Bringing Values to the Bitstream: A Framework for Digitally-Aware Professional Ethics of Curation

CHRISTOPHER A. LEE

Abstract: 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 curation of digital collections, they will increasingly discover that established sources of guidance suffer from "latent ambiguities" (Lessig, 1999). Digital resources are composed of interacting components that can be considered and accessed at different levels of representation (e.g. bitstream, through a filesystem, files as rendered through specific applications, records composed of multiple files, abstract "works"). To ensure integrity and future use, archivists and other information professionals must make decisions regarding treatment of materials at multiple levels of representation.

Digital curation professionals are faced with many new decisions, which require an understanding of underlying digital representations, in order to appropriately enact professional values. For example, when acquiring a disk as part of a collection, should an archivist create a bit-level image of the disk, in order to ensure the potential to recreate not only all the data but also system and program files? Should she retain "hidden" data in a Word document or only retain what she assumes to be the text that the author intended? If the disk includes a Microsoft Outlook .pst file (including saved and sent messages, calendar items, draft and deleted messages, address book, and possibly viruses), should she retain the whole .pst file, or simply extract messages and attachments that were sent and received?

This paper reports on a framework for ethics of digital curation that is designed to reflect both well-established professional values and the representational complexity of digital collections. I will report on development of the framework and efforts to incorporate it into professional education.

About the author:

Christopher (Cal) Lee is Assistant Professor at the School of Information and Library Science at the University of North Carolina, Chapel Hill. He teaches archival administration, records management, digital curation, resource selection and evaluation, understanding information technology for managing digital collections, and the construction of policies and rules for digital repository.

His primary area of research is the long-term curation of digital collections. He is particularly interested in the professionalization of this work and the diffusion of existing tools and methods (e.g. digital forensics, web archiving, automated implementation of policies) into professional practice.

Cal is editing and providing several chapters to a forthcoming book entitled, *I, Digital: Personal Collections in the Digital Era*.

Current projects include DigCCurr and DigCCurrII (http://ils.unc.edu/digccurr/), which are developing and implementing courses of study and practical engagement opportunities in digital curation. Cal has developed an extensive Matrix of Digital Curation Knowledge and Competencies (http://www.ils.unc.edu/digccurr/products.html), based on various data sources and grounded in professional literature. He is one of the lead organizers and instructors for the DigCCurr Professional Institute (http://ils.unc.edu/digccurr/institute.html).

The VidArch project (http://ils.unc.edu/vidarch/) investigated the collection of online video, with a particularly emphasis on contextual information. Cal's contributions included an information model for contextual information in digital collections (http://sils.unc.edu/research/publications/reports/TR_2007_04.pdf) and several empirical studies of online selection and collecting strategies. In collaboration with members of the Data-Intensive Cyber Environments (DICE) and VidArch project team, he is investigating various strategies for enhancing the sustainability of VidArch's products through the use of iRODS (Intergrated Rule-Oriented Data System).

Past research projects have included CAMiLEON (http://www.ils.unc.edu/callee/dig-pres_usersperspective.pdf), which examined migration vs. emulation as digital preservation strategies; and an in-depth case study of the development of the OAIS (http://www.ils.unc.edu/callee/dissertation_callee.pdf).

Cal has an MSI (with a concentration in Archives and Records Management) and PhD from the University of Michigan School of Information.