Digitization as Information Practics

Zack Lischer-Katz[1]

[1] Rutgers University

Abstract

This exploratory study sets out to better understand the information practices of academic library staff members engaged in digitization projects. Using semi-structured interviews and workplace observation, the research considers digitization within the framework of practice theory. Analysis of the data suggests that the use of information sources and the role of embodied knowledge in digitization work depends on the relationship between the media formats being digitized and the level of development of standards and policies for digitizing those formats. The research also suggests a variety of new avenues of inquiry for further conceptualizing digitization as information practice in archives and libraries, including the relationship between aesthetics and situated judgment in the creation of digitized library materials, and in the interplay between standards-following and improvisation supported by self-documentation.

Keywords: digitization, academic libraries, standards, practice theory

Contact: zack.lischerkatz@rutgers.edu

1   Introduction

Digitization projects are currently being conducted on a massive scale, with projects like the Library of Congress's American Memory project, the Google Books project, and the National Jukebox project (containing over 10,000 digitized 78 RPM recordings from years 1900 to 1925) offering expanded access to collections of various types of previously restricted materials via digital library systems. Scholarly access to preserved documents has been greatly enhanced through digitization, yet the sites where digitization take place have seen little in-depth qualitative analysis. This research project uses practice theory and qualitative methods to focus on one key site of digital production, a digitization lab at a research library (at a large university in the Northeastern United States).

2   Related Research

Past research on digitization practice has typically employed case studies of practical applications (e.g. Berger, 1999; Capell, 2010; Evens and Hauttekeete, 2011), or involved recommendations for standards development (e.g., Fleischhauer, 2010; Plichta and Kornbluh, 2002; Teper and Shaw, 2011). Lopatin's (2006) survey of literature on digitization projects from 2000-2005 suggests that studies on digitization tend to focus on such pragmatic issues as project management, funding sources, selection of materials for digitization, legal issues, metadata creation, interoperability, and digital preservation. In their review of existing literature on digital libraries, Landis and Chandler (2007) note that much of it "focuses on technical issues and solutions, and is driven by a 'proof of concept' research approach" (p. 2). Few researcher have pursued more theoretical approaches to digitization practices. While a variety of researchers have considered other facets of digital library development, such as Dalbello's (2005) social constructionist study of digital library administrators and institutional practices, and Rosenbaum and Joung's (2004) use of the concept of sociotechnical interaction networks as a model for understanding digital library enrollment, the key site of digitization, *the lab*, still remains conceptually unexplored.

3   Research Questions

- How do digitizers use information sources when digitizing materials for preservation?
- How do digitizers enact embodied knowledge in their digitization practice?
- How do digitizers enact standards at the local level through their practice?

4   Method and Methodology

4.1   Research Design

This exploratory research focuses closely on three library staff members engaged in digitization activities. Participants selected had similar graduate-level educational backgrounds, and similar job responsibilities (project management and digitization of special collections). Through the process of observation and inter-viewing, data were collected on the participants' digitization tasks, their use and creation of information resources, and their manipulation of a multi-faceted environment of physical and digital tools.

4.2   Theoretical Framework

The theoretical approach for this project is based on practice theory, which takes the basic unit of analysis as "sets of actions that are based on the interconnectedness of their various nonreducible elements" (Veinot, 2007, p. 159). In addition, "practice theory contains a unique understanding of the body, which 'highlight[s] embodied capacities such as know-how, skills, tacit understanding, and dispositions' as the basis of activity (Schatzki, 2001)" (Veinot, 2007, p. 160). This focus on forms of embodied knowing helps practice theory to move information studies research away from cognitivist understandings of individual action to considering social actors as operating "within expectations or 'the accountability of a shared way of doing' (Corradi, et al., 2010, p. 277) set up in a practice" (Cox, 2012, p. 177).

