NSF Personal Information Management Workshop Report:

Digital Memories & Ubiquitous Computing

Mary Czerwinski (Facilitator)
Microsoft Research
http://research.microsoft.com/users/marycz

Doug Gage (Facilitator)
XPM Technologies

Jim Gemmell (Facilitator)
Microsoft Research
http://research.microsoft.com/~jgemmell/

Tiziana Catarci
University of Rome
http://www.dis.uniroma1.it/~catarci/

Cathy Marshall
Microsoft Research
http://www.csdl.tamu.edu/~marshall

Manuel Perez-Quinones
Virginia Polytechnic Institute & State University
http://perez.cs.vt.edu/

Meredith M Skeels (Scribe)
University of Washington

Scenario

Gordon returns from his business trip and the photos taken automatically from his hat-mounted camera (Figure 1 - left) begin to appear on the screen saver of his fridge (Figure 1 - right). One picture of a lunch with colleagues reminds him of an email message he wants to look at again. He knows he read the email during a meeting after the lunch. On his tablet PC, he opens the list of photos played in his screen saver, and looks up the appointment associated with photo. He sees the subsequent meeting, and requests all emails accessed during the meeting.

After finding a reviewing the email he wanted, he decides to share some photos of the trip with friends. He wants to find a particular photo, and the first thing he remembers is that it was a very hot afternoon. So, he searches for all photos taken when his personal sensor read more than 80 degrees. This returns 500 photos, so he switches to map view, and looks at where the pictures were taken (GPS has auto-located the photos). Selecting a certain neighborhood, he is able to find a good photo of what he wanted. He also browses through the photos, marks a number of them as “share with friends”, marks all the events in his calendar as public, and an attractive story is automatically created on his blog (with access limited to his friends).

The next morning, Gordon’s body sensor (Figure 1 - middle) tells him he has a fever. His analysis software notes that he has been getting sick after business trips recently, and he forwards the analysis to his doctor to get advice on how to avoid this. He still feels well enough to go to work, but can’t find his hat. The last time he remembers seeing it was after doing his laundry. He accesses the log of his washing machine and finds the last time he used it. He asks for photos taken by his room-mounted cameras in his home just after that time and quickly scans through them until he sees a shot of himself tossing the hat on his bedside table. He looks behind the table and finds the hat.

Figure 1 – Left: wearable video camera from Deja View. Middle: wearable bio-metric sensors by BodyMedia. Right: LG Internet fridge

Overview

Everything is becoming smart and networked: objects like refrigerators and pens; places like meeting rooms and living rooms. A/V capture is becoming wearable. Bio-metric sensing is blossoming. The era of abundant storage we are entering makes keeping most of one’s life possible. The era of networking promises to allow one to view and manage from any device, any place. The combination of this technology will let us capture most of our lives in a passive way, so that one will no longer need to stop interacting to become the movie or picture taker.

We discussed what one would do with a complete life of digital memories, and looked at the possible applications over the course of a lifetime. We considered reasons why one might not want to keep everything, outlined some research challenges, and also identified the unique

leverage that having a complete (across both time and data types) life record brought to PIM problems.

Description

Perhaps the first question everyone asks about a completely digitized life is: why? What would you do with it? Some of the obvious answers are:

1. Recall
    a. Find things (such as keys and eyeglasses)
    b. Replay learning and teaching experiences
    c. Review past research and trips to places
    d. Remember names of people and places
    e. Discussions in meetings
2. Share experiences with others
    a. Relive experience of lost loved ones
    b. Grandparents to grandchildren
    c. Revisit a personal experience again
3. Personal reflection and analysis
    a. Understand personal development
    b. Review conflict situations
    c. Find patterns that are common to emotional states
4. Time management
    a. Improved health via medical monitoring

Figure 2 illustrates the space of applications by who controls and uses a person's digital memories. The applications will change over the course of one's life, as will the person using the application. It early life, one's parents and caregivers will use and control one's information. For example, your parents or the babysitter may need access to your dietary and homework information. One's adult life will be a phase of personal control and use. At the end of one's life, caregivers will again take a lead role, needing to look at your medical history or helping schedule appointments in your calendar. After one's life, the executor of your will can access your financial information, while your descendants can learn about their roots from your digital memories.

