Foundations and Trends® in
Human–Computer Interaction
Vol. 2, No. 2 (2008) 95–171

DOI: 10.1561/1100000014

Ubiquitous Computing for Capture and Access

Khai N. Truong[1] **and Gillian R. Hayes**[2]

[1] *University of Toronto, Canada, khai@cs.toronto.edu*
[2] *University of California, Irvine, USA, gillianrh@ics.uci.edu*

Abstract

People may want to recall a multitude of experiences and information from everyday life. Human memory, however, has its limitations and can be insufficient for capturing and allowing access to salient information and important details over time. A variety of tools — primitive, analog, or digital — can complement natural memories through recording. Throughout history, in fact, record keeping and documentation have become increasingly important. In recent years, ubiquitous computing researchers have also designed and constructed mechanisms to support people in gathering, archiving, and retrieving these artifacts, a broad class of applications known as *capture and access.*

In this paper, we overview the history of documentation and recording leading broadly from primitive tools into the current age of ubiquitous computing and automatic or semi-automatic recording technologies. We present historical visions motivating much of the early computing research in this area. We then outline the key problems that have been explored in the last three decades. Additionally, we chart future research directions and potential new focus areas in this space. This paper is based on a comprehensive analysis of the literature and both our experiences and those of many of our colleagues.

Foundations and Trends® in Human–Computer Interaction

Ubiquitous Computing for Capture and Access

By Khai N. Truong and Gillian R. Hayes

Contents

1

Introduction

Vannevar Bush was perhaps the first to write about the benefits of a generalized capture and access computing system. In his 1945 *Atlantic Monthly* article, he described his vision of the *memex* — a system intended to store all the artifacts that a person comes in contact with in her everyday life and the associations that she creates between them [23]. In pointing out the need to provide capture and access as a ubiquitous service, he noted that a "record ... must be continuously extended, it must be stored, and above all it must be consulted." His envisioned system includes a desk capable of instantly displaying any file and material that the user needs. Bush also envisioned other devices to support automatic gathering of information from other daily experiences for later retrieval, such as a camera that scientists wear on their foreheads to capture pictures during an experiment and a machine to record audio dictations. The goal of these imagined devices was to support the automated capture of common everyday experiences for later review, a concept reflected in much research since his article. Czerwinski et al. [37] more recently, reflected on Bush's vision within the context of modern recording technologies, identifying five reasons people capture digital versions of their lived experiences: memory, sharing, reflection and analysis, time management, and security.

In this paper, we briefly note the origins and natural human history of recording, providing only a selective overview of the extensive literature of note-taking, documentation, and the written word. We overview the early visions of computing systems designed and developed for these purposes and detail the myriad of research works in this area — broken down by domain of inquiry. We provide some information on information management and retrieval, but focus primarily on end-user applications of ubiquitous computing for capture and access of data surrounding lived human experiences. We close with an accounting of some of the near and long-term open questions for researchers.

There is no doubt that recording, whether of histories, rules, or data, has been a significant force in human history. From the early cave paintings and hieroglyphics, to the first Bibles printed for the masses using the Guttenberg press, to the Internet and publishing that is powered by the masses, the ability to record and then share information has changed the way people are able to interact, to empower themselves, and to spread knowledge. Certainly, accompanying the advances enabled through various recording media, sacrifices and losses have also impacted human history. Practices surrounding oral histories have been lost or reduced in many cultures. Documentation has sometimes replaced human rational thought in bureaucratic organizations. Peer-review and other quality controls have been lost in some arenas to the power of inexpensive, easy publication. Likewise, the advent of ubiquitous computing and the power of automated capture and access applications have continued the trend of recording as a core human need for which technology can be an important tool, support, and effecter of change.

Fernandez [53] provides a remarkable account of the way one lawyer in colonial America remained a welcomed member of an extremely conservative community, despite his unorthodox views of religion and morality. Thomas Lechford was most likely allowed to stay in the community despite his heretical views due to his role as one of the colony's "hired hands" paid to hand copy and to write legal documents. Ironically, these activities afforded him the power to insert several of his reforms, in particular around which elements of English common law to bring to American and which to leave as well as which elements Puritan

Jurisprudence might remain. The colonists were largely unable to read his reformation writings, and even if they could, they needed him to work as colonial copier, and thus tolerated reform activities that would normally have had him exiled. This single example demonstrates what many people know to be true intuitively: those who can document and share information have access to immense power and influence upon the society for which they are recording.

The power of records may also be discerned from the emphasis placed on them by certain professional fields [47, 157, 192]. Notes can serve as memory aids, decision-making tools, historical documents, backdrops for collaborative discussion and more. Furthermore, "record-keeping practices are a central means by which organizations demonstrate accountability" [192]. Thus, note-taking often can be a "requirement of professional practice" [157], governed by rules, regulations, and "best practices." However, many of the routines, procedures, and customs surrounding recording are in fact culturally and socially constructed. The context of the interaction has as much influence on the note-taking and record-keeping practices of the individuals as the content of these records.

Additionally, keeping records holds a significant historical place in the methods of the social sciences, many of which have been incorporated into HCI research. For example, Spindler and Spindler used films of research participants as "evocative stimuli" to encourage teacher reflection on classroom behavior, primarily their own [163, p. 19]. Goodwin described how videos could be used to develop a greater understanding of interactions by cataloging those interactions using similar methods to the cataloging of speech utterances by conversational analysts [64]. He also described the ways perceptions about those activities could be molded by the coding scheme used to catalogue and analyze them [63]. One goal of using video and audio records in some cases is to prevent some of the departure from reality that can be inherent to documentation manually recorded at the time of an incident or later. Even trained observers can make errors in judgment due to their own ingrained perceptions at the time of recording [163, pp. 219–221]. Video affords the possibility to return to those experiences at a later date for further analysis. Many education researchers have also examined the ways in which video can be used in teacher preparation and critique [105, 106, 121].