Cox (2012) suggests practice theory has seen relatively limited use in information behavior research outside of work in knowledge management (p. 186), and Moring and Lloyd (2013) suggest that practice theory is "still emerging in the information studies field" (para. 1). Practice theory has seen ongoing use in information literacy research, particularly in the work of Annemaree Lloyd (2009, 2010, 2011, 2012) and others (e.g., Sundin and Francke, 2009; Tuominen, Savolainen and Talja, 2005). Practice theory is seen as useful in this research because it enables a conceptualization of information literacy "as something that develops in social contexts and is specific to a particular community" (Talja and Lloyd, 2010, xii). Practice theory has also seen fruitful application to studying the information practices of blue-collar workers (Veinot, 2007). Sundin and Francke (2009) define practice as the "various manifestations of repeated activities, including historical, social, cultural and material ones" (para. 6). The emphasis on the materiality of practice is significant for this research. Rather than focusing on processes of information seeking in the cognitivist framework, practice theory opens up the field of information-related phenomena to include the physical manipulation of tools and documents in virtual and physical spaces, as well as emphasizing the specificity of interfaces, textual genres, etc. In this project, practice theory allows us to study digitization as an em-bodied and *situated* activity (following Suchman, 1987), requiring improvisation within dynamic conditions and utilizing an array of tools and documentary techniques.

4.3   Data Collection

Participants were recruited with the help of the manager of the library digitization lab. Each participant had recently graduated with a Master of Library and Information Science degree from the same library science program, and had been working at the library as part-time employees for a year or less. Each participant was working independently on a different preservation project. Data were gathered from the participants via semi-structured interviews, running 25 minutes to an hour, using a set of nine interview questions, based on Veinot (2007). In addition, participants were observed for up to two hours, conducting

their everyday digitization tasks, and asked to describe each step of the workflow and give examples of problems that had developed in the past and what information sources they utilized to solve them.

4.4　Data Analysis Procedures

The audio from the interviews and observation sessions was recorded and transcribed by the researcher immediately after each session. Coding of emergent themes began in the transcription process. In addition, the audio, images and notes produced during the observation sessions were transcribed and integrated with the existing data, identifying participants' uses of tools and their use and production of documentary forms. Initial open coding was conducted during the transcription process to begin to immediately identify emergent themes. After transcription and initial coding, the data were further analyzed to elaborate major themes using constant comparative methods (Corbin and Strauss, 2008). Additional analysis focused on identifying the specific tasks, tools, problematic situations, and information sources used within each participant's workflow. These elements of the workflow were compared between participants to identify common and divergent themes across each type of digitization project.

5　Findings/Analysis

5.1　Institutional Context of Digitization Work

Participants in the digitization lab are responsible for handling a variety of tasks, including retrieving library materials, digitizing materials, transcoding between digital formats, conducting quality control on resulting digital copies, creating metadata, and ingesting files into the digital repository for access and preservation. The ongoing digitization work in the lab is typically structured and monitored by workflow software, but breakdowns in the system and the needs of specialized, ad hoc projects require staff members to improvise documentation techniques (e.g., forms, policy documents, spreadsheets, logbooks, technical specifications for equipment settings) for keeping track of their current position in the project and noting any modifications they may make to established procedures. These documentation techniques help participants manage their work, but they also act as a means of consciously transmitting their knowledge to future workers who may hold their position in the future.

5.2　RQ 1: How do digitizers use information sources when digitizing materials for preservation?

Digitizers use standardized imaging targets and default settings to implement standardized baselines for digital scanning, and they are instructed by senior administrators not to manipulate tonal or chromatic relationships once documents have entered the digital domain. However, a digitizer must rely on her own *aesthetic judgment* (discussed below) for making decisions about the quality of digital output. Participants overcome problematic situations by searching for illustrative models of how others in the wider library community have overcome similar obstacles, following the online guidelines of such organizations as the Online Computer Library Center for metadata, as well as looking back to earlier records in the library's catalog as guides.

5.2.1　Aesthetic judgment

Participants displayed a form of embodied knowing that I have tentatively labeled *aesthetic judgment*, in which individuals integrate educated perceptual abilities and situated knowledge to come to a decision about the acceptable visual quality of digitized copies. The use of targets and color swatches takes away some of the need for aesthetic judgment, replacing it with objectively established measures, but other attributes such as depth of field, angle of lighting, etc. lack established standards or guidelines, so aesthetic judgment is required. For instance, Participant B explained the limitations of this process of making aesthetic judgments: "Sometimes this is as good it gets. I hate to say it. Some of the blurring on the items seeks to assist with that three dimensional feeling."

One unexpected finding was the creative use of color in various places in the workflow, which further suggests that participants are engaging in a practice that has a significant, and unexplored, aesthetic dimension. For instance, two of the participants use improvised systems of color coding as an important strategy for managing their file systems.

