Yours, Mine and (Not) Ours: Social Influences on Group Information Repositories

Emilee Rader
School of Information
University of Michigan
ejrader@umich.edu

ABSTRACT

Group information repositories are systems for storing and organizing files in a central location all group members can access. The functionality and capabilities of these systems are essentially the same as the desktop metaphor of personal information management (PIM) systems. Using a case study of a group information repository, I argue that social factors affect the information structure of the repository, and how it grows and evolves over time. Users restrict their activities to files they "own," are reluctant to delete files that might be useful to others, dislike the clutter that results, and can become demotivated if no one views files they uploaded.

ACM Classification Keywords

H.3.2 Information Storage and Retrieval: Information Storage – File organization; H.5.3 Information Interfaces and Presentation: Group and Organization Interfaces – collaborative computing

Author Keywords

information management, group memory, information producers and consumers, desktop metaphor

INTRODUCTION

Berlin et. al [1] introduced the idea of "group memory" systems in 1993, defining the term as "a common repository of on-line, minimally structured information of persistent value to a group". Fifteen years later *group information repository* systems are commonplace; many workgroups use them for storing and organizing files in a central location that all members can access. Examples include multiple-author blogs, wikis, content management systems, and version control systems for software code (e.g.: Drupal, Subversion, Sakai, OpenText Livelink, shared network folders, etc.). Despite the widespread adoption of these systems, they are far from perfect. Document management practices incur costs such as "ineffective access to information, duplication of effort, failure to share information, and information overload" [2]. A group information repository is more complex than just an "aggregate of every individual's contribution" [3]; both users' choices about what to contribute to a repository, and how the information within it is organized, are influenced by social practices. The information structure of the repository emerges through individual, idiosyncratic file naming and organizing choices and provides the means by which documents are located and accessed.

Repositories are collaborative systems, situated in a social context; however, group information management behaviors and systems are often studied in terms of an individual's efforts toward organizing and using information, rather than focusing on the aggregate or group level effects [4]. For example, Voida et al.[7] described a system for granting other individuals access to one's personal files that makes information about what was shared with whom more visually explicit. And, Tang et al.[6] identified similarities and differences among individual enterprise knowledge workers' files; they suggested this kind information could be used to help other individuals find information they need. Both examples focus on individual-level outcomes, rather than group level effects like how perceptions of ownership might exert a subtle yet long-term influence on how information in a group repository is structured and organized.

In this paper, I present findings from a case study of a group information repository system. I argue that the functionality and capabilities of the system are essentially the same as the file-and-folder hierarchy desktop metaphor for personal information management (PIM), but social factors affect the information structure and how the repository grows and evolves over time.

METHOD AND PARTICIPANTS

CTools is a group information repository and course management system used at the University of Michigan, and is an instance of the Sakai open-source application platform in use at over 150 institutions worlwide. Users can create *project sites* which support storing and sharing files online. CTools project sites are increasingly popular: the number of active project sites grew 213% from December 2005 (3170 sites) to December 2007 (9932 sites)[1]. The interface for storing and sharing files is similar to that of many other web-based systems for group file storage: it follows the file-and-folder desktop metaphor, meaning that files and folders can exist in one and only one location in the hierarchy. There

---

[1]Based on an analysis of system log data not included here.

Table 1. Descriptives about the CTools project sites used in this study. Data reflect site characteristics as of the date interview sessions began.

	type	users	site age	# files	accesses/ week
Site 1	extracurricular	26	2.5 yrs.	115	20.40
Site 2	staff/admin	89	1 yr.	452	76.08
Site 3	research group	11	4 yrs.	580	26.96
Site 4	research group	11	1 yr.	120	24.88
Site 5	research group	13	2.5 yrs.	894	22.73
Site 6	extracurricular	18	1.5 yrs.	407	19.10

is no version control, access control exists at the site level, and there is no search functionality. Users can specify that a notification be sent when adding a file, and set preferences for receiving notifications when another user adds a file.

