

Draft – Do not cite or quote – Draft
To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

Let's Shack Up: Getting Serious about GIM

David W. McDonald

University of Washington, Seattle, USA

Living with other people is always a challenge. As children we learn to live with a small group of others, our family. That learning takes time as we balance our selfish desires with those of our siblings and parents. Often that learning is successful and we establish regular patterns, habits and norms about how space is to be negotiated.

When we choose to share physical space with another person, like when we share an apartment, cohabitate, shack-up with a significant other, or when we start our own family, our understanding of how to share physical space is challenged anew as we renegotiate how the shared space will be used.

In my case I've lived most of my life with people who are savers; people who like to save stuff, things. My mother, a community college instructor, taught business and saved textbooks – all kinds of textbooks. You want the 3rd edition of

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

Marwick's Principles of Accounting ... no problem. Interested in the textbooks that cover personal finance from the late 1970's through early 1990's ... no problem, she's got that covered too. And then when I got married – I didn't quite realize it at the time – but I married another saver.

The only real tension in this is that I like to think that I'm a deleter; I just don't like clutter. It has taken me a several years of practice, but I'm ruthless with most junk mail and random bits. If it is a catalog, advertisement, solicitation or anything of that sort, then I consider it for about 5 seconds, and if I'm not going to use it right then, it is shredded! Quarterly I work to clean out my office. Once each year I unclutter my parts of our home; I consider the various financial papers, and items in my space – if it is not serving a purpose then it just **must** go!

Negotiating a tension like this is somewhat obvious in shared physical space. I'm not saying that the solution is obvious, but that the tangibility of physical objects in a shared space make the tension clear. When you save physical stuff it starts to accumulate; it piles up. And when you delete or remove physical stuff it is gone. This seemingly simple tension is just one of several consequences of a shared information space; consequences that we have not yet resolved.

In the early 1990's I started my Ph.D. and began working with my advisor Mark Ackerman and another student Brian Starr, on the next generation of his organizational memory system (Ackerman and McDonald 1996). We were doing a version in the world wide web (very early web) and one of my responsibilities

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

was to develop ways to help group members restructure the information in the system – plant and/or transplant trees and shrubs for those of you who like to keep up with the analogy.

The paper *Where Did You Put It? Issues in the Design and Use of a Group Memory* had recently been published by Berlin et al. (1993) and was required reading for our research group as we started our design and development. The paper describes the development of a group memory system for their research group. The system is relatively straightforward including a shared repository with hypertext linking ability and a method to submit items to the repository through email. The system had both automatic and user provided classification that facilitates search and link creation. Users could browse the category scheme to locate items and perform keyword and full text search to locate items.

But the paper also presented an interesting set of side observations. During the development, the research team understood that for their system to support their own needs they would need to settle on a common vocabulary for the data they submitted. The simple task of settling on a controlled vocabulary was not as simple as they had anticipated. As some luck would have it, they videotaped their own design sessions. The analysis of their design sessions revealed some interesting results that highlight key issues for shared information spaces.

First, they found that it was nigh impossible to settle on a universal set of terms to use to organize their own information. Their admonishment to us is that,

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

“It’s not enough to agree on a set of categories.” (p. 25) This is probably because the nature of the project and relevant problems shift as a project progresses over time as well, when conducting research the terms themselves have not necessarily been created. They point out in the paper that even after they had settled on some category terms they faced a recurrent problem of reconciling individual tendencies with the agreed group norm.

The issue that Berlin et al. (1993) raises with categories is interesting given the current interest in *tagging* systems; also known as *folksonomic classification*. Large-scale web based systems such as Flickr.com and del.icio.us have popularized user defined meta-data or *tags*. These are strings of text that serve as additional meta-data for an object like a picture or another web page. The tags can be searched and clustered much like any other categorization scheme, except in this case the scheme is created by the distributed efforts of many individuals. Tagging has shown some promising characteristics with the problem identified by Berlin et al., but this is not without problems (Golder and Huberman 2006; Guy and Tonkin 2006).

Returning to the analysis that Berlin et al. conducted on their design sessions, they also found that their small group exhibited a wide range of information management practices. That is, the way they engaged the problem of storing, organizing, and re-finding their information differed. Generalizing just a

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

bit, they identified five dimensions along which individuals differ when working with information.

Purist to Proliferator: A purist generally believes that each item belongs in its one place whereas a proliferator believes that items naturally belong in multiple places.

Semanticist to Syntacticist: A semanticist believes that each item has intrinsic categories that are obvious by inspection of the item itself. A syntacticist considers the context around an item as critical to the way it should be categorized and retrieved.

Scruffy to Neatnik: A scruffy prefers coarse, fat categories with potentially many items, whereas a neatnik prefers many fine-grained and often hierarchical categories.

Saver to Deleter: A saver prefers to keep many things, including items that might be tangentially valuable. A deleter wants to limit clutter and keep only essential information.

The last dimension, *Purpose based filing*, characterizes how individuals classify items based on how they anticipate the item being used at some point in the future. This one seems a bit more binary than the other dimensions and it is mentioned here for completeness.

