

# Group Information Repositories as Social Systems

**Emilee Rader**

School of Information  
University of Michigan  
ejrader@umich.edu

## 1 Introduction

Many workgroups use information systems for storing and organizing information online, in a central location that all members can access. Examples of such *group information repositories* 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.) Group information repositories are essential for document sharing, and can be greatly beneficial for organizational efficiency, communicating organizational goals, and also for learning and innovation. They can contain “mission critical information” such that if it were lost there would be serious consequences ([Blair, 2002](#)). Despite the growing importance of these systems, ineffective document management incurs costs such as “lost work time, ineffective access to information, duplication of effort, failure to share information, and information overload” ([Gordon, 1997](#)).

Information management behaviors and systems are often studied in terms of an individual’s efforts toward organizing and using their personal information (e.g., [Jones & Teevan, 2007](#)). However, group information repositories are collaborative systems, situated in a social context. A repository is more complex than just an “aggregate of every individual’s contribution” ([Jian & Jeffres, 2006](#)); both users’ choices about what to contribute to a repository, and how the information within it is organized, are influenced by social practices. In this paper, I argue that analyzing group information repositories from the perspective that they are social systems highlights ways they are different from other information management and document sharing domains.

## 2 CTools Case Study

### 2.1 Background

[CTools](#) is a group information repository and course-management system in use at the University of Michigan. Users can create *project sites*, with which group members can “make announcements

and share resources, such as documents or links to other resources on the web”<sup>1</sup>.

Eighteen users of six different CTools project sites participated in semi-structured interviews. 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. The interviews were recorded, transcribed and coded using qualitative data analysis (QDA) software, in a similar fashion to Miles & Huberman, 1994.

## 2.2 Sharing and Awareness

A repository user is generally familiar with other group members through interactions that take place face-to-face or via some communications medium, and with projects and joint work activities they are engaged in together. However, he can expect to be familiar with only some of the documents stored in a the repository, and he may or may not have been involved with contributing and organizing documents. Documents in a repository are *shared* in the sense that they can be accessed by anyone with permission to use the repository; however, the action of adding a document to a repository is more like making the document available to the users of the repository, than *sharing* it directly with any particular person.

Respondents who added files to the group’s CTools site sometimes had an expectation that if they put the information online, others would eventually see it. For example, Nancy (Site 3)<sup>2</sup> talked about putting some items online that were directly related to something she was working on, that she thought others should see: “I was looking at reference papers of a co-PI<sup>3</sup> 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.” Jennifer (Site 5) talked about 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 [the document] should be here? But the problem is that there are only like a couple of us right now that are really working 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.”

Not knowing whether anyone else ever looks at the information 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 document will be received by someone, 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 discourage participation.

---

1. from <https://collab.sakaiproject.org/portal/site/!gateway/page/!gateway-200>. CTools is an instance of the Sakai open-source application platform ([link](#)) and is currently in use by over 150 institutions.

2. All names are pseudonyms. Some names of files and other details have been omitted for privacy reasons.

3. Principal Investigator

## 2.3 Information Producers and Consumers

Group information repository users can be both information *producers*, document authors contribute content to the repository; and information *consumers*, or re-users of documents in the repository (Markus, 2001). In a personal repository like a laptop hard drive, the producer and consumer are necessarily the same person. However, in a situation where multiple users have access to a group information repository, the producer and consumer roles can be filled by the same individual, or any combination of users, resulting in a situation where a user might be trying to find documents with which she is unfamiliar or looking for familiar documents stored in unfamiliar places.

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.” And Jennifer (Site 5) described the same phenomenon: “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; all the “stuff” she thought she needed was in one place, where she put it. However, since this was the only part of the site she visited regularly, she did not know what information was available elsewhere. During the interview session 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 that Nancy (Site 3) 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, especially if they restrict themselves to the areas they are familiar with. These users prevent themselves from taking advantage of one of the greatest potential benefits of group information repositories—sharing information with other group members.