Sixteen users of six different CTools project sites (4 men, 12 women) participated in semi-structured interviews[2]. The sites were selected for variety in the type of work each group conducted, and for activity level: each site had at least three members active on an approximately weekly basis. Respondents ranged from undergraduates to graduate students in several departments across campus to University staff members from different units who had been part of the organization for decades; some used CTools several times a week, while others did so rarely. At least two members of each site were interviewed, in their normal work environments wherever possible and in front of a computer so they could access the CTools site of their group as needed. Respondents answered questions about the work being conducted by the group, about their use of the CTools site, and about interactions with other group members. They also walked through the CTools site with the interviewer, describing the information available on the site and how it was organized.

The interviews were recorded, transcribed and coded using using an inductive approach. Open, iterative coding took place in parallel with data collection, starting from an initial conceptual framework developed during a review of collaboration and information management literature. Early coding stages consisted of labeling and categorizing participants' reported attitudes and interactions with respect to other group members and the information repository, and observations about the work practices and environments unique to each group. In later stages of the analysis, high-level themes emerged as connections were drawn between codes, participants, and sites. The findings below report themes observed across all sites, despite differences in type of group, activity level, age and size (see Table 1).

FINDINGS
MY Stuff vs. YOUR Stuff

Group information repository users are both information *producers*—users who contribute files to the repository—and information *consumers*, or re-users of documents in the repository [5]. In a personal repository the producer and consumer are the same person. However, in a group information repository the producer and consumer roles can be filled by the same individual, or any combination of users.

---

[2]All names are pseudonyms. Some names of files and other details have been omitted to protect respondents' identities.

Complications arise when multiple information producers add files to a repository, because it is very difficult to both initially agree upon and then subsequently adhere to conventions for how files should be labeled and organized [1]. Nancy (Site 3) said,

> Probably the biggest problem we have with CTools is that people tend to organize information different ways, you know like you have a picture in your mind of how you think it should be organized, and that's not exactly how someone else's brain works to organize things.

Steve (Site 5) talked about trying to organize the site so that it would make sense to other people. He felt like he could only speculate on where other group members wanted things to be stored:

> I set up the directory structure and that kind of stuff. I'm always tweaking it. And I'm always looking for input, but I never get any input on stuff like, 'how would you like the stuff to be arranged'.

This tension leads to conflicts between "my stuff" and "your stuff" in a repository. Ideally, the contents of the repository are "our stuff", but in practice perceptions of ownership have very real consequences for repository structure.

Nancy and Zoe (Site 3) separately described the same incident that highlights consequences of this tension. Nancy was unhappy that members of her group all had different ideas about where things should be stored, but did not feel right about taking a hard-line approach to fixing the problem: "Unfortunately, I think it's important that people do things the same way but other people don't agree with me and I'm not an enforcer." However, when Zoe added a file to the site that Nancy felt should live in a different place, Nancy took action (described by Zoe): "So actually, I got an e-mail from Nancy telling me to move it to [a different folder]." When Zoe didn't move it, Nancy created a copy and put it in the new location herself without telling Zoe she had done so. During the interview, Zoe discovered her file in a different place than she had originally put it and said,

> It looks like she moved it to another folder but it's still here too... I'm the original one who posted it though, and it got moved around. But there's only two of us right now that are using it. And she has it in the place she wants it. And I have it in the place I put it.

Nancy took action to satisfy her own needs while preserving Zoe's original location and ownership of the file; however, this action also increased the clutter on the site. In another instance, Nancy described the history behind how two nearly identical folders on her group's site came to exist:

> And then [the folder called] 'Slides', this is where [the PI] sticks a bunch of powerpoint for talks that he's given on the whole concept of our lab... the 'Conferences' folder contains. Well. Yeah. Actually in a practical sense they should probably be the same. But the Conferences folder, contains stuff that's usually done by students, and the other Slides folder contains stuff that's done by [the PI].

Combining these folders might seem an obvious next step, and one that could even take place during the interview, but Nancy did not do it. The 'Slides' folder had been created by the PI, not by Nancy, and she was reluctant to make any changes to it. As a result, anyone needing presentations made by members of the group would have to look for the information in two places with different names.