Use of video as records in the social sciences has also been noted to have its downsides. Certainly, “the camera is selective” [163, p. 221]. In fact, so selective that Goodwin also described the ways coding techniques can be used to recreate “truth” from video within “socially situated, historically constituted” bodies of practice [63]. Thus, the old adage “seeing is believing” holds true. However, those who “see” using this constructed view of the video record, may see only what the constructor intends. The danger of acquiring a particular outlook regarding a setting either before or during the observation period can still be in place even when doing analysis of video records after direct observation. For example, Ochs [139] described the ways in which the process and style of transcription can influence the outcome of analysis. Significantly, she found this influence to be present whether the person doing the analysis was present for the initial observations or not.

The act of manual capture can also distract people from fully engaging in an experience. During the actual activity, people will not be able to devote full attention to the activity because they must devote time and effort towards the capture task. Conversely, when people wish to become engaged in the activity, they may not be able to take the time for recording enough details for later use. Bederson described the importance of using applications, like his NoteLens, to augment not to detract from human ability to stay in the moment [12] — in the “flow” of optimal human experience [35].

Recording adds extra challenges when considering storage of and access to that information over a long period of time. First, a large amount of information is generated, requiring intensive searching or browsing. Thus, a user’s ability to easily access a specific piece of information depends on not only where the user chooses to store the information but also how it is organized. Retrieval then becomes a matter of how to index into the collection of captured information. The ability to index into captured information flexibly and to correlate pieces of that information is important because the salient features for triggering the retrieval of the desired information can be a portion of the content or the surrounding context. Accessing content also involves more than the simple retrieval of content. Depending on the situation, users may want to recall information at different level of details.

Additionally, vast amounts of information stored initially for particular reasons can be used in very different ways once compiled together, raising a host of concerns about appropriate information use [91, 138]. In Europe, current laws prevent use of data for any purpose other than that which originally collected [52]. In the United States, however, as of this writing, no such restrictions are present.

Despite the many flaws of recording technologies and processes, keeping records is a significant part of the production of scientific and practical knowledge. By enhancing the ways in which people can document information, we can work towards reducing some of these issues. Following on the Memex vision, in 1960, Licklider presented his vision of the "mechanically extended man," describing how man–machine symbiosis can augment the human intellect in the future by freeing it from mundane tasks [112]. He emphasized the importance of the separable functions between the user and the computer in the symbiotic association based on what humans are good at (e.g., synthesizing an experience, making associations between information) and what computers are good at (e.g., capture, storage). Although he and Bush shared very similar visions, the technological progression over the 15 years between their writings helped Licklider to ground his idea with an understanding of the relevant issues and challenges that needed to be investigated, such as how to store the information and how to provide natural, ubiquitous interfaces for input and output tasks [35].

Like Bush and Licklider, Douglas Engelbart also believed that technology could be used to augment human intellect [51]. However, Engelbart believed in more than merely augmenting individual memory. In his work at the Bootstrap Institute, he coined the term "Collective IQ" to describe how a group can "leverage its collective memory, perception, planning, reasoning, foresight, and experience into applicable knowledge" to solve problems. The key factor in Collective IQ is the quality of the group's knowledge repository. However, augmentation also depends heavily on the speed and ease of creating and retrieving knowledge. Engelbart demonstrated the importance of these factors through the simple example of how tying a brick to a pencil can de-augment users.

These early visions were (and remain) inspirational to many areas of computer science. However, it was not until the beginning of the 1990s that these ideas were explored away from the desktop computer. In his seminal 1991, *Scientific American* article, Mark Weiser describes a vision of *ubiquitous computing* in which technology is seamlessly integrated into the environment and provides useful services to humans in their everyday activities [186]. Weiser described several scenarios demonstrating the benefits of automated capture and access, such as:

1. *Sal doesn't remember Mary, but she does vaguely remember the meeting. She quickly starts a search for meetings in the past two weeks with more than 6 people not previously in meetings with her, and finds the one.*
2. *Sal looks out her windows at her neighborhood. Sunlight and a fence are visible through one, and through others she sees electronic trails that have been kept for her of neighbors coming and going during the early morning.*

These scenarios illustrate two interesting ways that information automatically captured on behalf of the user could be used later for two different environments — at home and at work. However, Weiser left out many important details to inspire others to investigate creative applications for automated capture and access. The first scenario describes the user searching through a list of captured meetings for a particular meeting that satisfies the salient context about it that she remembers. However, we can imagine other desirable access behaviors that could have also helped Sal, such as content-based retrieval or browsable summaries of meetings. The second scenario, in addition to capturing nontraditional data types, demonstrates a very short-term access of the captured information, where walk trails are displayed hours afterward. In this application, the captured information is used only a short time after it occurred; it is conceivable this captured information is useful even after a long period of time passes. This scenario also introduces interesting privacy concerns relevant to the area of automated capture and access.

Ubiquitous computing technologies also can begin to remedy the selectivity and ambiguity of human memory. People often have difficulty foreseeing the value of information [188] or predicting when an event of significance (e.g., baby's first steps) might occur. Automated capture technologies can enable models of recording in which a person can determine the value of information about an event *after* that event rather than *before*.

As ubiquitous computing technologies improve, they can also reduce errors and challenges in external recorded memory as well. For example, we can reduce the selectivity of the camera by providing multiple fixed views of particular interactions. We can reduce the selectivity of coding of media by keeping the media and allowing access to varied coding schemes. Larger and more complex data provenance schemes can work towards ensuring the authenticity of the original record.

A large number of ubiquitous computing research projects focus on or incorporate technologies for capturing details of a live experience automatically and providing future access to those records. This theme of ubiquitous computing research is commonly referred to as *capture and access applications* [5]. We define *capture and access* as the task of preserving a record of some live experience for future review. Capture occurs when a tool records data that documents an experience. The *captured data* are recorded as *streams of information* that flow through time. The tools that record experiences are the *capture devices*; and the tools used to review captured experiences are the *access devices*. A *capture and access application* can exist in the simplest form through a single capture and access device or in a more complex form as a collection of capture and access devices [177]. Four common goals exist across ubiquitous computing applications focused on capture and access functionality:

1. *Capture should happen naturally*: The usefulness of this service often lies in its ability to remove the burden of recording from the human to support focusing attention on the live experience [5]. As a result, capture must be supported unobtrusively and should require little or no additional user effort. This behavior typically has been supported by (1) capturing raw audio and video of an experience and processing it later for additional semantic information or (2) augmenting the

devices that the user normally uses during an activity to log salient data.