5.3   RQ2: How do digitizers utilize embodied knowledge in the practice of digitization?

Participants commonly employed embodied knowledge for a variety of tasks, including assessing quality output, determining when files were done being processed, setting up items for digitization, and creating metadata records. These embodied practices draw on visual, aural and tactile ways of knowing to implement tacit knowledge gained through training and experimentation with equipment. For instance, in digitizing three dimensional artifacts, participant B employs methods of weighing and measuring with instruments and her hands to gather information in order to fill metadata fields in the digitized item's record, as well as to assess how difficult it will be to position a particular object for digitization.

5.3.1   Improvised self-documentation

An emergent theme I have termed *improvised self-documentation*, in which participants create their own forms and documents to record the improvised actions and decisions that they enact in their practice, produces important information sources utilized by participants and others in the lab. New knowledge about innovative digitization processes is preserved through self-documentation, helping digitizers to manage personal workflow organization, transmit new knowledge to other digitizers, and to communicate progress to supervisors. In addition, the materials produced through self-documentation, such as application profiles, user guides, and logs of digitized materials appear to serve as *de facto* local standards, working to ensure continuity in practice over time.

5.4   RQ3: How are standards enacted in preservation practice?

Participants self-consciously monitor their personal actions, and the actions of others working with them (i.e. workstudy students) for correspondence to accepted local standards. Participants are recent graduates from the same Master in Library and Information Science program, suggesting that their common self-conscious adherence to library norms may be tied to as shared educational background.

Once activities become "rote", internalized, and fully embodied, referring to standards and information sources typically becomes less frequent, until a problematic situation develops. For simple text documents, standards are strict, and all settings have been stored as defaults or presets in the hardware and software of the system. For other materials, such as three-dimensional artifacts, videos, audio recordings, or maps, standards for digitization and metadata creation are seeing ongoing development. The participants indicated a tension between the standards that they have agreed to follow and what they sometimes would like to do to improve the quality of the digital output. For instance, Participant B admits that "sometimes I ever so slightly fudge it in Photoshop, but we're really not supposed to do that. We are not supposed to beautify these [digitized objects]. We [only] remove distracting background elements." Beyond effacing marks of the context of image scanning (removing the background elements from the final image), after digitization, modification to the image is avoided. At other times, standards are followed without question, or are directly embedded within the technology in the form of defaults or presets that cannot be circumvented without technical intervention from administrators.

6   Summary

These findings suggest that digitization is an information practice that operates under varying conditions of standardization, improvisation, self-documentation and embodied knowledge. The lab's workflow software precisely defines and tracks each step in the process of scanning text-based documents, but other formats require varying degrees of deviation from the rule-based workflow model. Yet, even in the highly

rigid structure of the standards for digitizing text, participants still rely on embodied knowledge and tacit understanding to make judgments about whether items were scanned properly or not, often times improvising creative uses of color-coding in the file system to facilitate quick access to current files.

6.1   Limitations

The major limitation of this project is the small sample size of three participants. Additional participants and institutional sites will certainly help to enhance the validity and scope of these findings.

6.2   Future Research

This work is exploratory, and it is hoped that the patterns in digitization practice that emerged from the data can be further explored in a broader study. This research opens up the following avenues for future research:

- Exploring aesthetic judgment and the role of color in digitization practice and in human information behavior, more generally speaking.
- Further examining the role of texts produced through self-documentation in the social construction of digitization practices.
- Explore how the interaction between standards and self-documentation interact to *socially legitimize* (Cf. Byström & Lloyd, 2012) digitization practice.

7   Conclusion

While the use of practice theory in information research is "still emerging" (Moring and Lloyd, 2013), this research indicates the applicability of a practice theory approach to conducting research on information-related activities in the digitization labs of academic libraries. This research offers future directions for research into expanding theoretical understanding of digitization as a complex information practice that draws on multiple modalities of knowing, including the emergent theme of *aesthetic judgment*, which has not seen exploration in the field of Library and Information Science. This research also contributes to the conceptualization of the role of standards in local practice.

In addition, better understanding of digitization practices in academic libraries also points to practical implications, such as providing insight into how systems developers might create workflow software that can better assist staff members' changing documentation requirements when digitizing different media formats.

In terms of enhancing theoretical understanding of digitization, this exploratory research suggests that practice theory offers a rich qualitative lens through which to conceptualize digitization as a dynamic and embodied practice, instead of as a simple, specifiable task.