First Do No Harm

Information producers typically do not like to undertake the task of pruning the contents of a group information repository [5]. Users feel they should hang onto information they are not sure they need, just in case the need might arise later. Pruning decisions are especially difficult when repository contents are perceived as either yours or mine, but not "ours"; if users are bad at estimating their own future use, they are especially uncomfortable making assumptions about others' future information needs. The consequence of this uncertainty for respondents was they never deleted any files—even ones they had created! Josh (Site 6) said the only files he might feel comfortable deleting from his group's CTools site were outdated files he himself had created:

> I try not to [delete files]... so if another person goes looking for it, they can find it. The only times I delete, is when I post something more updated. And I'm sure that no one else will need the old version. Interviewer: How do you become sure that no one wants the old version? Josh: I only delete things I've posted.

However, he could not recollect a single specific instance when he had deleted a file. The only time Susan (Site 4) remembered deleting anything was, "...when we created the snack list, for last semester, and so I just created a new one for this semester and deleted the old one." When asked about what kinds of files she might delete, Linda (Site 2) said:

> Well, I can't think of anything. The only thing I can think of is if it's something really offensive, or doesn't pertain to [the project]. Then I would probably delete it. But I can't imagine anybody doing that. It hasn't really crossed my mind [to delete something].

Nancy (Site 3) recognized the value in periodically pruning the CTools site: "It seems like once a year somebody should sit down, like me or somebody else in the lab and kindof restructure all the new folders that pop up in here." Even so, when asked if such a cleanup had ever happened, Nancy's response was "No [giggle]." While exploring the site during the interview, she came across at least three different folders that had been created years ago by members who left the group. The folders were subsequently abandoned, new group members had no way of knowing why they existed, and yet they remained on the site. Despite the potential for positive outcomes from pruning, respondents kept outdated files and folders around just in case they or someone else might need them. They were unwilling to make a decision that might directly prevent another member from accessing the information, especially if they were not the original creator of the file. Even these decisions NOT to remove something have consequences for the structure and evolution of the site.

Consequences of Clutter

While there is no direct monetary cost or quota for sites that grow continuously and never shrink, users of cluttered and poorly organized repositories incur hidden costs in terms of lost time and effort. They may even be unaware of potentially important information: "It's not really people not knowing where to look, or not being able to find it once they're on there, but just not even knowing it's on there" (Zoe, Site 3).

David (Site 6) talked about how he was only familiar with the part of the CTools site where he kept files related to his part of the project: "So most of what I need is in the operations folder. To be honest I don't think I've ever really even looked in the marketing folder, because it's just not stuff that I need." Jennifer (Site 5) described the same phenomenon:

> See, you keep going and going and going and it's more folders and some more folders and some more folders, so if you don't really know, you're not really familiar with... let's see. 1,2,3,4,5 [levels]... yeah. So you really have to know your way around here or it can be quite intimidating. But like I know that my stuff is in here... wait a minute. Where is my stuff? Here it is! Like, I've been working on factor analysis, these are all my files. And I know that my reliability stuff is in here. So, I just know my spot, basically.

Jennifer did not explore other parts of the repository; she kept all the "stuff" she thought she needed in one place. However, because she did not visit other parts of the site regularly, she did not know what other information was available. During her interview, she came across a file she had not known was there: "And this one here... I totally need to read this, actually. This is really important, I should read this one. Because I'm doing work on that right now. It's my thing." Zoe (Site 3) had a similar experience during the interview:

> I don't really know what this is. It's probably just reference databases Nancy had on her computer and uploaded. I mean it would be beneficial for somebody to use... it looks like it hasn't been modified for a couple years. But some of it is interesting, and probably would have been useful for me when I was doing my quals.

As Jennifer and Zoe discovered, multiple people using a group information repository do not always know what other members might be adding, since they tend to restrict themselves voluntarily to the areas they are familiar with. In David's case, he voluntarily self-restricts because he is satisficing, only interacting with the parts of the site for which he has immediate need. The system provides access control only at the level of entire sites; one cannot programmatically hide individual files or sections of the hierarchy from specific members. And yet the self-restriction essentially creates the same kind of consequences.

