

Capturing and Restoring the Context of Everyday Work: A Case Study at a Law Office

Gaston R. Cangiano and James D. Hollan

Distributed Cognition and Human-Computer Interaction Laboratory
Cognitive Science, University of California at San Diego
{gaston,hollan}@cogsci.ucsd.edu

Abstract. Real-world activity is complex and increasingly involves use of multiple computer applications and communication devices over extended periods of time. To understand activity at the level of detail required to provide natural and comprehensive support for it necessitates appreciating both its richness and dynamically changing context. In this article, we (1) summarize field work in which we recorded the desktop activities of workers in a law office, (2) analyze interview data in detail to show the effects of context reinstatement when viewing video summaries of past desktop activity. We conclude by discussing the implications of our results for the design of software tools to assist work in office settings.

Keywords: User behavior, empirical study, screen recording, summarization.

1 Introduction

Observational studies of office workers reveal that real-life work is highly fragmented [1,2,3,4]. For instance, Mark et al. [1] found that during the course of a typical day information workers spend an average of only 12 minutes on any given task and most continuous (uninterrupted) “events” average about 3 minutes in duration. Furthermore, the nature of information-intensive work requires people to manage a complex mix of multiple tasks and activities, each one frequently requiring different collections of applications, devices, and other resources. Typically each individual application provides only partial support for the real task it is being used to help accomplish. For example, a task might require access to a conversation, consisting of a history of emails and instant messages between groups of individuals, interleaved with examining spreadsheets, notes, and bookmarked websites relevant to this task. But each application only provides support for a single aspect of the task. Users must decompose tasks into components appropriate for various applications, assemble the resulting outcomes, and maintain the overall context needed to complete the real task.

An interesting challenge for researchers today is posed by a seeming paradox in personal computing: while most applications are designed to support a single, often simplified task, most real tasks and activities necessitate coordinated use of multiple applications. As a consequence, the information needed to conduct daily life is increasingly spread across a variety of digital resources, resulting in fragmentation of our activities and increased complexity.

While individual applications, such as word processing, continue to accrue myriad new features, there is extremely limited support for aspects of tasks that span applications. For example, while change-tracking is a useful facility for highlighting the history of editing changes in a document, it captures none of the activity accomplished outside the application. That activity is necessary for a full picture of the history of the document's evolution. Even with application suites like Microsoft Office or Open Office, integration is typically document-centric. If we are to move towards understanding and supporting real activities, what is needed is access to *episodic* views of activity within and across applications and resources.

There are, of course, advantages and reasonable software engineering motivations for developing applications that focus on specific components of tasks. The longevity and continuing evolution of applications to support tasks like document authoring, email, browsing, and search are testimony to their usefulness and to this specific decomposition of activities. On the other hand, this focus may have created the paradox to which we refer to above and may be a factor motivating increasing interest in task-support tools [5,6,7].

The goal common to task-support tools is assisting task switching and restoring associated contexts. The underlying idea dates to the seminal work on the Rooms workspace manager [8], but there is renewed interest in this area due to new empirical evidence on the nature of modern office work [1,3,4] and widespread concern about the cognitive costs of multitasking [22]. But most approaches are still based on the assumption that task context can be represented as a static state without access to the history of interactions that led to that state. In this paper we question this assumption and argue for the importance of providing episodic access to the temporal dynamics of past interaction as well as associated activities that might otherwise be deemed interruptions. The motivation for questioning this assumption derives from data of desktop activity captured in a law office and interviews with the participants while they viewed their past activity.

Below we discuss related work, describe the details of our study and the methods we used to collect data, summarize analysis of the desktop activity data and interviews with participants, and conclude with a discussion on the direct implications of our analysis for developing tools to support reestablishing context and easing continuation of past activities.

2 Related Work

Logging Activity. There is a long history and recent resurgence of interest in recording personal activity. The history dates at least to Vannevar Bush's Memex [20] and recently on a series of ACM CARPE workshops [10] on capturing, archiving, and retrieval of personal experiences. There has been little empirical evaluation of any of these systems. One exception is recent work by Sellen and her colleagues [9]. Using a personal recording device (SenseCam <http://research.microsoft.com/sensecam/>) they were able to quantitatively test recall performance using self-recorded images. The images employed were taken from a first-person point-of-view, recorded either at random intervals or initiated by the participants. They tested recall of discrete items over a period ranging from 3 days to 4 months. Overall, they found positive recall with

the use of images, but the significant differences in memory performance were only revealed after making a distinction between *remembering* and *knowing*. The former process is that by which people re-experience the past. The latter has to do with inferring that something happened, based on cues available or habitual patterns. They found significant differences between short and long-term knowing when using images to recall the past.