8   References

Berger, M. (1999). Digitization for preservation and access: A case study. *Library Hi Tech,17*(2), 146-151.

Byström, K., & Lloyd, A. (2012). Practice theory and work task performance: How are they related and how can they contribute to a study of information practices. *Proceedings of the American Society for Information Science and Technology, 49*(1), 1-5.

Capell, L. (2010). Digitization as a preservation method for damaged acetate negatives: A case study. *American Archivist, 73*, 235-249.

Corbin, J. & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory, 3rd edition.* Los Angeles: Sage Publications.

Corradi, G., Gherardi, S., & Verzelloni, L. (2010). Through the practice lens: Where is the band wagon of practice-based studies heading? *Management Learning, 41*(3), 265-283.

Cox, A. M. (2012). An exploration of the practice approach and its place in information science. *Journal of Information Science, 38*(2), 176-188.

Dalbello, M. (2005). A phenomenological study of an emergent national digital library, part I: Theory and methodological framework. *The Library Quarterly, 75*(4), 391-420.

Evens, T., & Hauttekeete, L. (2011). Challenges of digital preservation for cultural heritage institutions. *Journal of Librarianship and Information Science, 43*(3), 157-165.

Fleischhauer, C. (2010). Format considerations in audio-visual preservation reformatting: Snapshots from the Federal Agencies Digitization Guidelines Initiative. *Information Standards Quarterly, 22*(2), 34-43.

Johannisson, J., & Sundin, O. (2007). Putting discourse to work: Information practices and the professional project of nurses. *Library Quarterly, 77*(2), 199-218.

Landis, W. E., & Chandler, R. L. (2007). Archives and archivists: Key collaborators in the digital library. *Journal of Archival Organization, 4*(1-2), 1-9.

Lloyd, A. (2009). Informing practice: Information experiences of ambulance officers in training and on-road practice. *Journal of Documentation, 65*(3), 396-419.

Lloyd, A. (2010). Framing information literacy as information practice: Site ontology and practice theory. *Journal of Documentation, 66*(2), 245-258.

Lloyd, A. (2011). Between a rock and a hard place: What counts as information literacy in the workplace and how is it conceptualized? *Library Trends, 60*(2), 277-296.

Lloyd, A. (2012). Information literacy as a socially enacted practice. *Journal of Documentation, 68*(6), 772-783.

Lopatin, L. (2006). Library digitization projects, issues and guidelines: A survey of the literature. *Library Hi Tech, 24*(2), 273-289.

Moring, C. & Lloyd, A. (2013). Analytical implications of using practice theory in workplace information literacy research. *Information Research, 18*(3). Available at http://InformationR.net/ir/18-3/colis/paperC35.html

Plichta, B., & Kornbluh, M. (2002). Digitizing speech recordings for archival purposes. Retrieved 3-2-13, http://www.historicalvoices.org/papers/audio_digitization.pdf

Rosenbaum, H. and Joung, K. (2004), Socio-technical interaction networks as a tool for under standing digital libraries. *Proceedings of the American Society for Information Science and Technology, 41*, 206–212. doi: 10.1002/meet.1450410124

Schatzki, T. (2000). Introduction: Practice theory. In Schatzki, T., Knorr-Cetina, K., and von Savigny, E. (Eds.), *The practice turn in contemporary theory* (pp. 10-23). London: Routledge.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication.* New York: Cambridge University Press.

Sundin, O., & Francke, H. (2009). In search of credibility: Pupils' information practices in learning environments. *Information Research, 14*(4). Retrieved from www.informationr.net/ir/144/paper418.html

Talja, S. & Lloyd, A. (2010). Integrating theories of learning, literacies and information practice. In Lloyd, A. and Talja. S. (Eds.), *Practising information literacy. Bringing theories of learning, practice and literacy together* (viiii-xx). Wagga Wagga: Centre for Information Studies, Charles Sturt University.

Teper, J. & Shaw, E. (2011). Planning for preservation during mass digitization projects. *Libraries and the Academy, 11*(2), 717-739.

Tuominen, K., Savolainen, R., & Talja, S. (2005). Information literacy as sociotechnical practice. *Library Quarterly, 75*(3), 329-345.

Veinot, T. C. (2007). "The eyes of the power company": Workplace information practices of a vault in-
spector. *The Library Quarterly, 77*(2), 157-179.