Defining What Digital Curators Do and What they Need to Know: The DigCCurr Project

Christopher A. Lee, Helen R. Tibbo, John C. Schaefer
School of Information and Library Science
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599
+ 1-919-962-7024
[callee, tibbo, jschaefe@ils.unc.edu]

ABSTRACT
The DigCCurr (Digital Curation Curriculum) project is developing a graduate-level curricular framework, course modules, and experiential components to prepare students for digital curation in various environments. This paper summarizes an initial draft and guiding principles behind a matrix of digital curation knowledge and competencies, which are acting as the basis for our curriculum design efforts.

Categories and Subject Descriptors
K.3.2. Computer and Information Science Education. K.7. The Computing Profession. K.7.1. Occupations.

General Terms
Measurement, Documentation**,** Design

Keywords
Digital archives, digital curation, digital preservation, professional education, functional model.

1. INTRODUCTION
Cultural heritage, science, commerce, education and government services all depend upon long-term access and use of meaningful and authentic digital resources. This will require concerted, appropriate digital curation efforts. Several disciplines and professions have developed de facto practices and expertise in aspects of digital curation. Professional education for digital curation has generally involved on-the-job training and experimentation, possibly supplemented by workshops lasting a few days.

The IMLS-funded DigCCurr (Digital Curation Curriculum) project [2] is a collaboration of the School of Information and Library Science at the University of North Carolina at Chapel Hill and the National Archives and Records Administration. We are developing a graduate-level curricular framework, course modules, and experiential components to prepare students for digital curation in various environments. DigCCurr brings together an Advisory Board of experts from Australia, Canada, Italy, the Netherlands, New Zealand, the United Kingdom and the United States. The grant includes funding for a set of Digital Curation Fellows, who will pursue degrees at SILS beginning in fall 2007. Repositories at UNC will provide practical experience opportunities for the Fellows. Two international symposia will bring the issues of digital curation and this curriculum to the attention of librarians, archivists, museum professionals, data curators, scholars, and the general public.

2. MATRIX OF DIGITAL CURATION KNOWLEDGE AND COMPETENCIES
Based on a review of existing literature, syllabi, job descriptions, and semi-structured interviews with the members of our Advisory Board, we have developed a 6-dimensional matrix for identifying and organizing the material to be covered in a digital curation curriculum. Each unit of curriculum content will address one or more dimensions. The six dimensions are:

1. Type of Resource
- Level of Aggregation
- Level of Abstraction
- Medium
- Format
- Genre

2. Functions and Skills [see below]

3. Professional, Disciplinary or Institutional/Organizational Context
- Professional Context
- Disciplinary Context
- Institutional/Organizational Context

4. Mandates, Values and Principles
- Ethics
- Legal Requirements
- Standards
- Interoperability and Sustainability Requirements

5. Prerequisite Knowledge
- Terminology
- Characteristics of Technologies

6. Lifecycle Stage
- Pre-Creation Design and Planning
- Creation
- Primary Use Environment (Active Use)
- Transfer to Archives
- Archives (Preservation Environment)
- Transfer Copies or Surrogates to Secondary Use Environment
- Secondary Use Environment

A 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 selection and appraisal (main considerations and practices), but one might also

want to teach a unit specifically on selection of web pages, selection during the active use stage, selection within a corporate recordkeeping context, ethical values at stake in selection, or some combination thereof.

3. GUIDING PRINCIPLES

Our design of both the matrix and curriculum units is guided by a set of basic principles.

Principle 1: Build on an installed based

We are drawing heavily from both existing and emerging models, frameworks, guidance documents and educational initiatives.

Principle 2: Digital curation activities span the entire lifecycle of digital resources

Responsible digital curation is much more than preservation of bits. It reaches into both the producer and user environments associated with collections of digital objects. Several members of our Advisory Board have suggested that one of the major benefits of adopting the label "digital curation" is that it foregrounds the importance of issues and activities that fall outside the scope often associated with "digital preservation."

Principle 3: Keep lifecycle stages simple, and move complexity into the functions

We have identified and analyzed numerous lifecycle models. The lifecycle models often include considerable detail about the processes and workflow of a particular domain or context. We have attempted to identify a high-level set of lifecycle stages for two reasons. First, this allows us to map our lifecycle stages to existing domain-specific lifecycle models. Second, many activities that are important to digital curation (e.g. classification, metadata extraction, characterization, file format transformations) can occur at more than one stage of the information lifecycle. For example, many of the activities that the Reference Model for an Open Archival Information System (OAIS) [1] places within the Ingest function can be important and valuable to carry out, not only during transfer to an archives, but also during system design, creation, active use, within the preservation environment, during transfer to a secondary use environment and within the secondary use environment.

Principle 4: Build from modules, rather than entire courses

We have identified several categories of digital curation knowledge and competency that should be addressed within a curriculum, but (1) do not warrant an entire course and (2) may not fit well into existing courses or areas of instructor competence. Two such categories that have emerged from our analysis are legal requirements (e.g. compliance with Sarbanes-Oxley, Freedom of Information Act) and understanding technologies (e.g. basic data structures, operating systems, storage environments). We have been exploring approaches for integrating small modules into the curriculum, which address such particular topics. This can be a challenge in many university environments, in which the 3- or 4-credit-hour course is the basic unit for administering curricula.

Principle 5: Emphasize core, generalizable modules

If each of the six dimensions in the matrix had 10 categories of values (rows), there would be one million individual cells. Obviously, no curriculum could devote attention to all of the cells individually. The aim should be to teach units that address a large number of cells simultaneously, while also supporting a drilling down into specific cells, when necessary.

4. FUNCTIONS AND SKILLS

During the first phase of DigCCurr, we have devoted considerable attention to the development of a taxonomy of functions and skills. This is an iterative process and is still likely to undergo significant revision. The following characterizes the current draft of our taxonomy.

First-Level Functions and Skills:

Access
Administration
Advocacy and Outreach
Analysis and Characterization of Digital Objects/Packages
Analysis and Evaluation of Producer Information Environments
Archival Storage
Collaboration, Coordination and Contracting with External Actors
Common Services
Data Management
Description, Organization and Intellectual Control
Identifying, Locating and Harvest
Ingest
Management
Preservation Planning and Implementation
Production
Reference and User Support Services
Removal from Archive
Selection and Appraisal
Systems Engineering and Development
Transfer
Transformation of Digital Objects/Packages
Use, Reuse and Adding Value by User

Meta-Level Functions and Skills:

Analysis and Documentation of Archives Functions
Evaluation and Audit of Archives Functions
Research and Development to Support Archives Functions

5. ACKNOWLEDGMENTS

This work was supported through IMLS Grant # RE-05-06-0044. Thanks to all the partners and advisory board members of the DigCCurr project. Thank you also to Timothy (Pete) Ramsey for alerting us to the EBSCO lifecycle.

6. REFERENCES

[1] ISO 14721. 2003. Space data and information transfer systems – Open Archival information system – Reference Model.

[2] Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum. http://www.ils.unc.edu/digccurr.