2. *Information should be accessible with minimal effort*: The design of capture and access applications involves more than the development of unobtrusive capture services. The usefulness of these services becomes evident in the access phase when users need to review the information. Brotherton previously defined successful *access* as situations in which information can be found at the proper level of detail (as accepted by the user doing the access) with minimal effort [19]. Additionally, as Abowd and Mynatt point out, these information access services are most useful when they are ubiquitous [5]. Together, these two points indicate that users require interfaces that support access whenever, wherever, and however they need.
3. *Records should be cross-indexed by content and context*: Over a long period of time, automated capture results in the recording of a large amount of information. This amassing of data may cause users to experience difficulty in finding desired points of interest in the captured streams. To help users better navigate through these streams of information, applications often support many forms of indices so a user can jump directly to relevant points in a stream [127]. Additionally, access tools should be able to easily correlate and compute over events across streams and levels of detail, as motivated by user retrieval needs. For example, a person might remember a portion of the content or the surrounding context; as such, records should be cross-indexed by content as well as context.
4. *Records should be created, stored, and accessed in socially, ethically, and legally responsible ways*: Development of ubiquitous computing, database, and storage and processing technologies has led to a new "recording age," in which immense quantities of data about everyday life are being amassed. To ensure that users maintain a natural experience, capture technologies must be created in ways that

reflect their comfort levels with them. Likewise, to ensure appropriate protections against misuse and other social and societal challenges, storage and access portions of capture and access systems should ensure these same requirements are upheld.

These requirements stem not only from formative, philosophical, and analytical inquiries into capture and access technologies but also from a review of successful applications of ubiquitous computing technologies to the problem of large scale and naturalistic recording. In the previous work, we also identified five dimensions in the design space for capture and access applications (see Table 1.1). These dimensions are explored in more depth in the summary section as a means for understanding the applications overviewed in this paper.

In this paper, we outline the history of such projects from early visions at the dawn of computing to current and future trends. We divide these technologies and applications initially by domain area: workplace (Section 2), educational and scientific (Section 3), and personal (Section 4). This division enables the reader to follow how specific human problems have motivated and driven many researchers to design,

Table 1.1 Key dimensions in the design of capture and access applications.

Dimension	Description
When & where capture occurs	When and where does capture occur — how ubiquitous is recording? • Fixed locations vs mobile or wearable • Continuous, at scheduled times, or only when explicitly specified?
Methods for capturing and annotating the live experience	How is recording activated? How are annotations and other meta-data associated to the raw data?
Number of devices	How many and what types of recording devices are associated with the capture application? How many and what types of storage devices are associated?
Techniques for reviewing captured information	How do users access search, browse, index, and retrieve captured content?
Length of time captured information persists	How long will data persist? How will it age (e.g., automatic deletion, user-assisted deletion, time-based degradation)?

develop, and deploy capture and access technologies. In Section 5, we describe the techniques for capture and access across these domain problems, with a more explicitly technological focus. In Section 6, we close with a summary of the design space and some indication of the major open challenges that remain in this research area.

2

Capture and Access in the Workplace

Marc Weiser's article previewed many ubiquitous computing projects conducted at Xerox PARC under his supervision as well as those at Xerox's research facility in Europe (Rank Xerox Research Center, RXRC). Because Xerox at the time had a vested interest in focusing on the technology for the workplace, these projects focused on capture and access for meeting rooms and office environments.

The scenario in which Sal searches previous meetings to recall whom a person she has met in the past was based on an application developed by Lamming and Flynn. Their application, Forget-Me-Not, was the first to demonstrate the continuous capture of information for a user as she moves about an instrumented capture environment, the office [103]. Designed to leverage the ParcTab devices that a user had with her constantly [159], Forget-Me-Not continuously stored the user's location and the people she encountered. Additionally, the application captured the user's workstation activities, file exchange and printing activities, and phone call activities. Forget-Me-Not included an interface that allowed the user to navigate through the captured history or search for documents and other captured data. Users could also apply context filters to narrow their search.

Researchers at RXRC also developed many other systems to enhance a user's retrospective memory and prospective memory. Newman et al. [137] developed the Pepys diary application to collect people's movements around the office. Using the Olivetti infrared active badge network [184], the application simply logged raw location data. Techniques were then developed to extract significant episodes from the raw data. Because episodes recognized by Pepys were purely location-based information, they lacked details that might be considered useful. Included in the application was the ability to augment the Pepys Diary with video snapshots [48]. This Video Diary relied on a video network [24] instrumented throughout the office and controlled by the active badge network to take a snapshot of a person using a camera closest to her location when a significant episode is recognized. Other retrospective memory supporting applications included NoTime, a note-taking application that captures and synchronizes the user's handwritten notes with audio and video of a meeting [104], and Marcel, a system that monitored paperwork activities at an actual desk using overhead video cameras and computer vision [136]. This collection of research projects, in addition to influencing the design of the Forget-Me-Not application, also resulted in a number of general guidelines for building memory prostheses. These guidelines include the need for sensing to exist in the physical environment, support for both automatic as well as manual data capture, and the development of access interfaces that facilitate finding relevant information. Lamming et al. [102] emphasized that "a successful memory prosthesis will integrate seamlessly into the user's normal everyday activities and be available to provide help when needed."

At Xerox PARC during this time, researchers also began to investigate the benefits of capture and access technologies for small working meetings. Such meetings often involve coordinating team projects and supporting communication between the project members. As a result, these meetings are rich in knowledge that has been transferred between the participants. Researchers were largely focused on replacing or augmenting the manual capture, in the form of writing notes, being used most often in meetings. These projects were designed to reduce the burden and interference experienced from taking notes that often served to disengage one or more members of the meeting.

The Tivoli application, an influential meeting capture system, was designed to capture small co-located group meetings [144]. Tivoli focuses on the capture of interactions centered on large whiteboards. The Tivoli application runs on an electronic whiteboard, known as the LiveBoard [48]. In the degenerate case, Tivoli captures the user's annotations as raw ink. The user can also interact with the ink annotations as "domain objects," which represent specific kinds of captured information (such as Intellectual Property). As a whiteboard surface, Tivoli supports the capture of ink annotations. It also recognizes special strokes and gestures. In the simplest case, this feature allows the user to create, edit, manipulate and relate the contents on the interface [128]. With an additional understanding of the application domain, Tivoli would be able to operate on the ink annotations; for example, the application could recognize numbers and automatically add them in a math problem.