Task Support Systems. Research on task support has been gaining momentum in the last five years in HCI. The main notion is that a task can be defined as a discrete group of resources, such as documents in the file system, URLs, emails and so on. There is an implicit assumption in this area of research that these static groups are a sufficient representation of the *context* for a given task. The main approach therefore has been to support grouping windows, since window can represent any number of types of documents [5,6,7]. For instance the SWISH project (part of the Microsoft Vibe Technologies group <http://research.microsoft.com/vibe/>) applies machine learning technology to cluster windows based on their title and temporal proximity. Even though this kind of approach greatly facilitates switching and restoring sets of windows, it only represents a *partial* and *static* view of the context. More may be needed to reconstruct the mental representation that was active at the time the activity took place. Part of the reason why this view of context is incomplete is that it does not contain the *history* of interaction nor is flexible enough to include sporadic items that can act as memory triggers. We elaborate this further below as we describe our data in detail.

Other task support approaches provide users with ways to organize unstructured data, such as conversations and interactions, without restrictions on the type of resource involved (e.g., email, documents, IM, web pages, etc). For instance, the Haystack project [13] provides a sophisticated user interface that can access almost any element from users' activities. Haystack provides the facilities to view emails, chats, web pages and documents under user-defined categories. In principle, this type of approach could solve the support challenges we described so far, especially if histories of interactions are available for navigation. But there is an important caveat: users are required to explicitly define the parameters of the associations between resources (i.e., the categories). Having to manually define every *context* that a user is engaged in a work environment (e.g., by tagging or annotating), significantly increases the cost-benefit ratio of the approach. In addition, people don't have conscious access to the complete context of their activity nor know the memory triggers and cues needed to reinstate it, further constraining the impact of this approach.

3 Overview of Study

Over the course of the last year, we have conducted an ethnographic study of a law office. The main goal was to understand the culture, work practices, and demands of this information-intensive work setting. Our longer term goal is to design and prototype applications to assist office workers. To ground design in the real activity of the office, we built a data-collecting tool called *ActivityTrails* that enables the capture of all desktop activity. One motivation for this tool was to be able to replay specific

episodes of desktop activity to encourage participants to talk in detail about their work and the way they think about it. In addition, we conjectured that summaries and visual representations of the activity have promise in helping users to reestablish the context of previous work and to assist in resuming it.

We conducted a series of detailed interviews to learn about the tools and practices of this office setting and to explore the use of video summaries as an aid to recall of past activity.

The Office Setting. The field site for this study is an Immigration Law firm located in the San Francisco Bay Area. The firm has a staff of eight and has been in operation for over 3 years. Its members have worked together as a team for almost a decade, starting as employees of a large immigration law firm before leaving to start their current office. Shaped by their history of working together they have well established work practices for assisting clients with immigration issues. The work flow of the office is smooth and efficient.

The division labor at this office involves two main echelons: *paralegals and attorneys*. The role of attorneys is to address legal questions from the paralegals and clients, and to keep up with legal research, the industry, and government legislation. Paralegals' main job is to compile cases for filing. Compiling cases means to gather all the required documentation and information about foreign nationals in order to fill out the required forms for each visa category. It also means following up on filing deadlines, including sending notifications to clients. In addition, paralegals are responsible for maintaining updated client profiles, entering all communication, forms, and deadlines into an online database.

Because this is an Immigration Law firm, billing is primarily done on a filing basis. Unlike other types of law firms that have a need to keep track of billable hours, billing in this office is done on a per visa or work permit category basis. The focus is on keeping track of deadlines, client communication, and documentation. It is important for paralegals to keep case logs, not with the time spend on each case, but rather with a timeline of chronological interactions with clients, filing deadlines and follow up alerts. Because of the emphasis on "time-lining" rather than "time-tracking", both paralegals and attorney often multitask more frequently than they might if specific time was being charged to each client. They can address issues of many different clients in an interleaved fashion, not having to worry how much time they spend on each email or phone call. We mention this because it affects their work practice. A common pattern that emerges is that everyone usually works on multiple cases at the same time; the only exception is when a case is being compiled prior to a filing deadline or when a new case is opened and substantial data entry is needed.

