EXPLORING PERSONAL INFORMATION

By Edward Cutrell and Susan T. Dumais

When we look for information in large or unfamiliar sources such as the Web or an encyclopedia, it is almost second nature to use a search engine to help us find what we need. Yet until recently, people have had to be more creative when trying to find information on their own computers; locating an email message, a bank statement, or a short movie clip of a friend on your own computer can be an exercise in patience and luck. While research in the area of personal information access has been ongoing for several years [1, 3], we are only now beginning to see widespread use of such tools. Rich search and browsing capabilities to support exploration are now being built into the next generation of PC operating systems (for example, Apple's Spotlight for Tiger OS X and Microsoft's Vista Search) and are also available in a variety of standalone desktop search tools.

One might wonder why people would need to "explore" their own information. An important reason is that it is difficult for users to unambiguously specify what they looking for, even in their own collections (as described by Marchionini in this section). To make matters worse, human memory is often vague and dependent on context. Therefore, users must be presented with a wide variety of techniques to articulate and refine information needs (for example, keywords, document summaries, or metadata about document content and use).

Since personal collections are often stored locally, highly dynamic interfaces can be used to help users quickly iterate queries and explore their content.

Designing interfaces for accessing personal information presents several unique challenges and opportunities. An important feature of personal collections is that people are familiar with many details and characteristics about their information, as well as the contexts surrounding their use of it. When looking for personal information, you may remember the general topic of the item, who it was from, where you filed it, or roughly when you saw it. The challenge is in creating a user interface that exploits the wide and varied associations and contextual cues that people remember about their information, while maintaining the simplicity of keyword search that makes Web search so powerful and easy. Such an interface must support search as well as browsing among many different kinds of metadata. Metadata can serve as a query (for example, "find me email that I saw yesterday"), or as a cue allowing a user to recognize an item more easily. In many ways this is similar to the category interfaces described by Hearst in this section, providing an organizing context for results and future queries.

Personal information has several other characteristics that are important in designing effective interfaces. Personal information cuts across the many "silos" of information that exist today. For example, the address of a business contact may be in your address book, an email message, a white paper document, or even in the browser history. Interfaces that focus on only a single domain, such as, email, documents, music,

photos, or real estate, are able to leverage the uniform nature of metadata to tailor the interface for that domain, (for example, see Shneiderman et al. in this section). This is much more challenging when designing for a variety of information sources. While thumbnail images can be critical for finding photos, they are almost useless for email; and the author may be important for manuscripts, but does not even exist for Instant Messaging conversations. Some kinds of metadata such as people and time are particularly useful for retrieving your own content [1]. Finally, metadata comes from many different sources—some are inherent properties of the objects (file type), some are activity based (how many times you've looked at it), and others are explicitly added by users (tags or folders). Tags are particularly interesting because they are explicitly added to aid future use of the content [2].

In an effort to explore this UI challenge, we created the Phlat interface for personal search (see the figure). Phlat combines keyword search and metadata browsing in a seamless manner, allowing people to quickly and flexibly find information based on whatever they may remember about the information they are looking for. In addition, Phlat provides a facility for tagging items with a uniform system of user-created metadata. Such tagging enables people to add information they think will be useful in getting back to their content.

A key to the design of Phlat is the tight coupling of searching and browsing. Rather than viewing search and browse as separate behaviors, Phlat treats them as two ends of a smooth continuum of exploratory search. To reinforce this unification, keyword and metadata search terms "look" very similar and are located in the same query box (in the upper left of the figure). A searcher looking for a photo may remember the name of the person who sent it to her, the approximate date it was taken, or perhaps some text in an email message about the photo. From any broad starting point, she may then rapidly filter, sort, and iterate on her query based on

what she sees and remembers until she finds it. Phlat allows for the fluid exploration of a person's own content using any of these cues.

We have deployed Phlat to about 500 volunteers at Microsoft and are currently studying how people use it. As the guest editors of this section note, evaluation of such systems can be very challenging. By studying detailed interaction logs and
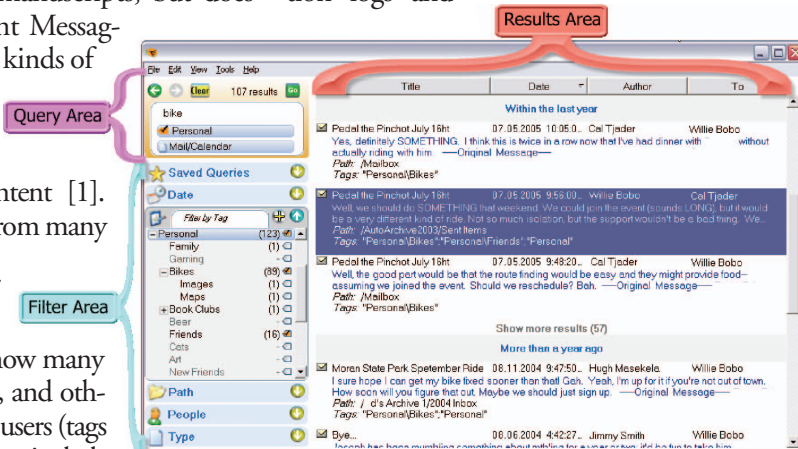
Screenshot of Phlat interface. Note the integration of search for keywords and property values in the interface.

interviewing our participants, we hope to better understand how people want to interact with rich search systems for their own content. Terabytes of personal storage will be commonplace in a few years. With systems like Phlat, we hope to make it as easy for people to find, explore, and share their own information as it is for them to acquire it in the first place. **C**

REFERENCES
1. Dumais, S.T., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R. and Robbins D.C. Stuff I've seen: A system for personal information retrieval and re-use. In *Proceedings of SIGIR 2003.*
2. Jones, W., Bruce, H., Foxley, A., and Munat, C.F. The universal labeler: Plan the project and let your information follow. In *Proceedings of ASIST'05 (2005).*
3. Quan, D., Bakshi, K., Huynh, D., and Karger, D.R. (2003). User interfaces for supporting multiple categorization. *Interact 2003–9th IFIP TC13 Intl. Conf. on HCI*, 228–235.

EDWARD CUTRELL (cutrell@microsoft.com) is a researcher in the Adaptive Systems and Interaction group at Microsoft Research, Redmond, WA.
SUSAN DUMAIS (sdumais@microsoft.com) is a senior researcher in the Adaptive Systems and Interaction group at Microsoft Research, Redmond, WA.