Unmet Social Expectations

Respondents who added files to the group's CTools site expected that other group members would be aware of their contribution, and that some would make use of it. For example, Nancy (Site 3) talked about putting some items on-

line that were directly related to something she was working on, that she thought others would want to see: "I was looking at reference papers of a co-PI who works with our lab and I posted them to our site because I thought they were papers that people should probably read and be aware of." However, the only way for her to tell whether anyone had accessed them was to ask directly, in person: "At our lab meeting I mentioned that I had posted these as well. But nobody said they had looked at them." Susan (Site 4) had a similar experience, posting manuals and research papers to the site. And Zoe (Site 3) posted her procedures for operating research equipment so that others could take advantage of them. Jennifer (Site 5) explained her thought process regarding whether to add a particular document she and another group member had been working on together:

> Doesn't it sound like it should be here? But the problem is that there are only like a couple of us right now that are really working on it, that really have our hands on it, and so we have our own versions. But yeah, this is important for... people should have access to that.

However, on the flip side, Frances (Site 2) spoke negatively of information posted by another group member: "[information] that [the group member] found, that she thought would be useful [for the group]". The implication of Frances' comment was that only the original poster, whom she called a "walking brainstorm", would find the information to be useful. Not knowing whether anyone else ever looks at the information one has added leads to the assumption that nobody is interested, and that can be demotivating. Linda (Site 2) experienced this, and talked about it in the interview:

> I did stop doing it [posting meeting minutes to CTools] for like a month, in January. The January meeting minutes are missing from CTools. I didn't do it and nobody said anything. And I'm like, why do I keep doing this?

The act of sharing a document via the repository carries with it an expectation that the audience is out there, despite the lack of a targeted recipient and the asynchronous nature of the potential exchange. When this expectation is violated in a group information repository it can reinforce the tendency to simply stick to one's own self-designated area of the site.

DISCUSSION AND IMPLICATIONS

Each finding above points to a way system designers can shift focus from the individual and improve information flow between users, to better support the social aspects of the group information repository system:

1. **MY Stuff vs. YOUR Stuff**: Users' knowledge of repository contents is necessarily biased toward what they've added themselves. Feedback supporting awareness of others' use of the system [7] is essential for instilling a greater feeling that the repository contains "our" information.

2. **First Do No Harm**: It can be difficult for individual users to differentiate useful, relevant files from older, outdated ones. The interface might provide information about which files are frequently accessed and which are not, serving as a basis for decisions about what to keep and what to delete. Further research is necessary to determine whether this information can empower users to overcome their reluctance to delete files "belonging" to someone else.

3. **Consequences of Clutter**: The system should provide support for identifying cross-folder commonalities and linkages when information producers make decisions about where in the hierarchy to store new files. This information could help users recognize connections between "my" stuff and "your" stuff, much the way that the *Frequently Bought Together* and *Customers Who Bought This Item Also Bought* fields on Amazon.com bring items to the attention of online shoppers.

4. **Unmet Social Expectations**: On the information consumer end, the system should provide easy access to a summary of the information available, and recent changes to the site. This approach is used successfully on Facebook to keep Friends in the loop on each other's status and activities. Without a summary, the only way to become familiar with files added by other group members is brute-force browsing of the entire site.

REFERENCES
1. L. M. Berlin, R. Jeffries, V. L. O'Day, A. Paepcke, and C. Wharton. Where did you put it? Issues in the design and use of a group memory. In *CHI '93*, 23–30, Amsterdam, The Netherlands, 1993. ACM Press.

2. M. D. Gordon. It's 10 a.m. Do you know where your documents are? The nature and scope of information retrieval problems in business. *Information Processing & Management*, 33(1):107–122, 1997.

3. G. Jian and L. Jeffres. Understanding employees' willingness to contribute to shared electronic databases: A three dimensional framework. *Communication Research*, 33(4):242–261, 2006.

4. W. G. Lutters, M. S. Ackerman, and X. Zhou. Group Information Management. In W. Jones and J. Teevan, editors, *Personal Information Management*. Univ. of Washington Press, 2007.

5. L. M. Markus. Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success. *Journal of MIS*, 18(1):57 – 93, 2001.

6. J. C. Tang, C. Drews, M. Smith, F. Wu, A. Sue, and T. Lau. Exploring patterns of social commonality among file directories at work. In *CHI '07*, 951–960, New York, NY, USA, 2007. ACM.

7. S. Voida, W. K. Edwards, M. W. Newman, R. E. Grinter, and N. Ducheneaut. Share and share alike: exploring the user interface affordances of file sharing. In *CHI '06*, Montreal, Quebec, Canada, 2006.