4 Data Collection

Researchers from many disciplines are taking advantage of increasingly inexpensive digital video and storage facilities to assemble extensive data collections of human activity captured in real-world settings. Hollan and Hutchins [21] have argued that the ability to record and share such data has created a critical moment in the practice and scope of behavioral research. The use of video recording for the study of workplace

activity is increasingly central to advancing HCI research [15]. Lately, workplace ethnography has incorporated the use of computer screen recording [16]. One of the advantages of using screen capture is that it provides researchers with detailed records of activity. Our study makes use of screen recordings as a means to elicit data about users' perception of workflow. We employed a technique known as "auto-confrontation", proposed as an effective tool for participatory design [17].

4.1 Interviews

Our research consisted of two interview phases with four participants: one attorney and three paralegals. During the first phase the primary goal was to learn about the work practices in the law office. We recorded over fifty hours of workstation activity spanning a period of several days, and also videotaped hour long interview sessions. During the interviews, we showed participants visual summaries of their desktop activity taken from recordings (recorded not more than three days prior and edited manually to create video summaries). The interviews in the first phase were bound with a set of open-ended questions to elicit descriptions of a participant's work activities. During the second phase of interviews, our goal was to obtain qualitative data about how they described their activities and associated context while viewing visual summaries. We subsequently recorded the same participants for several days and conducted another round of interviews a week after these recordings were made.

Video summaries were generated from snapshots of the desktop taken every five seconds. These were played at one frame per second. Therefore time is compressed by a factor of five. The viewing experience is that of a slightly speeded-up movie but still very comprehensible. The figure below shows the current interface used for displaying video summaries using ActivityTrails. The bottom portion of the interface shows the playback area (a full capture of the desktop). The top area of the interface shows a timeline composed of thumbnails indexing specific episodes. Users can scrub through the video or access specific points in it by using the control at the bottom of the window, or jump in time by double-clicking on a thumbnail at the top.

The following interview excerpts are representative exemplars from our first phase of recorded interviews. The interviewer portions of the transcript and some other irrelevant data have been removed to keep the excerpt short. Line numbers are kept for referencing purposes and ellipses inserted to indicate where text was left out.

The first excerpt involves a senior attorney at the firm. She uses two computer screens on her desk. On one screen is her email inbox; on the other one are several instant messaging (IM) windows. She is writing several instant messages while simultaneously going back and forth her inbox.

- 1 this is one of my ahh legal assistant paralegals asking me a legal question
- 3 that's related to an email I had received that I am supposed to give response to this is another paralegal asking me ... and I just responded she usually just writes hi to see whether I am there and then waits to my response
- 5 ... here she is telling me to politely to review a file that she can file and so then I am going into the email to check for
- 7 for that the email that she wants me to respond to
- 9 here she's letting me know about the client I am starting to go through her emails so ah this is another legal assistant so I am am dual tasking? between IM and email

We see from the transcription that she is engaged in talking to three different people, while also going through her email. According to her report, this is something that she typically does each morning. We refer to this activity as “getting situated”. She spends a significant portion of her mornings getting situated by instant messaging everyone else in the office and looking over her inbox to see “what is going on”. Instant messaging facilitates the ability to have multiple conversations at the same time and to quickly get a glimpse of everyone’s concern.

It is interesting she uses the term “dual tasking” on line 9 above, adopting a common conception of multitasking. An alternative is to characterize it not as multitasking but as an instance of her typical morning routine of getting situated. This makes a difference in what might be relevant aspects of the context. The former case fits with an application-centric view of activity, while the later highlights her routine as central.

As a thought experiment, imagine that a month later she needs to recall the reasons she gave a particular piece of legal advice that morning. What aspects of the context of her activity from earlier would be most useful in helping her to remember or reconstruct those reasons? Should the “interruptions” and other aspects of multitasking be removed or might they actually be integral elements to assist her memory? Perhaps her legal advice came after reading an email and discussing another related issue with a paralegal. If she only has separate application-based access to individual emails or IM conversations, her ability to effectively recall her reasoning might be diminished. It takes considerable extra effort to bring the separate pieces of information together, and even then she would not have access to perceptual cues associated with the unfolding of the original episode in time to help trigger her memory.

