. 140–54; Joel A. Baum and Paul Ingram, "Survival-Enhancing Learning in the Manhattan Hotel Industry, 1898–1980," *Management Science*, vol. 44, no. 7 (1998), pp. 996–1016; Marci J. Tyre and Wanda J. Orlikowski, "Windows of Opportunity: Temporal Patterns of Technological Adaptation in Organizations," *Organization Science*, vol. 5, no. 1 (1994), pp. 98–118.

118 Society of American Archivists, *Core Values of Archivists* (May 2011), http://www2.archivists.org/statements/core-values-of-archivists (accessed on 13 July 2011).

whole new set of settlements in the system of professions[119] associated with the curation of digital materials.

Digital Archiving Continuing Education (DACE)

The DigCCurr Matrix of Knowledge and Skills was developed within the context of graduate education. While we are preparing an emerging generation of digital curation professionals (including archivists responsible for digital materials) within schools such as UNC SILS, a large cadre of established professionals who cannot go back to universities for certificates of advanced study or other degrees, need robust continuing education programs through which they can acquire the skills and knowledge to manage and preserve digital materials. UNC SILS, through IMLS funding, is filling some of this need for continuing education with the annual, week-long DigCCurr Professional Institute.[120] We also adapted the Institute materials to offer a more focused workshop entitled "An Introduction to Digital Curation for Public Records Professionals" in association with the joint Annual Meeting of the National Association of Government Archives and Records Administrators (NAGARA) and the Council of State Archives (CoSA) in July 2011. The Digital Preservation Outreach and Education (DPOE) Program of the Library of Congress has been surveying needs, assembling a calendar of training opportunities, and is starting to prepare potential instructors and outreach materials for working professionals.[121] Of particular note, the SAA is currently launching the Digital Archiving Specialist (DAS) curriculum and certificate program.

In 2010, the SAA formed the Digital Archives Continuing Education Task Force (DACE), to be "responsible for developing a detailed professional development curriculum on the subject of digital archives."[122] Significantly, the Task Force agreed to focus on born-digital records in recognition of the centrality of electronic records to the archival enterprise today. They also argued that

119 According to Andrew Abbott (*The System of Professions: An Essay on the Division of Expert Labor* [Chicago, 1988], pp. 69–79), "[t]he claim to full and final jurisdiction is only one of the possible settlements of jurisdictional dispute," i.e., arrangements for the division of labour across different groups related to a particular type of work. There are at least five other forms of settlement: 1) subordination of one under the other; 2) division of labour; 3) "intellectual jurisdiction," in which one profession retains control of cognitive elements but allows others to engage in the work's practice; 4) one profession has advisory control over some aspects of the work; and 5) division according to the nature of the client.

120 See http://ils.unc.edu/digccurr/institute.html (accessed on 15 July 2011).

121 See the Digital Preservation Outreach and Education (DPOE) website, http://www.digitalpreservation.gov/education/ (accessed on 15 July 2011).

122 See the Digital Archives Continuing Education (DACE) Task Force from the Society of American Archivists Council, 26 May 2010, http://saa.archivists.org/Scripts/4Disapi.dll/4DCGI/committees/SAATF-DACE.html?CommCode=SAA**TF-DACE (accessed on 13 July 2011).

electronic records training needs to be integrated into the work of archives rather than treated as something separate from, or in addition to, that work. The Task Force used the DigCCurr Matrix to shape the resulting Digital Archives Specialist (DAS) curriculum, which incorporates much from the inclusive perspectives seen in digital curation, postcustodialism, and the continuum model, as well as content from other disciplines.