The challenge for any shared information system is accounting for the wide range of user approaches to saving, organizing and refinding the

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

information. Berlin and her colleagues had a term for the problem of accommodating all of these different styles – *cognitive cohabitation*. Clearly, our behavior in a shared information system is somehow a reflection of our style (cognitive or otherwise), and getting different styles to live together in an information system is a challenge of cohabitation. The challenges of cognitive cohabitation are, perhaps, not too different from the challenges of learning to live together in a shared physical space – but without all of the social and physical cues.

The Berlin et al piece influenced our thinking about not only how shared information systems work, but how people ‘cognitively cohabit’ in everyday life. As we worked on our project that would eventually become Answer Garden 2 we were careful to consider how to support different views of the same information, how to support different styles of contribution, and how to support gradually diverging models of interaction with the system. For example, users could initially share the same view of an information repository, but as one or another made changes, another person could be shown those changes and decide if they wanted to keep them or not. We didn’t handle what to do as views continued to diverge, but allowing the views to diverge was an interesting contribution to the development of shared repositories.

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

The subtle influence of this piece continued through some of my thesis work and my development of the Expertise Recommender (McDonald and Ackerman 2000; McDonald 2001). In particular our notions of how the social milieu reflect and shape expectations of individuals' expertise can be viewed as a social manifestation of cognitive cohabitation.

Yet, still, the Berlin et al paper has perhaps been underappreciated by researchers who study and build shared information systems. The dimensions that that the paper outlines form a veritable cornucopia of research possibilities. The challenge of understanding the ways in which people organize information could have some profound impacts on the design of file systems and shared information stores. Even researchers in Human Information Behavior have not directly addressed these dimensions; instead mostly opting for a scientific rationalist view of information needs, searching and retrieval behaviors.

During that last few years we have seen the development of a new area known as PIM (Personal Information Management). The approaches to PIM can be organized along two general approaches; improved search and improved structure. One approach to PIM based around search and retrieval techniques. Motivated by the success of Google and other 'network influence' based algorithms, the basic stance of researchers promoting this approach is that one can solve personal information management problems with better search. The jury is

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

still out on this. In fact, some recent results suggest that people are quite attached to the information organizing structures that they create for themselves.

The second approach is to develop tools to improve structuring of the information. Researchers taking this approach recognize that the predominate strategy for structuring information is based on the hierarchical file systems supported by modern operating systems. This approach recognizes the need to support other organizing schemes, but is often hamstrung by the existing technical constraints of the operating system (Dourish, Edwards et al. 1999).

From my perspective both approaches leave something important on the table. As someone who considers groups and their information sharing practices a fundamental problem for systems, the principle drawback to both approaches is that they are largely focused on individuals. Sure, it is always good to help individuals with their solo information problems. But for me, some of the most valuable work we accomplish is only accomplished through collaboration. That is, the way I see it, a fundamental characteristic of our work and social lives is the exchange and sharing of something important to us; information.

This implies that the real challenge for researchers, designers and developers is supporting a diverse range of behaviors when groups cognitively cohabituate – when they shack up in a shared information space. The Berlin et al. paper was not just about a specific group or organizational memory system, indeed it remains one of the earliest articulations of challenges for GIM (Group

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

Information Management (Erickson 2006)) as distinct from the current framing of PIM. This is not specific to work systems but pervades our day-to-day interactions with information. Indeed, this may be a source of the problems with tagging systems (folksonomies) described earlier.

What can we do to resolve these problems? Certainly, a better understanding of the dimensions described by Berlin et al. would be a good start. In particular, we should maintain a focus on the everyday information practices of groups currently using shared information spaces. File shares and shared web spaces are widely available and seem to frustrate groups. It is also instructive to see how critical group-reflection on collective practices can provide important insights. Lastly, self-reflection can provide some important distinctions between the group's information practices and our own.

Applying that same self-reflection I ask myself how I can resolve the challenge of being a deleter who happens to live with savers? Really, this is not as bad as it sounds. I understand why they are savers. I recognize that I'm a product of my environment – over the years I've learned when to say something about it and when to just let go. As well, since I'm a product of my environment – I recognize that I too have saver tendencies. Perhaps I'm just masquerading as a deleter and my email shows how badly I am at the masquerade.

Draft – Do not cite or quote – Draft

To appear in Erickson, T. and McDonald, D.W. (eds.). *HCI Remixed: Essays on Works that have Influenced the HCI Community*. MIT Press, 2007.

References

- Ackerman, M. S. and D. W. McDonald (1996). Answer Garden 2: Merging Organizational Memory with Collaborative Help. Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work: 97-105.
- Dourish, P., W. K. Edwards, A. LaMarca and M. Salisbury (1999). "Presto: An Experimental Architecture for Fluid Interactive Document Spaces." ACM Transactions on Computer-Human Interaction **6**(2): 133 - 161.
- Erickson, T. (2006). "From PIM to GIM: Personal Information Management in Group Contexts." Communications of the ACM **49**(1): 74-75.
- Golder, S. A. and B. A. Huberman (2006). "The Structure of Collaborative Tagging Systems." Journal of Information Science: 198-208.
- Guy, M. and E. Tonkin (2006). Folksonomies: Tidying up Tags? D-Lib Magazine. **12**.
- McDonald, D. W. (2001). Evaluating Expertise Recommendations. Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work: 214-223.
- McDonald, D. W. and M. S. Ackerman (2000). Expertise Recommender: A Flexible Recommendation System and Architecture. Proceedings of the 2000 ACM Conference on Computer-Supported Cooperative Work: 231-240.