A second excerpt involves a paralegal. In this segment, she is using her Outlook calendar, drafting an email, and switching to her browser repeatedly to navigate the case database. It turns out that she uses her calendar initially as a visual reference to know what to do next. She realizes that she needs to work on a case, but before she starts working on it (by drafting an email) she also realizes she needs some information from another case that shares a similar legal resolution to use as a reference before she can draft the email.

- 1 I am going through my calendar cuz i probably had to schedule a call
that's why i went the calendar
- 2 ahhhh hmmm then I went into another case which was the one that i was
looking previously for a resolution
- 3 and i guess i FOUND a resolution for what i needed and that's why i went
to the actual case and looked at the database and started working on the
case

This is an example of a typical use of multiple tools to accomplish a single task, in this case drafting a client email to schedule a conference call. We can see from the narrative above how she gradually recalls the entire episode while narrating the sequence of her interactions with the different tools: her calendar, her email folders and a web-based database. The history of interaction between these different tools seems fundamental to reestablishing context.

From the point of view of design, these and other episodes in our data highlight the importance of cross application and temporal cues in participants’ descriptions of meaningful episodes of their activity. For example in this case, the fact that her task was initiated by looking at her calendar, and that she then realized she needed to look at a different case for reference purposes, is information not available from access to individual applications separately.

4.2 Visual Summaries

The following interviews excerpts were obtained during the second phase of our study. We recorded continuous desktop activity for periods of a week to a week and a half. We interviewed participants using the same auto-confrontation method as in the earlier phase, one week after the recordings were made. We employed the same type of visual summaries but the guiding questions for the interviews were different. The goal for these second interviews focused on participants' reaction to and recall from the visual summaries.

The first excerpt corresponds to a paralegal describing how she resumes prior client communication. The images show her drafting an email with several interruptions from an IM conversation window. She recalls during her narrative how she was prompted by a client fax to resume prior communication.

- 1 sometimes I have to figure where i left off, and I am like should I do something new?
- 2 but then should I wait for her and then you kinda go back
- 3 NOW I AM emailing on so i got a fax i emailed XXX (attorney's name) or IM'ed XXX about the fax
- 4 and now i am emailing them saying (pause) call to discuss the fax

The capital letters in the excerpt indicate emphasis in speech. The switch to the present tense ("I AM") and use of the word "NOW" in line 3 may signal the cognitive reinstatement of an episode and its associated context. We noticed that as soon as she *shifts* tense and begins to speak of herself as if *being there*, she is able to recall the causes and details of why and how she wrote that email.

The next excerpt is from a paralegal recalling a case filing. The screen shows her having a lengthy IM conversation while having a Yahoo Portal window open in the background. She then switches to the web browser and navigates to the finance section and starts reading information about a specific company.

- 1 ok, so this guy calls or I called him so then the form or the forms came in that Fedex that we were waiting so then I called him to get a paystub so we can file the case today
- 2 and then he is saying about the XXX (company name) bought out his company XXX (other company name) or whatever
- 3 Interviewer: sooo what made you remember that?
- 4 that's just my memory I have a very odd memory cuz I just saw the word "paystub" and then i remember that this guy works for XXX (company name)
- 5 (silence pause) CAUSE I AM ON THE PHONE WITH HIM i'm on the phone with him and IM'ing XXX (attorney's name) and then because he wanted me

In line 4 above she says that seeing word "paystub" on the screen reminded her about the situation surrounding her contact with the client. This visual cue together with the name of the company showing on the screen (in the opened background window) facilitated her recall of the episode. We see in line 5 that after a pause, she raises the tone of her voice when she realizes that she was actually on the phone with the client *while* having an IM conversation and researching the company online. The full context of this episode, in this case extending to the surrounding physical context, is gradually available to her through the combination of visual and temporal cues. These types of cues appear to be a powerful and effective mechanism to reinstate a past episode, as we can see from her even recalling her physical context. Our data contains several instances similar to this, where a participant recalls the physical surroundings of an episode from viewing only images of the desktop. For example, one

of our participants recalled someone walking into her office; another participant recalled a courier service employee walking up to her desk and dropping a package.

