

Digital Libraries & Desktops



How Users View & Use Digital Libraries in Relation to their Own Private Collections

Christina M. Finneran
Syracuse University
School of Information Studies

Outline of Talk



- Introduction to the Problem
- Background
- Research Questions
- Theoretical Framework
- Method
- Study Limitations
- Expected Contribution

The Problem



- People are exposed to a plethora of information from an array of sources
 - A particular digital library is just one source
- The workplace is more amorphous and dynamic
 - Judging (future) usefulness of information is difficult
- Digital libraries' aim for content to be accessible
 - What about re-finding content?
 - Providing portals, personal DL for users manage materials
- How can digital libraries enable re-access of material found within them?

Background



- Personal Information Management
 - Location-based factors
 - Situational factors
 - Temporal factors
 - Unprocessed information
- *Keeping Found Things Found* project
- Personal Digital Libraries

Location-based Factors



■ Malone, 1983

- People organized their paper documents in “piles” and “files”
 - Piles: unordered, unnamed documents
 - Files: named and ordered documents
- People used location of documents as **reminders** in their work
 - Intentions are important when considering users

Situational Factors (Contextual Metadata)



- Metadata can be personal, transient, and situation-specific (Marshall, 1998; Bishop, 1999)
- Ducheneaut & Bellotti (2001) - embedded-ness of transmitting documents w/in email
- Kwasnik (1991)
 - situation attributes accounted for 33%, while the document attributes accounted for 29%
 - Use/Purpose (13%)
 - Time (9%)
 - Location (8%)

Temporal Factors (Document Types)



- Cole, 1982; Barreau & Nardi, 1995; Sellen & Harper, 2002, etc.
 - Ephemeral (“Hot”)
 - | Documents that are needed for short periods of time and quickly thereafter lose their value
 - Ex: Scraps of paper
 - Not given enough attention in tool development (Barreau & Nardi, 1995)
 - Working (“Warm”)
 - | Documents that are related to active projects and have ongoing value during the project
 - Ex: Project plan, Proposed budget
 - Archival (“Cold”)
 - | Documents that represent a completed project and have unknown future value

Unprocessed Information?



- Whittaker & Hirschberg (2001) study of paper files when office was moving
 - | 15% of paper archives were unread
 - | Because of time constraints, users were not able to read and evaluate the usefulness of documents
- With electronic documents that can be transmitted and stored so easily, likely to be a larger percentage of documents in this category?
 - | Email: messages of “undetermined status” sits in inbox (Whittaker & Sidner)

Keeping Found Things Found



- Jones, Bruce, & Dumais project developed list of features considered when deciding how to keep web site information for future access
 - | Portability
 - | Number of access points
 - | Persistence of information
 - | Preservation in current state
 - | Currency of information
 - | Context
 - | Reminding
 - | Ease of Integration
 - | Communication and information sharing
 - | Ease of maintenance

To Keep or Not to Keep?



- *Keeping Found Things Found* project
 - Sometimes users do not “keep” the web site because they would not have trouble finding it again
- Given search capability of DL, if the user knew to look in the DL, they should find it again
 - Why spend their time organizing the content?

Personal Digital Libraries



- PDL: users can store their own content and make it available to others if they wish
- 2003 NSF Workshop on DL stated that personal digital libraries were important for research
- *Managing the Digital University Desktop* (Tibbo)
- ADEPT added PDL for instructors
 - “architecture will allow collections and services to be tailored to individual practices without compromising the organization of the larger database (Borgman et al, 2005)

Research Questions




1. How do users manage the material that they find useful from digital libraries?
2. How do users conceptualize the distinctions between, and the inter-relationships of, digital libraries and their private collections?
3. How does the way users conceptualize a digital library impact the management of items from a digital library?

RQ1: How do users manage digital library items?



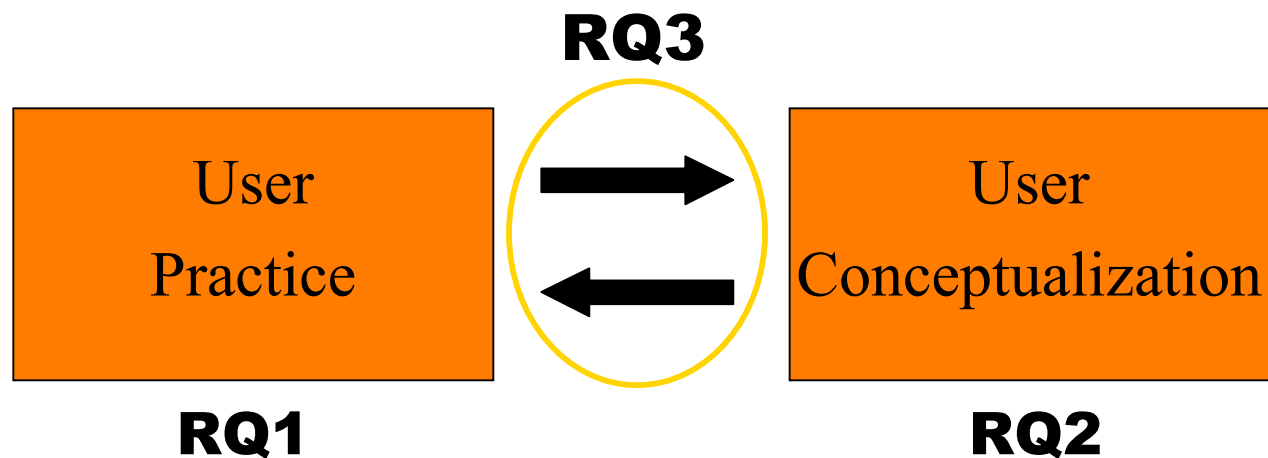
- In what situations do users decide to (or not to) keep information?
 - Vary by ephemeral, working, archival?
 - Is there a "To Be Processed" category?
- How do users keep items?
 - In own desktop, DL portal, PDL, bookmarks, etc?
 - How does this compare to keeping web sites (*KFTF*)?
- In what contexts, and why, do users store items within their private collection? DL portal/PDL?
- What metadata are users trying to retain through their management techniques?