Beyond the whiteboard application, Tivoli has been extended in a number of ways. These extensions include audio salvaging capabilities, supporting ways to create annotations and indices for better playback of the captured content [129]. Tivoli is one of the few systems that have been deployed and studied over a prolonged period of use, and the evaluation results have been reported. Although deployment and field evaluations are becoming more and more common, during the late 20th century, they were nearly unheard of in ubiquitous computing. Moran et al. [129] observed the behavior of an individual user reviewing information to write reports of intellectual property meetings. In this particular work, Moran et al. provided some initial understanding of how a user would further organize and structure the information. Their study uncovered several salvaging strategies employed by the user to prepare reports about the meetings; and furthermore, it provided an understanding of how these behaviors changed over time. Moran et al. uncovered that the user eventually adopted salvaging behaviors on top of the features supported by the interface. The user created marks during the meeting to help him later index into exact portions of the audio that he would review. That evaluation was the first to study how capture and access is actually used in an authentic setting. It demonstrates the importance of evaluation because, in this instance, the user adopted an unexpected practice surrounding features provided by the interface. This work also demonstrates advanced access services such as information summarization.

In addition to capture and access of discussions centered around whiteboards and presentation surfaces becoming a common theme of exploration, Tivoli in turn inspired a number of other projects. In the remainder of this chapter, we describe projects that explore preserving details of the user's experiences in meeting rooms and offices for later access.

2.1 Supporting Planned and Structured Meetings

Meetings are an important communication and coordination activity for any group of people performing a shared task. During such an activity, much knowledge is transferred between the participants. As a result, meetings are rich in content that might need to be reviewed in the future. Because discourse and communication are important during meetings, often the task of taking notes can interfere with the participants' opportunity to fully engage in the meetings themselves. As a result, a variety of applications have been built over the years to explore the capture and access of a number of different types of meetings.

The SAAMPad application is a system that supports the capture of software architecture discussion sessions [152]. During these sessions, the rationale behind the architecture of a system is presented verbally and through architectural diagrams. The diagrams and discussions are, therefore, important aspects of the meetings that are captured and related later on. An electronic whiteboard is instrumented to capture the information that would normally be presented on a public display and ties that information with additional captured audio and video streams. Because it is known ahead of time the kind of annotations users would draw on the board, the application supports application specific objects, such as types of architectural blocks in an architectural diagram.

Similarly, in the domain of military strategic planning sessions, users would navigate over a high definition map to place annotations of military symbols on it as they present a strategic operation. This information is best captured in a way that supports full playback to allow those not present to hear how the operation should play out and the rationale behind some of the tactics. Because the set of annotations drawn over the map belong to a well-defined set of symbols, it is possible to interpret these ink strokes into recognized objects. The Rasa system applies vision techniques to capture and interpret a strategic planning session discussed over both SmartBoards as well as actual paper maps [122].

Despite their support for special domain objects, SAAMPad and Rasa capture information in similar manners to the applications for education that are described in the next chapter. In these applications, again we see the practice of either augmenting physical devices or using vision techniques previously discussed to capture information being applied to a different domain.

Project group meetings are used to discuss various aspects of a group project. Meetings can be devoted to understanding the team's progress; specifics on how important parts of the projects are implemented (or will be implemented) are sometimes presented, agendas are drawn out, and schedules and responsibilities are defined. The TeamSpace project [151] supports the capture of these meetings as multimedia meeting notes as part of a larger set of shared artifacts created and maintained for each project. Traditionally, these meetings involve multiple people who come together at a mutual location. As companies look to grow worldwide, the nature of the work place is now a distributed environment with multiple people at different geographical locations collaborating in a large project. The TeamSpace application can be launched at these different sites involved in the meeting to capture and share streams of information between the remote locations. The different streams of information the application supports include presentation slides, annotations, agenda items, action items, and video frames. Telephone connections are used to provide an audio connection between these physical spaces. Thus, audio is captured through the phone line, although potentially a voice over IP solution could be instrumented as well.

Like Tivoli, TeamSpace was deployed for a prolonged period of time. Richter [150] studied its authentic use and also conducted laboratory studies to understand how users review captured meetings. In the laboratory study, she uncovered six browsing techniques (scan, skim, jump, honing, replay, and random). However, during a study in which the technology was deployed to users, people rarely reviewed the meetings. Users review meetings when they need to recover important information that could not be otherwise obtained. From her study, Richter concluded the cost of capture must be lowered significantly beyond the need to review. Because review seldom happens capture should done at a minimal cost to the users, but capture must be performed so that it is beneficial the few times the users do review information. Richter also created the tandem Tagger and TagViewer applications to study the capture of meetings where users have a high need for such records. The Tagger application allows software engineers to annotate a text transcript of a recorded requirements gathering session. The TagViewer application allows users to review the requirements gathering sessions. In a controlled study, Richter studied the effect of the TagViewer application versus just the captured video that software engineers typically use today. She reports that users extract nearly the same number of requirements from both, though there are fewer errors when using the TagViewer application.

There are a number of other applications that provide similar functionality to TeamSpace. For example, the Workspace Navigator application from Stanford [88] captures student project group meetings within a dedicated physical workspace instrumented with cameras that capture snapshots of the room, whiteboard, and physical objects. The Workspace Navigator also recorded screenshots from the computers. The system helps students record design information for later reuse.

In comparison to other meeting capture and access applications, TeamSpace supports the capture a single collocated collaborative meeting but also has support for multiple people to collaborate in the capture of the streams of information. Similarly, the DOLPHIN application [169] was designed for capturing small freeform meetings where group members may be either in the same room or in different locations taking notes on computers. Like Tivoli, DOLPHIN also supports gesture input for interacting with domain objects created by the user. Different from most capture and access system, DOLPHIN structures the captured information as a hyperlinked set of nodes instead of temporally contiguously sheets or slides of notes. DOLPHIN supports both shared and private annotations. These systems go beyond just allowing multiple devices to control a single meeting surface (such as Pebbles [133]. More compellingly, they provide multiple people and multiple locations with the chance to participate in the capture of information. The key difference is everyone can capture information and it must be shared across all locations.