The following excerpt involves an attorney while she researches legal websites and drafts a related email.

- 1 the news come daily so ahm I look through the emails regularly i don't scan their websites regularly but since we are in a marketing mode right now and I am thinking of redoing my website some this was a good occasion to do that
- 3 this is a website i really like (silence pause) oh and then I sent an email because I really liked it I sent an email to somebody who wants to do the marketing for me more regularly to let her know that this is one
- 4 (pause) [in lower voice] I don't remember getting a response to that I guess come to think of it

In the excerpt above we see the participant browsing through a website, following links, reading and then drafting an email. As the narrative develops, her recall indicates a rich recollection of details. For example, she remembers that she had completely forgotten to follow up on the email she drafted.

In information-intensive work environments like this one, not everything we are exposed to gets annotated, saved, bookmarked or filed. It may therefore be exceedingly valuable to have access to what we see and do on a daily basis without the need for explicit tagging or saving. Of course, much work remains on how to summarize and provide access to such rich activity records.

5 Conclusions and Future Work

Technological advances have dramatically changed work practices. Everyday work increasingly involves switching between multiple applications and responding to myriad “interruptions.” To understand these practices and thus be able to provide comprehensive support for them requires detailed characterizations. One contribution of the work reported here is a case-study description of everyday work practice in a legal office. We also briefly describe ActivityTrails, a tool developed to transparently record desktop activity and enable navigation and replaying of episodes.

Examining the fabric of everyday work activity reveals a dialectic between what people want to accomplish and how it can be accomplished within the constraints of the facilities and applications available. Although work activity from an application perspective can appear fragmented and disjoint, it is important to note that our participants still perceived themselves as working on a *single* activity in a number of those same episodes. Similarly, the many instances of self-initiated “interruptions” we saw in the data may constitute the very fabric of work patterns and function as key components of the activity context.

The increasing complexity of work and frequent use of multiple independently designed applications limit the ability of developers of any single application to provide comprehensive support for work practices. This results in increased cognitive load and fragmented contexts. We think there is promise in providing users with facilities to record activity and mechanisms to exploit those records to help restore previous context. While we have only begun to develop such facilities we are encouraged by the data we report here that demonstrates how visual summaries of past activity influence recall not only of the activity but its wider context.

Human memory is finely tuned to reconstruct the past. We argue that it is important to understand the power of visual and temporal cues to evoke the meaningful structure and context of past activity. Activity summaries, to be effective, must cut across artificial application boundaries and the resulting fragmented workflow to engage episodic memory to help reinstate the full rich context of previous activity. Such summaries will, of course, need to do more than provide access to raw activity recordings. Understanding how to create effective summaries is a key direction for future research.

We are quite a way from being able to provide the facilities needed for this scenario. Our current plan is to first continue with data collection in order to develop a quantitative measure of the effects of visual summaries for context reinstatement. Surprising little work has been done in this area and we think it is crucial to understanding how to design effective visual summaries. We are also starting to explore algorithmic mechanisms to automatically generate activity summaries at different timescales. Identifying natural activity breakpoints and meaningful scales are important steps towards summarization. In parallel with developing automatic algorithms, we will investigate the effectiveness of human identified activity breakpoints. We are especially interested in comparing how breakpoints differ when viewing one's own activity and when others are viewing it. We expect, as we saw in our study, that viewing one's own activity provides particularly rich access to the context and purpose of the activity.

We have argued in this paper for the importance of capturing and examining activity beyond the boundaries established by individual applications. We believe better understandings of the rich dynamics of human behavior in complex settings and of the dialectic nature of interaction between new digital tools and the co-evolution of new work practices is fundamental to HCI advancement. We are encouraged by the data from our initial field study that shows that reviewing activity episodes seems to trigger vivid recollections of the past. While there are many challenges involved in understanding how to effectively represent such activity histories, we see it as an exceedingly promising direction for our future research.