RQ2: How do users conceptualize DLs and their private collections?



- A. What metaphors do users use when talking about the digital libraries that they regularly visit in relation to their own private collection?
- B. To what extent, and in what manner, is the users' view of the digital library impacted by their utilization of the portal/personal digital library tools available within the DL?

RQ3: How does user conceptualization impact practice?



Theoretical Framework

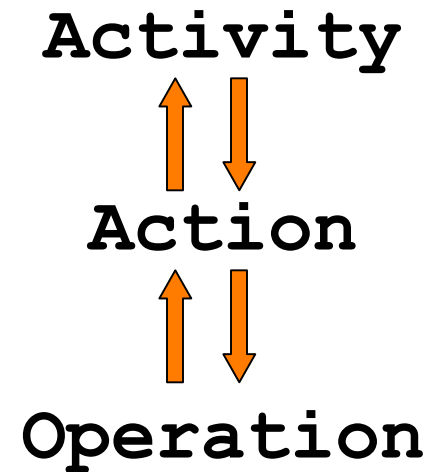


- Why Activity Theory?
 - Merges thought and action, considers practice
 - Enables understanding tool within activity
 - Handles dynamic levels of users "task"

Activity Theory



- 3-level model of activities (Leont'ev)
 - Hierarchical
 - Dynamic

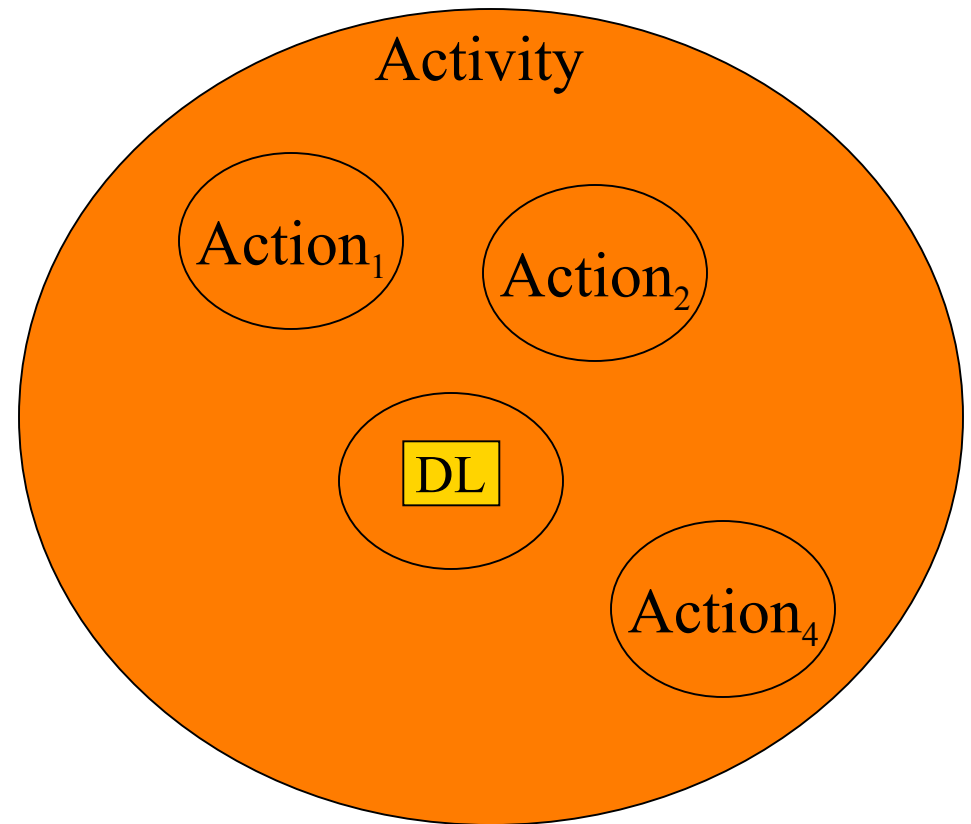
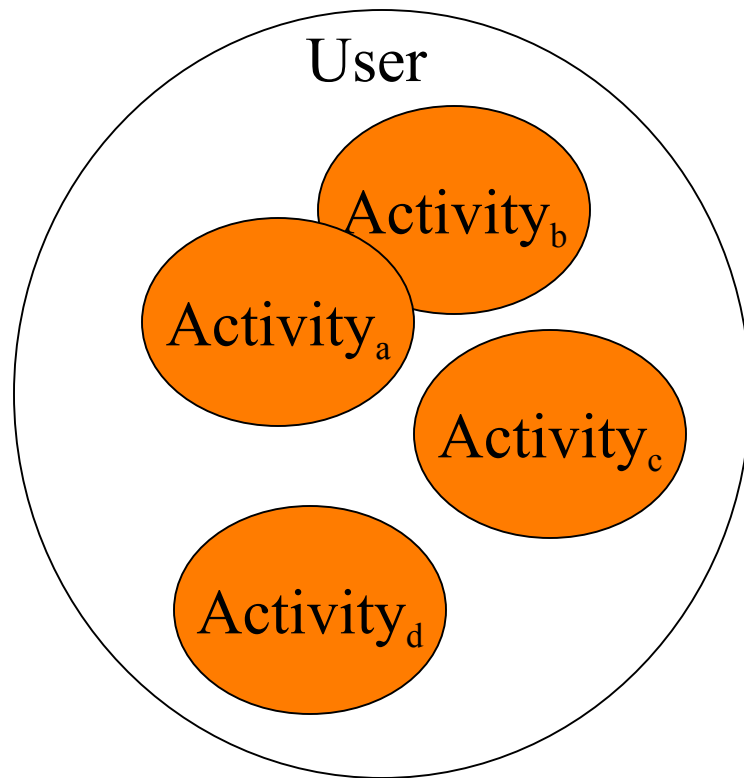