## 2.4 Naming and Organizing

Conventions are spoken or unspoken rules for how people should behave in social situations. Regardless of the importance of information stored in group information repositories, document labeling and organizing actions are generally not governed by explicit conventions; even if users agree upon conventions in advance for how documents should be labeled and organized, they find it difficult to adhere to them (Berlin, Jeffries, O'Day, Paepcke, & Wharton, 1993). Nancy (Site 3) was unhappy that members of her group all had different ideas about where things should be stored: “I

don't know, I feel like i don't really have time to sit there and enforce [the file hierarchy]... 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." Steve (Site 5) talked about how he tried to organize things such that it would make sense to other people, but all he could do was speculate about where other members of his group 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'."

Zoe (Site 3) talked about a time she'd added a file to the site, and Nancy (Site 3) thought it should live in a different place: "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 still here too, so I don't know. 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 used e-mail in an attempt to enforce her preferences, but when that didn't work she took matters into her own hands. In other social situations, conventions exist as guidelines for appropriate behavior. However, in group information repositories conventions are not effective; there is no explicit mechanism for providing feedback when conventions have been violated, or on whether the naming and organization system is working well for other members of the group.

## 2.5 Ownership and Pruning (Deleting Files)

Pruning the contents of a group information repository is an activity that information producers typically do not like to do (Markus, 2001), and increases in digital storage space mean that people are able to store more information than ever before. Users feel like they should hang onto information they are not sure they need, just in case the need might arise later. Often, later never comes, and people generally do not go back and purge without an incentive or triggering event (Whittaker & Hirschberg, 2001).

Josh (Site 6) said that the only files he might feel comfortable deleting from his group's CTools site are outdated files that he had originally 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." 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." And when asked about what kinds of things she might consider deleting, 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].” And Linda (Site 2) also described the benefit of pruning: “I think when we first started [working on the project] we just had stuff. We just had like these files everywhere. and [another group member] went in here and created the folders. She’s pretty organized. She started filing stuff. And it made it easier because we didn’t have a thousand of these little word documents.”

However, despite the potential for positive outcomes from pruning, users keep outdated files around just in case they or someone else might need them. They do not want to make a decision with consequences that might affect another member with a future need for the file, especially if they were not the original creator of the file. While there is no direct monetary cost or quota for the extra storage space, there are hidden costs in terms of lost time and effort when old information is kept around and the site becomes cluttered. For example, Nancy described the history behind how two nearly identical folders on her group’s site: “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, even during the interview, but Nancy didn’t 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 in need of presentations made by members of the group would have to look for the information in two places with different names.

## Conclusion

In this paper, I have highlighted four ways in which a group information repository can be shaped by social context. Users expect that when they put a file online in a repository, others in the group will see it; when this does not happen, it can be demotivating for the information producer, and the information consumer misses out on something that might be valuable for them to have seen. There are no explicit controls for how information should be structured or organized, and yet a structure emerges; users add their own files in the areas with which they are most familiar, regardless of whether that works for other members of the group as well. And finally, the social nature of group information repositories means that any decision to delete information can potentially affect everyone in the group, so evaluation decisions are much more difficult. Group information repositories are different from tools for personal information management, and should be analyzed not just from the perspective of the tools they employ and the information they contain, but also the social aspects of the context in which they are situated.

## References

- Berlin, L. M., Jeffries, R., O'Day, V. L., Paepcke, A., & Wharton, C. (1993). Where did you put it? issues in the design and use of a group memory. In *Sigchi conference on human factors in computing systems* (p. 23-30). Amsterdam, The Netherlands: ACM Press.
- Blair, D. C. (2002). Information retrieval and the philosophy of language. In B. Cronin (Ed.), *Annual review of information science and technology* (Vol. 37, p. 3-50). Medford, NJ: The American Society for Information Science and Technology.
- Gordon, M. D. (1997). 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.
- Jian, G., & Jeffres, L. (2006). Understanding employees' willingness to contribute to shared electronic databases: A three dimensional framework. *Communication Research*, 33(4), 242-261.
- Jones, W., & Teevan, J. (Eds.). (2007). *Personal information management*. Seattle, WA: University of Washington Press.
- Markus, L. M. (2001). Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success. *Journal of Management Information Systems*, 18(1), 57 - 93.
- Miles, M. B., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, California: Sage Publications, Inc.
- Whittaker, S., & Hirschberg, J. (2001). The character, value, and management of personal paper archives. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 8(2), 150 - 170.