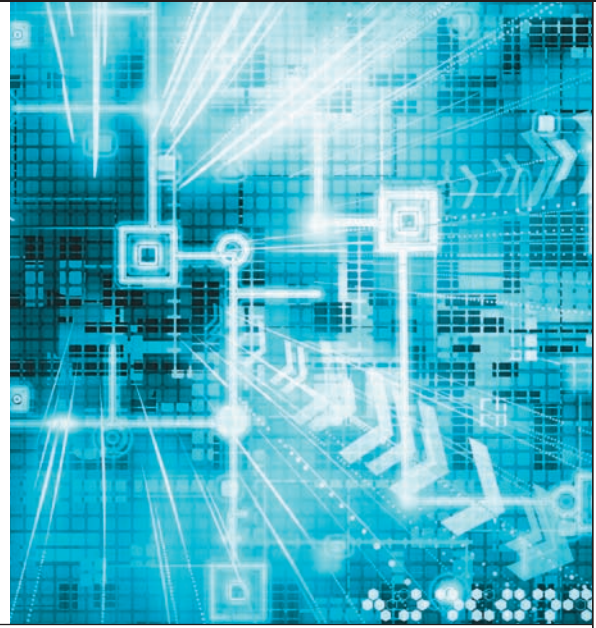# Activity-Enriched Computing: Capturing and Mining Activity Histories

James D. Hollan
University of California, San Diego

Ubiquitous data capture and visualization technologies make it possible to exploit histories of computer-mediated activity to develop activity-enriched computing.

People create, share, and organize information in myriad ways, and these activities increasingly involve computers. Yet, for all their capacity and speed, computers can be difficult and frustrating to use. Even after six decades of design evolution, computers offer little of the naturalness or flexibility of working with tangible media.

Most of our time using computers is spent within applications. Typically, each application only partially supports the task it is helping to accomplish. Completing a task might require reviewing disjoint collections of documents, spreadsheets, notes, websites, emails, and instant messages. To find and assemble all the needed information, a user might have to continuously jump between applications and data sources.

At the same time, the demands of everyday life and rapidly expanding network and mobile connectivity bring an ever growing number and variety of interruptions. Although interruptions can be essential for the smooth flow of collaborative work, resuming a task and reestablishing the often complex personal context of the interrupted activity can be difficult.

Thus, a critical research challenge is how to mitigate the impact of interruptions and help to reestablish the context of interrupted activities. One way to accomplish this is by capturing and mining activity history. Two prototype systems being developed at the Distributed Cognition and Human-Computer Interaction Laboratory at the University of California, San Diego—ChronoViz and Activity Trails—illustrate some of the potential benefits of this *activity-enriched computing*.

UBIQUITOUS CAPTURE

Traditionally, computing systems have largely ignored users' activity histories and, even when they are retained, rarely exploited them. When available, activity histories typically consist of application-specific events such as command or browser histories rather than comprehensive records of activity across multiple applications, devices, and media.

However, this is beginning to change.

As everyday life and computation become increasingly intertwined, a new generation of inexpensive sensors and digital recording devices is revolutionizing data capture, extending it to situations that typically have not been accessible and for the first time enabling examination of the structure and fine detail of human activity in real-world settings.

The ability to unobtrusively record, store, and mine activity data, coupled with a rapidly evolving pervasive computing infrastructure, have the potential to radically improve how people live, work, communicate, and collaborate.

A world in which we can digitally capture virtually everything we do, see, and hear is on the horizon, and many futurists are enthusiastic about the possibilities afforded by "lifelogging." In *Total Recall: How the E-Memory Revolution Will Change Everything* (Dutton, 2009), Gordon

Bell and Jim Gemmell described how the ability to archive our activities, augmented with biosensor and other personal data, will enable us to remember everything we have learned and experienced.

The interest in creating personal archives dates at least to Vannevar Bush's 1945 seminal vision of a memex (memory extension). From the beginning, researchers have focused on the technical challenges of capturing, storing, managing, and retrieving information from such archives. However, only recently have researchers begun to explore the practical implications for system design.

A cognitive perspective is fundamental to enabling activity-enhanced computing. In "Beyond Total Capture: A Constructive Critique of Lifelogging" (*Comm. ACM*, May 2010, pp. 70-77), Abigail Sellen and Steve Whittaker argued that rather than try to capture the totality of experience, a personal archive system should maximize human memory's strengths and minimize its weaknesses by supporting the "five Rs": recollecting, reminiscing, retrieving, reflecting, and remembering.