The CALO Meeting Assistant explores how to produce meaningful records of the interactions during a meeting through speech processing and dialogue modeling [46]. Whereas TeamSpace leverages explicit user-performed actions to produce rich records of meetings, including agendas and action items produced during the meetings, the CALO Meeting Assistant captures and processes audio from meetings to detect topics, question-answer pairs and action items. Through an offline interface, the system emails participants transcripts of meetings in which they participated and distilled information that each user might want [45]. The system leverages implicit user actions (such as the addition of action items onto personal to-do lists) to retrain the classifier models without explicitly asking for further user feedback.

Evaluation of these meeting capture systems proves to be an incredibly large challenge. In deployment tests of these systems, they have often been limited in their uses. Whittaker et al. proposed that this limited use was at least in part due to their limitations in provision of interfaces for Information Capture and Retrieval (ICR) and then developed and tested more advanced ICR functionality for providing the "gist" of meetings [189]. In response to this concern, Post et al. [146] developed and validated an instrument for assessing what they term *meetingware*. This instrument proved to be somewhat generalizable but is likely limited by the focus on students and on long meeting cycles, such as one might find in long-term projects.

In early phases of website design, ideas of the eventual site structure are collected and arranged through paper Post-It notes on large walls or tables, where large amounts of information can persist for as long as needed. To support this practice in the electronic world (thereby making it easier to share the information as well as to maintain design changes), the Designer's Outpost application uses two digital cameras and a rear-projected electronic whiteboard to bring the artifacts manipulated in the physical world into the electronic world [99]. The whiteboard provides a large augmented surface to which notes can be added and links and annotations can be written, much like a wall or a table in the existing practice. When a designer adds or removes paper notes from the whiteboard, the two digital cameras are used to determine the changes and to capture the notes. After each interactive Outpost session, the notes and the links and annotations on the whiteboard (effectively a sitemap) can be saved into a DENIM input file, while still physically persisting on the whiteboard for as long as needed. DENIM presents the design at many different levels such as the sitemap, storyboard, and page schematic representations of a web site and allows the user to continue to refine the design, making it a convenient (and appropriate) access application of the captured information [114].

2.2 Impromptu and Unstructured Meetings

Serendipitous encounters present the challenge of not knowing who the participants are and when the meetings could potentially occur. Because whiteboards are the site of where a lot of these types of informal meetings take place. These boards are often placed in locations where there is a reasonable flow of traffic to encourage anyone who passes to discuss ideas and to brainstorm with one another. The DUMMBO application uses a nonprojecting SmartBoard with an attached sound system, to capture informal and opportunistic meetings [21]. When anyone approaches the board and picks up a pen to write, the board automatically begins to capture the writing and discussion. After a certain period of inactivity, recording will stop. Sensing technology is instrumented near the whiteboard to detect the people present during each meeting. If two or more people are known to be near the board, then recording of the conversation will occur even if no writing appears on the electronic whiteboard. A Web interface is provided to support the access of this collection of unstructured meetings. The context of an informal meeting (who was there, when and where it occurred) is used to help an individual find a meeting of interest. Users may browse through a timeline displaying periods of activity at the board and may apply filters (who, where, when) to pinpoint a meeting of interest. Once an appropriate time period has been selected, and the correct meeting has been retrieved, the access interface allows the user to replay the whiteboard activity, synchronized with the audio.

Several years later, Hayes et al. [74] revisited the problem of *impromptu* meetings in the same context, semi-public spaces in a research building, with the BufferWare project. The BufferWare system made use of the concept of Selective Archiving, in which recording services in an environment are always on and available, but they require explicit action to save any recorded data. In that work, they found that physical, social, and experiential knowledge could help users decide how to react to and potentially to adapt and to adopt new capture and access technologies [73].

Xerox PARC's Flatland project [134] was also focused on the capture and access of informal activities, and uses time as the mechanism for retrieval of historical information. Flatland was designed to support informal activities within a private office. More structured activities in the office, such as actions performed on desktop computers, can be logged and visualized in peripheral displays as a montage of images to remind users of past actions. The Kimura project supports this type of capture and access of office activity to assist the user in managing multiple "working contexts" [117]. Specifically, Kimura allows the user to switch between working contexts by moving different "tasks" throughout the office. Within a task, there may be multiple branches of activities or contexts as well. Similarly, the Designer's Outpost, which is focused on *ad hoc* design sessions, allows individuals and groups to manage and visualize a task that branches and has multiple working contexts. They accomplish this functionality through a history interface that includes a main timeline, a local timeline and a synopsis view [100].

2.3 Personalizing Meeting Capture

Personal meeting capture can be achieved through personal electronic notebooks users would have at their seats or wherever they are. The Filochat system [188] and Dynomite [190] are built on pen-tablet computers instrumented with a soundcard. In these systems, notes can be scribbled and synchronized to captured audio content. Dynomite builds on the list of issues uncovered by the Filochat study and as a result has some additional features that include organizing the notes based on user queries and assigning keywords to blocks of ink. The benefit of such systems was later evaluated by Kalnikaité and Whittaker [89] in study which compares the accuracy, efficiency and user preference for four different capture and access system: organic memory, pen and paper, a dictaphone, and ChittyChatty (a PDA application supporting temporal co-indexing of handwritten notes and audio streams). Their study showed that when users are confident in their ability to recall information without any aid, they are less likely to use any prosthetic memory device. Users were able to remember information after seven days and thirty days using the ChittyChatty system as accurately as with a dictaphone while being more efficient with the ChittyChatty system than a dictaphone. At the same time, users were not as efficient with the ChittyChatty system as with pen and paper, but they were more accurate.