Method



- Multiple embedded case study
- Each user is a case
 - Each activity of the user is an embedded unit
- Recruit users who regularly visit an established DL that has a PDL/portal

Study Perspective



Data Collection



- In-depth interviews about the activities for which they use that DL content
 - What activities are they working on?
 - How do they generally handle info that is relevant?
 - Think of an example when you decided not to keep something useful?
 - Why, and how, they keep items in their private collection?
- Ask about other DL they use for those activities
 - Repeat questions about each DL

Data Analysis



- Analyze data by activity
 - How are useful items kept?
 - In what contexts, and why, do users store items within their private collection (desktop)? PDL or DL portal?
 - What metadata are users trying to retain through their management techniques?
 - What do users decide to (or not to) keep/store?
 - Vary by ephemeral, working, archival?
 - Is there a "To Be Processed" category?
- Look for patterns within activities, users, and across users using analytical induction (Miles & Huberman)

Conceptualization of DL



- Analyze transcripts of interviews looking for metaphors depicting the relationship between the digital library and user's own collection
 - At individual user-DL relationship
 - At user level
 - Across users

Study Limitations



- Only looks at current DL users
- Primarily focus on individual collections, though discuss collaborative

Expected Contribution



- Refine new approach studying DL users
 - Considers library within the life of the user
- Understand how, and why, users chose to keep and manage information from DL
 - Help DL designers with personalization and customization features

Thank You



Questions?

References



- Barreau, D. & Nardi, B.A. (1996). Finding and Reminding: file Organization from the Desktop. *SIGCHI Bulletin*, 27(3), 39-43.
- Bishop, A.P. (1999). Document structure and digital libraries: How researchers mobilize information in journal articles. *Information Processing and Management*, 35, 255-279.
- Borgman, C.L. et al. (2005). Comparing faculty information seeking in teaching and research: Implications for the design of digital libraries. *Journal of the American Society for Information Science and Technology*, 56(6), 636-657.
- Jones, W., Dumais, S., & Bruce, H. (2002). *Once found, what then? A study of "keeping" behaviors in personal use of web information*. Paper presented at the American Society for Information Science & Technology, Philadelphia, PA.
- Ducheneaut, N. and Bellotti, V. (2002). E-mail as Habitat. *Interactions*, 30-38.
- Kwasnik, B. H. (1991). The importance of factors that are not document attributes in the organization of personal documents. *Journal of Documentation*, 47(4), 389-398.

References (cont.)



- Leont'ev, A.N. (1978). *Activity, consciousness, and personality*. Englewood Cliffs, NJ: Prentice-Hall.
- Malone, T. W. (1983). How do people organize their desks? Implications for the design of office information systems. *ACM Transactions on Office Information Systems*, 1(1), 99-112.
- Marshall, C.C. (1998). *Making metadata: A study of metadata creation and use in a mixed physical-digital collection*. Paper presented at the 3rd ACM International Conference on Digital Libraries
- Nardi, B.A. (1996). *Context and consciousness: Activity theory and human-computer interaction*. Cambridge, MA: MIT Press.
- Sellen, A.J. & Harper, R.H.R. (2002). *The myth of the paperless office*. Cambridge, MA: MIT Press.
- Whittaker, S., & Hirschberg, J. (2001). The character, value, and management of personal paper archives. *ACM Transactions on Computer-Human Interaction*, 8(2), 150-170.