References

1. Mark, G., Gonzalez, V.M., Harris, J.: No task left behind?: examining the nature of fragmented work. In: *Proc. CHI 2005*, pp. 321–330 (2005)
2. Brush, A.B., Meyers, B.R., Tan, D.S., Czerwinski, M.: Understanding memory triggers for task tracking. In: *Proc. CHI 2007*, pp. 947–950 (2007)
3. Czerwinski, M., Horvitz, E., Wilhite, S.: A diary study of task switching and interruptions. In: *CHI 2004*, pp. 175–182 (2004)
4. Iqbal, S.T., Horvitz, E.: Disruption and recovery of computing tasks: field study, analysis, and directions. In: *Proc. CHI 2007*, pp. 677–686 (2007)
5. Oliver, N., Smith, G., Thakkar, C., Surendran, A.C.: SWISH: semantic analysis of window titles and switching history. In: *Proc. IUI 2006*, pp. 194–201 (2006)
6. Dragunov, A.N., Dietterich, T.G., Johnsrude, K., McLaughlin, M., Li, L., Herlocker, J.L.: TaskTracer: a desktop environment to support multi-tasking knowledge workers. In: *Proc. IUI 2005*, pp. 75–82 (2005)

7. Rattenbury, T., Canny, J.: CAAD: an automatic task support system. In: *Proc. CHI 2007*, pp. 687–696 (2007)
8. Card, S.K., Henderson, A.: Rooms: A multiple, virtual-workspace interface to support user task switching. In: *Proc. CHI 1987*, pp. 53–59 (1987)
9. Sellen, A.J., Fogg, A., Aitken, M., Hodges, S., Rother, C., Wood, K.: Do life-logging technologies support memory for the past?: an experimental study using sensecam. In: *Proc. CHI 2007*, pp. 81–90 (2007)
10. Gemmell, J., Williams, L., Wood, K., Lueder, R., Bell, G.: Passive capture and ensuing issues for a personal lifetime store. In: *Proc. CARPE 2004*, pp. 48–55 (2004)
11. Czerwinski, M., Horvitz, E.: An Investigation of Memory for Daily Computing Events. In: *Proc. British HCI Group Annual Conference* (2002)
12. Lamming, M., Brown, P., Carter, K., Eldridge, M., Flynn, M., Louie, G., Robinson, P., Sellen, A.: The Design of a Human Memory Prosthesis. *The Computer Journal* 37(3), 153–163 (1994)
13. Karger, D., Bakshi, K., Huynh, D., Quan, D., Sinha, V.: Haystack: a customizable general-purpose information management tool for end users of semistructured data. In: *CIDR 2003* (2003)
14. Ruhleder, K., Jordan, B.: Capturing Complex, Distributed Activities: Video-Based Interaction Analysis as a Component of Workplace Ethnography. In: Lee, A.S., Liebenau, J., DeGross, J.I. (eds.) *Information Systems and Qualitative Research*, pp. 246–275. Chapman and Hall, London (1997)
15. Tang, J.C., Liu, S.B., Muller, M., Lin, J., Drews, C.: Unobtrusive but invasive: using screen recording to collect field data on computer-mediated interaction. In: *Proc. CSCW 2006*, pp. 479–482 (2006)
16. Kuorinka, I.: Tools and means of implementing participatory ergonomics. In: *Papers from the Nordic Ergonomics Society Conference, April 1997*, vol. 19(4), pp. 267–270 (1997)
17. Redish, J., Wixon, D.: Task Analysis. In: Jacko, J.A., Sears, A. (eds.) *The Human Computer Interaction Handbook*, pp. 922–940. Lawrence Erlbaum Associates, Mahwah (2003)
18. Rekimoto, J.: Time-machine computing. In: *12th annual symposium on user interface software and technology*, Asheville (1999)
19. Chau, D.H., Myers, B., Faulring, A.: What to Do When Search Fails: Finding Information by Association. In: *Proc. CHI 2008* (2008)
20. Bush, V.: *As We may Think*. The Atlantic Monthly, Boston (1973)
21. Hollan, J.D., Hutchins, E.L.: Opportunities and Challenges for Augmented Environments: A Distributed Cognition Perspective. In: Lahlou, S. (ed.) *Designing User Friendly Augmented Work Environments: From Meeting Rooms to Digital Collaborative Spaces*. Springer, Heidelberg (in press, 2009)
22. Bailey, B.P., Iqbal, S.T.: Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Trans. Comput.-Hum. Interact.* 14(4), 1–28 (2008)