PERSONAL INFORMATION MANAGEMENT

In addition to the current interest in lifelogging, there is extensive literature on personal information management—primarily finding saved information. Researchers have addressed the tradeoffs in deciding what information to keep and whether to file or pile—for example, to file email messages in task-related folders or leave them in the inbox as reminders.

Variants of widely used Web search tools are available as local search applications—for example, Google Desktop and Windows Desktop Search. These applications and other experimental systems, such as Stuff I've Seen (S. Dumais et al., "Stuff I've Seen: A System for Personal Retrieval and Re-use," *Proc. 26th Ann.*

Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, ACM, 2003, pp. 72-79), demonstrate the importance of accessing rich metadata as well as content.

Personal information management systems characteristically retrieve the results of *what* users did (perhaps with some metadata), rather than *how* they did it. That is to say, the search mechanism focuses almost exclusively on semantic rather than episodic retrieval, with little support for restoring the often rich dynamic context of earlier activities.

Existing systems assume that a hybrid collection of resources—documents, instant messages, emails, webpages, and so on—is sufficient for a person to reconstruct an activity and its context. However, each resource makes available only the last state of activity. This provides neither a view of the history of interaction among resources nor how each arrived at its current state, leaving the user responsible for assembling an integrated view of the past activity's context.

Even if the appropriate set of information can be retrieved, the user must access each individual element in a separate application. Having access to this collection of resources does not reveal the chronological and progressive development of the activity.

HISTORY-ENRICHED DIGITAL OBJECTS

An effective and intuitive way to capture activity history is through visualization.

In an early research project, my research group modified an editor to collect detailed histories of users editing and reading text or code and present graphical abstractions of those histories in scrollbars, like those shown in Figure 1, to inform subsequent activity (W.C. Hill et al., "Edit Wear and Read Wear," *Proc. SIGCHI Conf. Human Factors in Computing Systems* [CHI 92], ACM, 1992, pp. 3-9).
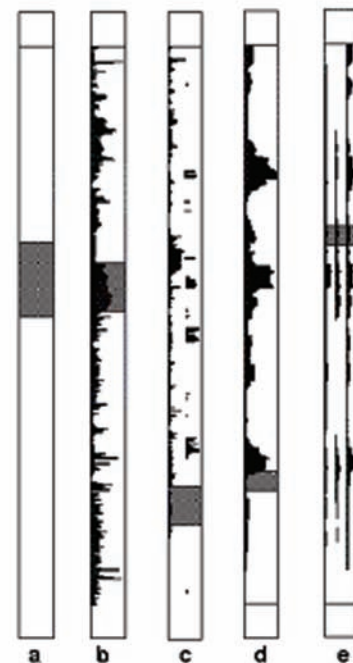

Figure 1. Attribute-mapped scrollbars in Edit Wear and Read Wear visualize user activity histories.

A key requirement of this tool was that users need not modify their behavior: they edited or read the way they normally did, and the system automatically compiled and visualized their activity history. The scrollbars represented various activity-based attributes, such as who edited a specific section or how long it took. For example, the scrollbar for a jointly authored paper might reveal that one author recently had spent considerable time editing a specific section and that another author had not yet read the changed section. Clicking on a graphical attribute took users to the associated section of the file.