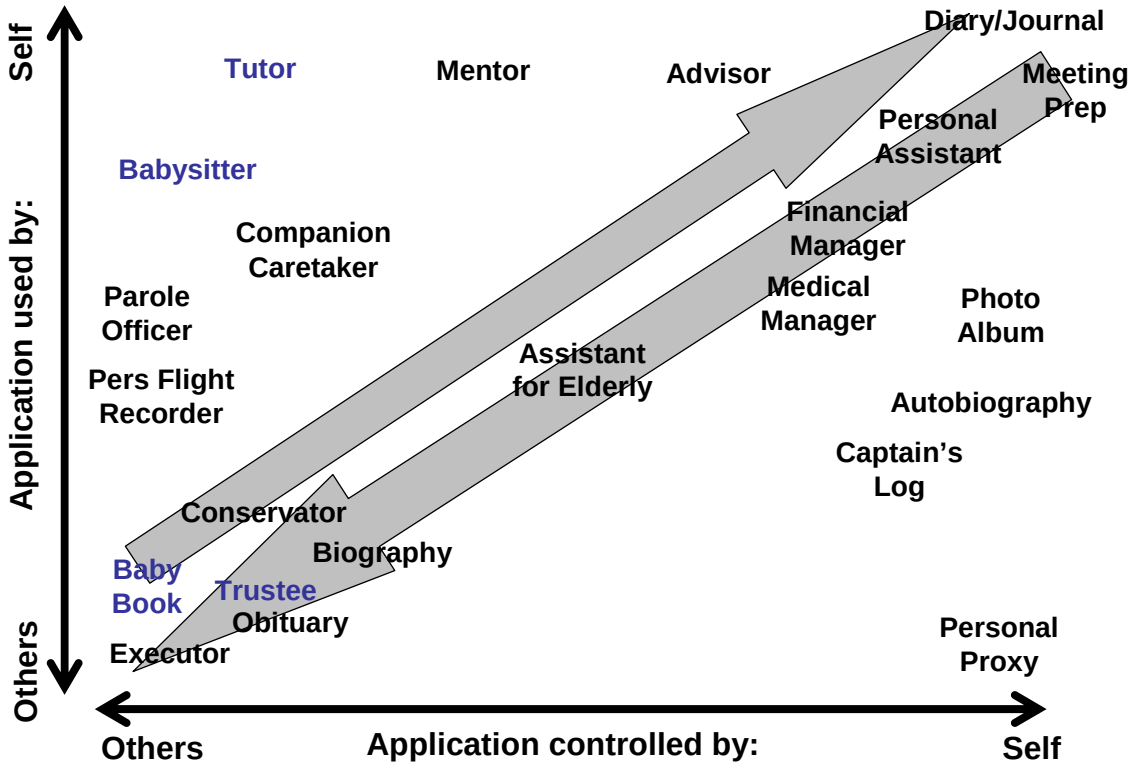

Figure 2

It is fascinating to speculate about inferences that can be derived from digital memories. Poor health might be correlated with certain locations or activities. The onset of poor health may be detected before the user is aware of the condition. Perhaps the system could even identify a pattern of poor choices in dating partners.

There are, however, arguments to be made against keeping everything in one's life. It may be that I don't want an accurate memory of the past in some instances. The memory may be painful. In fact, psychologists indicate that forgetting negative events (or at least having the member fade and become less clear) is critical to recovering from trauma. I may also want to forget for legal reasons: I don't want my digital memories subpoenaed so that they can be forced to testify against me in a way my real memories cannot. Plausible deniability may be lost if I keep a recording of everything. Furthermore, there are privacy concerns: the more I store, the more might be seen or stolen. Admittedly, this is only a quantitative difference from the privacy concerns I have with my PC today, not qualitative. Nonetheless, it makes a scary situation scarier. Perhaps it is beneficial to receive a good scare now so that we will begin dealing with the issue rather than gradually creeping towards a privacy crisis[1].

There is also a natural concern that clutter could keep one from finding things and that the increased volume would lead to more management work for the user. However, it also seems clear that anything one might have deleted could just be marked to be hidden by default; this would eliminate the clutter and management while still retaining the item for possible use in the future. Furthermore, it is often impossible to predict whether some thing will be needed again in the future. This is why some of us have large filing cabinets full of paper today: not because we will want to access every piece of paper, but because we cannot predict the very few papers that actually will needed again. Keeping as much as possible avoids destroying something that may turn out to be valuable after all.

---

[1] People are generally ignorant of the level of risk to their privacy they face right now, and don't realize how much it has been compromised already.

Fortunately, the scaling of the number and types of data is an advantage as well as a challenge. With the increased scale comes more opportunities to correlate – mostly likely based on time or place, but possibly on any common attribute – and such correlation can clearly be leveraged to help find things and to help users tell stories about their lives. For example, having a personal location record allows one to find the document you edited while in Seattle, and photos can be connected calendar events with the same time value to turn a calendar into a photo-diary.

Key research challenges

The primary research challenge for digital lifetime memories is coping with the sheer quantity of material. Summarization, abstraction, and data mining approaches must be investigated to identify “important” items, although what is important to one person is not important to another, and what is important today may not be in the future. Multiple levels of detail and resolution are desirable for all captured media, and especially sensor data.

Making use of the increasing number and types of data sources (primarily from sensors) poses another challenge. The information must be abstracted and displayed in useful and attractive visualizations. It is not clear that all the details of all of one’s devices should be part of one’s digital memories; perhaps devices will have their own digital memories and the user will only want summaries. For example, my washing machine may have the complete record of all its RPM values sampled every ten seconds, while my digital memory only cares how many loads of laundry I did.

If the question is just whether to go digital or not, security is not an issue. I can lock up my hard drive in exactly the same way that I would have locked up my papers and photos. Furthermore, it is easier to make a perfect copy and store it at another location for disaster survival. However, the convenience of access makes it desirable to attach my digital memories to the Internet – one of the key premises of this breakout is ubiquitous access. Now, instead of one locked door in one place that a burglar may attack to get at my data, I have put a digital door into millions of virtual neighborhoods for every burglar in the world to take a crack at.

Even presuming that systems can keep my private information safe, simply specifying what should be private or public is a challenge: it is critical to get right for privacy, so it must be handled correctly, but without imposing undue burden on the user. Even the binary choice of private vs. public could be onerous, but a really satisfactory solution should have designations of exactly who information gets shared with and when. Different layers of security may be desirable with different types of information having a default layer assignment.