The NotePals application also allows users to each privately capture their notes during the live experience [38, 107]. After the experience, all the user's notes are gathered to form a collective view of the experience during the access phase. This approach takes into consideration that some points may be missing in different people's notes, or that the users' views may be different. The NoteLook system also supports the integration of both public and private content [26]. The NoteLook system provides users with an array of camera views that when a seminar participant requests can be used to take snapshots of the public presentation. Once the snapshot is integrated into the user's private notebook, private annotations can be placed on top of it. Additional work at FXPAL also investigated the automated video capture of meetings and generation of a comic book layout summarization of the session. This application, known as Manga, suggests high-level points of interest in the video through different image processing techniques (such as keyframing) and allows the user to drill down into video to find segments for playback [180].

Many of the applications for capture and access were originally designed in workplace settings, often by researchers at industrial labs interested in boosting workplace productivity. The examples here, which focused on personalized recording in the work setting, can also be considered as generalized free-form or structured note-taking applications. Thus, these systems can be and have been considered for a variety of domain problems and human activities.

3

Capture and Access for Educational Purposes

As greater amounts of technology are moving into classrooms and homes, teachers of all kinds — from parents to therapists to college professors — have the ability to present more information, communicate more readily, and record more experiences about the lessons at hand. As a result, some "information overload" has come about on the part of students trying to note what is important, instructors trying to prepare and administer lectures, and everyone attempting to assess progress made and learning achieved. A variety of research projects have focused on addressing these challenges from all angles, which we describe in this section.

3.1 Lectures and Higher Education

In college and even high school classrooms, today, instructors use a variety of technological tools during each lecture, with the goal of providing a richer learning experience. Often, this information is presented in the form of complex slide and multimedia presentations alongside more traditional writing on whiteboards or chalkboards. As a result, students may be drowned with information and forced into a "heads down" approach to learning. While students are busy copying down everything presented in class, they may be distracted from attending to the lecture itself. Instructors produce numerous artifacts while teaching (lecture slides, links to web sites visited in class, audio and video presentations, handwritten annotations, and spoken words), all of which students attempt to preserve in analog paper-based notes or the occasional personal voice recorder.

The eClass project (formerly known as Classroom 2000) aimed to alleviate some of the student's burden by recording much of the public lecture experience [1, 20]. To capture what the instructor writes, electronic whiteboards, such as the LiveBoard [50] or the SmartBoard[1] are employed. Prepared presentations can be automatically converted into slides displayed on an electronic whiteboard on which additional annotations can be added by hand. To capture what the instructor says and does, the classroom contains microphones used to record the audio and a single camera to capture a fixed view of the classroom. Finally, to capture other web accessible media the instructor may want to present, a web proxy was used to monitor and record the web pages visited during each class. Immediately after each class, all the different captured streams of information are processed to create an on-line multimedia-augmented set of lecture notes in a form that supports student review [22]. Storing the notes on the Web also allowed students to review the notes at their own convenience.

This work demonstrates how capture and access applications typically comprise a confederation of components that must work together seamlessly [127]. As a result, application developers spend a lot of time creating the glue that allows these independent, distributed, and heterogeneous systems to work together.

Other work that investigates the recording and playback of presentations include STREAMS [34] and Authoring on the Fly [8]. The STREAMS work introduced a technique for capturing multiple streams of information as separate, single medium streams that can be temporally integrated. As many streams of information as possible can be captured for later access, when the user may decide which stream to

[1] http://smarttech.com.

focus on as the most significant. The Authoring on the Fly system also captures any programs running on a computer as a multimedia document for later playback.

Rather than instrumenting the classroom with augmented capture devices (such as the LiveBoard as an augmented whiteboard and servers that pull web pages automatically from a logging web proxy), the Lecture Browser application [132], AutoAuditorium [17], MSR's automatic camera management system [115], Virtual Videography [62], and other whiteboard applications such as the ZombieBoard [18] and BrightBoard [164] rely on cameras and computer vision techniques to capture the materials written and presented on the boards, as well as to detect changes. Additional cameras can track people and provide different images or videos of the classroom that are integrated with the captured presentation when the system automatically generates a multimedia document for the captured lecture experience. Beyond just snapshots of the whiteboard, the Lecture Browser provides a structured interface for accessing the captured lecture; the interface uses a timeline that facilitates temporal navigation through the lecture as well as nonsequential indexing into the content. Similar both in work and in name, the MIT Lecture Browser focuses on using automatic speech recognition and natural language processing to help transcribe, summarize, and index into recorded lectures [60]. Work performed at Microsoft has investigated the summarization of the captured presentations as well [115].

The trade-off between these two approaches (the instrumentation of the classroom with augmented capture devices versus passively capturing using cameras and vision techniques) lies in the granularity of capture as well as the level of intelligence built into the capture systems. Systems using augmented physical objects the user interacts with which are able to obtain a finer level of granularity in the interaction history without needing to apply much intelligence into the system. For example, when the instructor writes on an electronic whiteboard, stroke level information can be easily obtained. Capture devices that rely on machine vision face a greater challenge to extract this level of information. For example, occlusion by the lecturer can prevent the system from seeing all of the writing as it is being written. As a result, the change detected is not at the stroke level, but at a cluster level (or a coarser level of granularity). Virtual Videography overcomes occlusions in one camera view by using multiple cameras strategically aimed at the same location. This solution assumes that at least one camera will have an unobstructed view of the scene; this situation may not always be true and furthermore the solution does not scale well. The Virtual Videography technique also post-processes captured video after the live experience. Because this processing occurs after the experience, the technique can rely on footage of the scene after the instructor moves away from the whiteboard to compute the ink written by the instructor and uses this information to fill in gaps in the captured video where she occluded the whiteboard.

3.2 Personalizing Educational Notes

Many capture and access applications for education have been focused on collaborative recording and review. Personalized notation systems and digital notebooks have emerged as a way to enhance that group capture experience or to record important information for solely an individual user.

To support the personalization of the captured lecture experience, the Student Notepad (or StuPad) system [176] was added to the eClass system to provide students with an interface that is capable of integrating the prepared presentation, digital ink annotations and Web pages browsed from the public classroom notes into each student's private notebook (during the capture phase). Students' desks were instrumented with networked video tablet technology supporting the act of writing (which is more natural for students to perform and less distracting than typing). Outside of class, it is hard to predict when and where students will review the notes; therefore, the access application was designed to run on networked computers with the more traditional keyboard/mouse interface. The personalized notes are reviewed over the Web to facilitate students to be able to review the notes anywhere anytime. The subtle difference between NoteLook (described in the personalized workplace capture section) and StuPad lies in NoteLook's reliance on the participants to devote effort and awareness (as well as a little anticipation) on when to request the public information to be added into their personal notebook.