HARNESSING THE DATA REVOLUTION

Today there is an opportunity to move from history-enriched objects to activity-enriched computing. By combining visualization with ubiquitous capture technologies, researchers can begin to understand, as never before, the structure and dynamics of real-world human
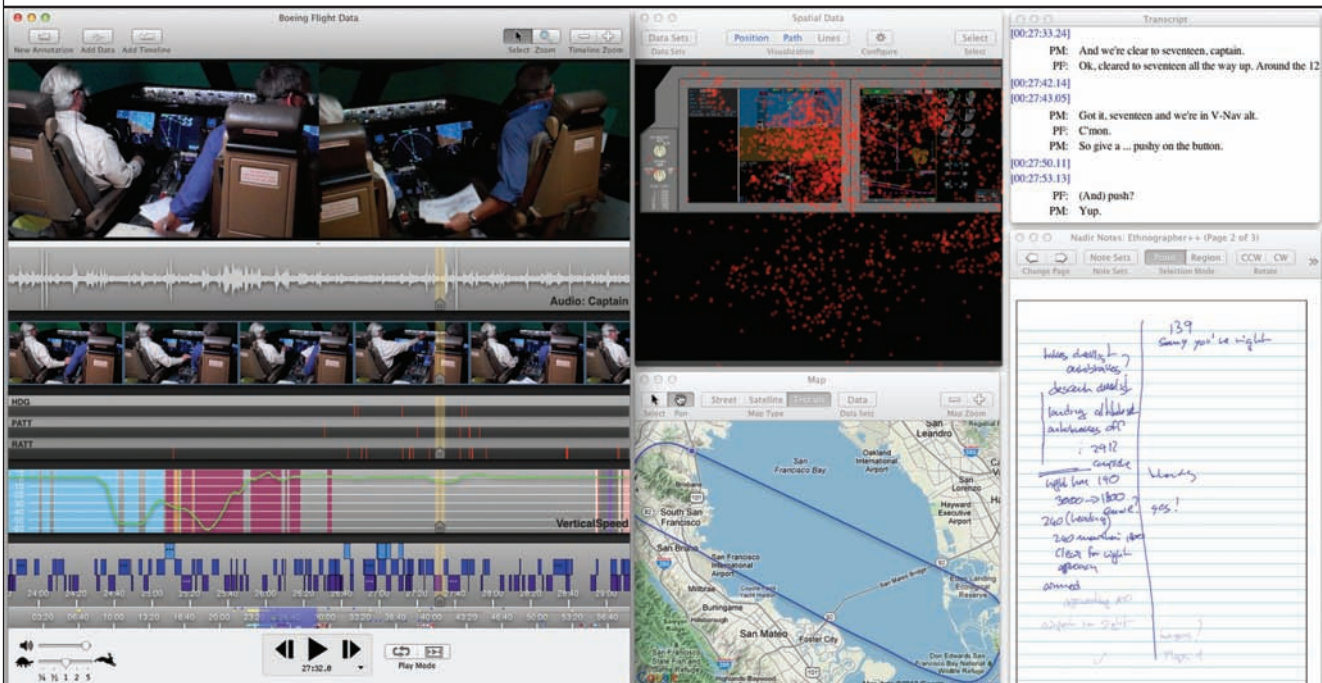
Figure 2. ChronoViz enables researchers to visualize and flexibly navigate time-based data from multiple digital sources and code the data with structured or unstructured text-based annotations. This example shows analysis of data from a flight simulation using two videos, a waveform of the audio, individual video frames, and a graph of the simulated flight's altitude with overlaid annotations.

activity. However, this requires harnessing the ongoing data revolution.

Various efforts in our lab each involve capturing activity histories, usually with video, and then analyzing those histories for trends and patterns. Examples of video data we've collected include 10 video streams recorded in a car during freeway driving, pilot activity during flight in a business jet and in a flight simulator, children interacting with a bilingual museum exhibit, interviews with deaf patients and a nonsigning physician, and seniors using a multitouch table to access medical information.

While data is crucial to scientific advances in this area, simply gathering more data will not by itself yield insights—many researchers already feel that they are drowning in data. To exploit human activity data, we also need appropriate analytical frameworks to guide research.

CHRONOVIZ

We designed the ChronoViz system (www.chronoviz.com) to aid in the analysis of multimodal time-coded activity. While many video data analysis systems are available, ChronoViz uniquely enables researchers to visualize and flexibly navigate time-based data from multiple digital sources and code the data with structured or unstructured text-based annotations.

Data sources can include video and audio files, computer logs, sensor readings, and transcriptions. Synchronizing these artifacts can be difficult given the various devices used to create them. As Figure 2 shows, ChronoViz provides interactive mechanisms to align data streams; create annotations and define and evolve category structures for them; and construct visualization filters. It can also automatically generate annotations via an analysis plug-in framework that supports custom scripts.

ChronoViz exploits Anoto digital pen technology to integrate handwritten notation with digital data from other sources. For example, it is possible to access a section of video recorded at the same time a particular note was taken or control video playback using the pen.

ACTIVITY TRAILS

Activity Trails transparently records desktop activity, allowing users to conduct their work naturally without interference. In addition to screen snapshots, it can capture applications in use, open document titles, window coordinates and focus, keyboard input, and textual context from applications such as the content of a webpage being viewed in the browser.

For example, we use symbolic aggregate approximation to transform time-series data into discrete symbols to facilitate the discovery of recurring patterns. We process word data separately and index this data with a search engine. From these two sources, we build a visual index of episodes in the form of time-stamped video frames that span applications.

Activity Trails provides a frameless widget similar to Google Desktop,

Spotlight for the Mac OS, and Enso for Windows. A program runs as an invisible background process and lets users search for episodes by typing words and specifying time ranges with a slider. The program then returns a chronology of matching events as video clip thumbnails. Users can click on these thumbnails to replay activities recorded during the episode in a larger window, as Figure 3 shows.