We hope that the DAS curriculum will help to move many of the considerations outlined above from the DigCCurr Matrix into the mainstream of professional archival education. The curriculum and the associated Digital Archiving Specialist certificate were created in response to one of the most challenging issues identified in SAA's strategic plan: "Rapidly changing information technologies challenge archival principles, practices, and communication protocols, demanding effective leadership from the archives community to access, capture, and preserve records in all formats." This "Technology" priority carries with it the following objective: "SAA will provide education and training to its members to ensure that they are aware of relevant standards and adopt appropriate practices for appraising, capturing, preserving, and providing access to electronic records."[123]

The DAS curriculum has several learning objectives that resonate with the DigCCurr Matrix. Upon completion of the curriculum, students will be able to:

- Understand the nature of records in electronic form, including the function of various storage media, nature of system dependence, and the effect on integrity of the records over time.
- Communicate and define requirements, roles, and responsibilities related to digital archives to a variety of partners and audiences.
- Formulate tactics and strategies for the appraisal, description, management, organization, and preservation of digital archives.
- Integrate technologies, tools, software, and media within existing functions for the appraisal, capture, preservation, and access to digital collections.
- Plan for the integration of new tools or successive generations of emerging technologies, software, and media.
- Curate, store and retrieve original masters, and access copies of digital archives.
- Provide dependable organization and service to designated communities across networks.[124]

123 To view the strategic plan, see http://www2.archivists.org/sites/all/files/0511-StratPlan_PublicPosting_060111.pdf (accessed on 13 July 2011).

124 Action Item: Digital Archives Specialist Curriculum and Certificate Program (Geof Huth, *DACE Task Force Report*; Council Action Item prepared by Solveig De Sutter), 11 May 2011, p. 14, http://www.archivists.org/council/Council0511/0511-III-D-DAS.pdf (accessed on 13 July 2011).

Moreover, the DACE Task Force notes in its report to the SAA Council that "[t]o provide dependable organization and service, Digital Archives Specialists must be able to integrate new tools within existing functions to implement strategies serving the needs of designated communities across networks. Digital Archives Specialists must be able to appraise, capture, preserve, and make materials accessible through successive generations of emerging technologies, software, and media."[125] These are issues stressed throughout the DigCCurr Matrix and especially in the notion of transition points in the lifespan of digital objects.

The DACE Task Force made several additional recommendations to shape the program. One of the recommendations is to build a structured curriculum involving four tiers of study: foundational courses, tactical and strategic courses, tools and services courses, and transformational courses. Foundational courses "focus on the essential skills that archivists will need to manage digital archives. These focus primarily, but not exclusively, on the needs of practitioners."[126] Tactical and strategic courses "focus on the skills archivists need to make significant changes in their organizations so that they can develop a digital archives and work seriously on managing electronic records. These focus primarily, but not exclusively, on the needs of managers."[127] Tools and services courses "focus on specific tools and services that archivists need to use for their work with digital archives. These are practical courses focused on specific software products and other tools. These courses focus primarily, but not exclusively, on the needs of practitioner archivists."[128] Transformational courses "focus on the skills archivists need to change their working life dramatically and transform their institutions into full-fledged digital archives. These courses focus primarily, but not exclusively, on the needs of administrators."[129]

The DACE Task Force also recommended that there be Tracks of Study for various audiences including Practitioner, Manager, Administrator, IT Professional, Librarian, Legal Professional, and Records Manager. While the task force was charged with developing a Digital Archiving Specialist Curriculum, these tracks of study, aimed at a wide array of individuals, speak to the necessity for collaboration across professional lines and much of the material highlighted in the DigCCurr Matrix. Participants who complete a specified number of requirements will earn a Digital Archives Specialist (DAS) certificate. The DAS certificate requires evidence that students have knowledge of technical standards and of core archival activities as they relate to digital archives. This combination very much parallels the knowledge, attitudes, functions, and skills

125 Ibid., p. 14.
126 Ibid., p. 6.
127 Ibid.
128 Ibid., p. 7.
129 Ibid., p. 8.

outlined in the DigCCurr curriculum. Built on the foundation of long-standing archival principles, the DAS Program will evolve as new technologies emerge and repositories change. Here the world's largest professional archival organization is embracing the elements of digital curation while foregrounding the contributions of archivists.

Conclusion

Professional education is a process that is never completed. This is particularly true in the dynamic and rapidly evolving field of digital curation. As we learn together what digital curation means and what is required to do it well, we look forward to continuing engagement with students, scholars, and allied professionals. The DACE recommendations and curricular framework move the digital archivist into the digital curation space. What evolves will be tested in the repository workplace from archives to libraries to corporate data centres. All of this will test the principles, structure, and robustness of the DigCCurr Matrix, and provide feedback for future development of graduate and professional archival and digital curation education.