The DEBBIE system employs much of the same features supported in StuPad [16]. In this system, what an instructor presents on the electronic whiteboard is broadcasted to computers at students' desks and is added into their electronic notebooks. Both systems have separate areas for private and public annotation, but StuPad allows students to add their annotations on top of the instructor's slides. The DEBBIE system was built as one large application that starts with the capture of the instructor's lecture as well as the students' personal annotations, while StuPad extended the base eClass system — exploring the personalization of public experiences.

More recently, Barreau et al. [9] have examined how personal capture devices, such as Microsoft SenseCam [80] and Tablet PCs can be used to capture individual experiences with biological specimens in university classes. They then integrate those experiences to create an entire classroom-learning repository [9]. In this example, rather than personal recordings being used to augment those made for the group, the personal recordings in large part become the group educational record.

3.3 Recording for Scientific Inquiry

Many scientific discoveries can only be made with the collection and examination of a large amount of data around specific phenomena. By enabling simpler collection, storage, and analysis of scientific data, capture and access systems enable researchers outside of ubiquitous computing to make advances in their own disciplines.

The Audio Notebook is a private device used to capture an experience for just its user [167]. This electronic notebook supports the recording of audio, which it integrates automatically with handwritten annotations. The Audio Notebook, initially developed as an augmented research notebook, also includes technologies such as page detection and an annotatable timeline to facilitate the access of specific portions of the captured experience. Because all information captured by the audio notebook effectively reside on the same device, information streams do not need to be integrated. However, the device uses time to synchronize and index into the captured information for playback. Because it is a mobile device, the Audio Notebook could support the capture of almost any general user experience. In her dissertation, Stifelman demonstrates the general purpose ability of this device through the study of its use in three domains: when a student uses it to take notes during class, when a journalist uses it to capture interviews, and when she uses it to take mundane everyday notes [167]. She later analyzed the results of this study and provided an understanding of how users, two students and two reporters, coped with the interface to retrieve the information they wanted to review [168]. This five-month study resulted in several key findings that can be applied in new note-taking devices. First, users often skim through notes and only review in full detail the material that was unclear during lectures or meetings that was unclear. This discovery suggests the need for the ability to skim audio at faster than normal speed, a feature Vemuri et al. [182] later accomplished.

Users also noted that playback often started in the middle of a phrase when they index the audio using page level or ink indices. This finding resulted in the development of an "audio snap-to-grid" feature using phrase detection. Stifelman et al. also processed the audio to predict topic introductions through an analysis of the pitch, pausing and energy. The Audio Notebook device indicates changes in topics on a physical scrollbar that allows the user to snap quickly to new discussion threads within the recorded meeting or lecture. Not surprisingly, Richter et al. [151] found similar user behaviors and needs in their TeamSpace project, which we will describe in the next section. In contrast to the Audio Notebook, the TeamSpace project ran for a longer period of time and supported more captured sessions and users.

Through participatory design with scientists, archivists and managers, Mackay et al. [118] developed an system called the *a-book* to augment the way that biologists currently write, annotate and interact with their notes. The system uses a graphic tablet to capture ink annotations written on paper notebooks and includes a PDA that acts as an electronic Interaction Lens that sits above the layer of ink. By capturing the users's gestures and identifying the associated documents, users can name or tag information, search, and create links between pages and with external information.

ButterflyNet is a mobile capture and access system for biologists, biology students, and other scientists working in the field [193]. Using technology-augmented notebooks and other devices, scientists can record and later find handwritten notes, photographs, sensor readings, GPS track logs, and other research content. Data recorded by the scientists manually is integrated automatically with the other data to provide structure and metadata with which to search and explore findings later.

Imaging technologies (e.g., X-RAY, MRI, CT scanning) and computer-aided diagnostic tools have greatly advanced our understanding in medical sciences (e.g., orthopedics, neurology, oncology). Likewise, researchers have begun using capture and access technologies to better understand phenomena involving externally observable human behavior. For example, Morris et al. [130] introduced the concept of *embedded assessment* for diagnosing and understanding cognitive and physical decline earlier through the long-term embedded capture of physiological data. Similarly, researchers have shown how capture and access technologies might be used to support caregivers of children with autism [70, 95]. Later, Kientz et al. described how such technologies might be used to diagnosis developmental disabilities — specifically autism [92] and Hayes et al. investigated how capture and access technologies might be used to understand behavior and outcomes related to chronic cancer care [69]. All of these projects involve capture and access technologies being used to create an archivable *image* of human *behavior*. Behavior *imaging* (BI), then, is a collection of tools and techniques that allow behavioral scientists to better understand externally observable human behaviors [2]. As more automated BI techniques are developed, based on advances in human activity recognition, they will likely enable scientists in exploring fundamental questions about the behavioral phenotypes that underlie important yet poorly understood chronic conditions or even changes in human behavior unrelated to a health condition or disorder.

3.4 Record Keeping for Childhood Education

Significant to the discussion of capture and access in education is a focus on the ways in which progress in childhood education can be measured. In the 1980s and 1990s in most of the industrialized world a trend towards concepts of "outcomes" and "performance indicators" began to appear in education. Such movements have generally fallen under the categories of outcomes-based education (OBE) and standards based education [124]. Nearly two decades later, a debate rages on within the education and sociology of education communities regarding the (de)merits of these metrics [61, 67, 123, 158, 162]. OBE has created a significant push towards documentation in many schools around the world, and thus, we summarize briefly here the primary issues and arguments inherent to this debate.

In addition to standardized record keeping for documentation and accountability, capture and access systems provide a variety of supports to educators in internal efforts. Use of video to record best practices for teachers can help identify those tacit differences in behavior that make one teacher significantly more successful than another [111]. In other schools, video has been used as a way to conduct performance assessments for hiring and promotion of teachers [135]. Others have suggested that teachers do self-evaluation using video and possibly computer technology to augment the video record [85, 65]. Furthermore, Pailliotet [141] found that teachers were able to "see" those things of which they had not been aware during an actual classroom interaction and therefore make better judgments about the children's behaviors and their own.