The key notion behind the design is to allow users to efficiently find relevant and meaningful episodes—for example, a discussion about an insurance claim in an exchange involving email, instant messages, and access to various websites. To illustrate how this might work in practice, consider the following scenario.

Gail leads a computer-vision research group at a major research university. Her group has recently developed a novel stochastic modeling technique, and she has been working on an implementation of it. She would like to demo it tomorrow as part of a National Science Foundation site visit and finally has time after dinner with the visiting team to work on it.

Gail last worked on the code just before attending a conference the week before. She can't remember where she left off and decides to examine recently changed files. She doesn't recognize some of the code and wonders if someone else has changed it. The version control system might help, but the real question is what she was planning to do next on the code.

Using Activity Trails, Gail can quickly locate images of her programming activity just before the conference. To help her recall exactly what she was doing at the time, she plays a short video clip of the activity. Nearby images on the timeline also remind her that while working on the code, she visited a former colleague's website to better understand the details of an algorithm he recently published that might make her implementation finally run in real time.
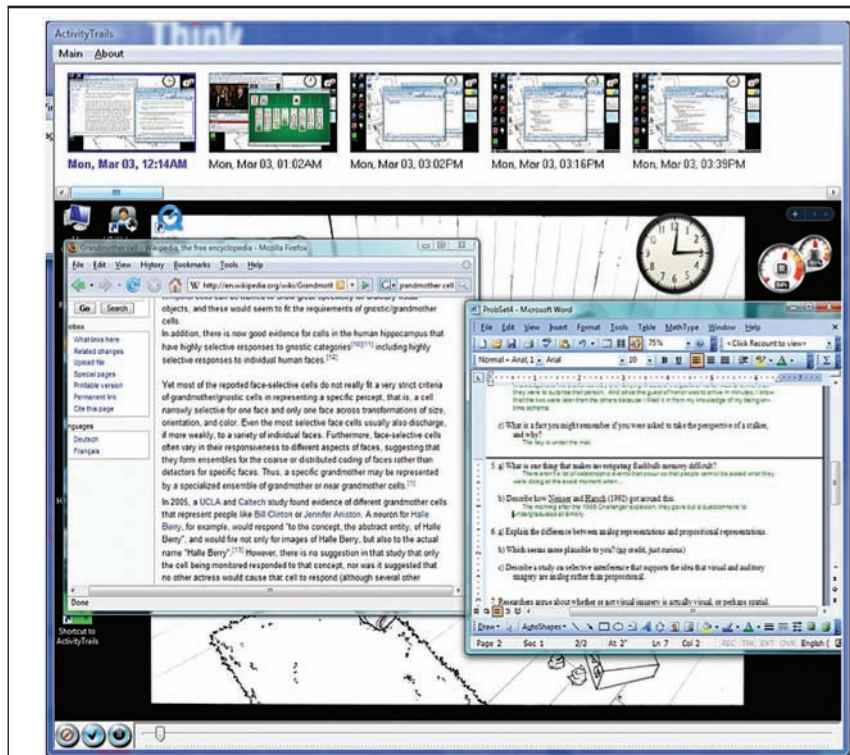

Figure 3. Activity Trails lets researchers analyze captured time-series data of desktop activity to identify salient episodes. Video clip thumbnails index episodes of activity (top). Users can click on these thumbnails to replay activities recorded during the episode in a larger window (bottom).

Seeing a mosaic of images of the website leads Gail to those sections of her colleague's paper she had spent the most time reading. This in turn reminds her of notes she made back at her office while initially reading the paper. She follows the links to the online digital version of her notebook, touches a section of the notes, and hears audio of her excited voice talking about the algorithm's potential for her present project. With context restored, she can resume her programming activity.

Of course, this simple scenario glosses over a host of research challenges, but a key insight is that it might not require a particularly complicated mechanism to radically improve reestablishment of context. In a system such as Activity Trails that exploits visual and episodic memory to restore context, representations only need to be meaningful for the person who had the experience—just as when you look at piles of papers on your desk, you see order that others do not.

ChronoViz and Activity Trails are just two examples of the many ways to capture and mine the rich history of users' computer-mediated activities. We continue to explore other exciting aspects of the emerging technology of activity-enriched computing.

James D. Hollan is a professor in the Departments of Cognitive Science and Computer Science at the University of California, San Diego, where he is co-director of the Distributed Cognition and Human-Computer Interaction Laboratory and the Ubiquitous Computing and Social Dynamics research group. Contact him at hollan@ucsd.edu.