Several projects have explored use of video for the assessment of programs and interventions in childhood education. The VideoShare project focuses on use of video in caring for children with disabilities [183]. It encourages use of video for communicating between schools and families, increasing the effectiveness of the care team, and improving therapeutic interventions. The Walden Monitor [187] is a prototype of a mobile application that supports the capture of handwritten notes (on a TabletPC) synchronized with captured video as part of a prescribed, timed observation process. The process of recording this information was already a detailed, structured practice that required a dedicated observer. Thus, the addition of capture technologies did not disrupt the practice and at the same time provided the ability to share and review the data more easily.

CareLog [69] is a semi-automated capture system for conducting Functional Behavior Assessments (FBA) in classrooms. This system allows for teachers and school staff members to notice an inappropriate behavior after it happens, note its occurrence with a small remote control, and then automatically retrieve audio and video recordings of a classroom in the minutes before and after that notation. These recordings can be labeled, analyzed, and stored as part of a generalized educational record. As with the BufferWare project, this model of selective archiving of data was chosen to ensure that an appropriate balance of needs and concerns of users and those who would be recorded was achieved [68]. Users access stored information at multiple levels of detail, including overview graphs, the videos themselves, and summary information of the tagging they may apply to the videos (e.g., comments about the occurrence, key tags for antecedents and consequences surrounding that behavior). In a five month quasi-controlled study of its use, Hayes et al. found that using this model of capture and access greatly reduced the burden and error rates of teachers conducting assessments in classrooms. At the same time, the embedded nature of the recording equipment, the buffering model of data recording, and the integration of data access into current needs for reporting and accountability led to teachers and staff members being comfortable with the presence and use of recording in their classrooms.

Abaris [93, 94] is an automated capture system for the recording of instructional data during one-on-one Discrete Trial Training (DTT) sessions [77] usually performed at a table top or in another defined area. The very nature of DTT makes it an ideal candidate for exploring capture and access applications for recording data about individual student progress. Therapists use controlled and conditioned training sessions to help children with moderate to severe disabilities learn new skills. In a pilot deployment study, Kientz et al. [93] found that the high level of structure inherent to the therapy provided useful indices into richer video data that made that data both accessible and useful in ways it had never been before. Furthermore, progress through DTT can often be slow and hard to track. Because new intervention and instructional plans are created through the careful analysis of detailed annotations about a child's performance on particular tasks, the ability for therapists to easily access this data is a capability that was needed. Abaris supports a problem that includes a high inherent level of structure useful in capture applications. Furthermore, this domain problem includes a high need for recording, accessing and understanding instructional data that generally indicates a high likelihood for adoption, as suggested by Richter [150].

In addition to capture and access of classroom activities for student use, researchers have examined recording of classroom and one on one instructional settings for instructor analysis and record keeping. Rosensteing provides an extensive review of uses of video technologies in social science research, which can include observation, data collection, and analysis [154].

4

Continual Personal Capture and a Lifetime of Recording

Many research systems have previously focused on personal capture of live experiences, with varying levels of manual (explicit) and automatic (implicit) capture. These projects attempt to provide individuals with tools for remembering details from regular activities, both structured and unstructured. In the long term, these projects enable the documentation of a lifetime of experiences across an immense population, a concept that has come to be known as behavior imaging. In this section, we describe several of these projects and their approaches to capturing relevant information in socially appropriate ways and providing access to this information later.

4.1 Continual Personal Capture

The goal of using external tools, such as recordings and documentation, to augment human memory is certainly not a new concept. Lifelogging tools record and archive all information from one's life, which includes all the text and media that one interacts with on a daily basis as well as any visual and auditory information and other biological data from sensors on one's body. The advent of mobile and ubiquitous computing is making it possible to collect and to store ever more data about an individual and to make those data accessible anywhere at any time, functionality we refer to as continual personal capture.

As described in the early investigations chapter, Forget-Me-Not was arguably the first application to demonstrate the concept of continuously capturing a mobile user's experience in an instrumented capture environment — the office [103]. Since that work, many other research projects have revisited the same concept either in spaces rich with information, such as museums and academic conferences or continuously throughout a person's life.

Want et al. [185] have also explored this concept but from a mobile infrastructure standpoint with the Personal Server project, a device that users can carry with them capable of storing and providing access to all their personal information as needed through available devices and local wireless connections. This device contains no integral user interface, but instead can wirelessly communicate with input and output devices (via Bluetooth).

Similarly, Clarkson [29] developed a wearable system with "insect-like perception via low-resolution but wide field-of-view sensors" using microphone, front and rear facing cameras and three gyros to determine the capture system's orientation. With the goal of sensing and recording people's life patterns, this system continuously captures "I Sensed" series data, which is processed to determine high-level user activities on a day-to-day basis. In capturing and analyzing data from 100 days in one person's life, Clarkson was able to determine that such a system does not need complex techniques for classifying a user's life, but rather a simple alignment and matching techniques at the pixel level is sufficient. This is in part because, 50% of people's daily situations are typically limit to only four choices of activities.

In 2001, Bell [13] became interested in how the rapidly increasing affordability and size of disk storage can be used to fulfill Vannevar Bush's vision of the Memex. His project, initially named as CyberAll and later renamed to MyLifeBits [58], explored issues related to encoding, storing and retrieving all of the user's personal and professional information. Gemmell et al. provided rough estimates on how much disk space is required to store different media that a person would create or interact with in her life. For example, Gemmel et al. estimated that 1TB, costing <$300 by 2007, would be sufficient for holding a years worth of document, even if it were to store 9800 pictures, 2900 documents, 26 1-h audio clips, or four 1-h video per day.

In the early stages of the project, Bell made the decision not to use a database. He believed that the Windows file system organized into a relatively flat three level hierarchy, with about a dozen first level folders and an average of four folders in the second level, would provide more flexibility than a database. Bell believed that the need to maintain the database columns and metadata was an unnecessary cost. To meet the user's need, Bell believed that applications could support ordinary indexing techniques, such as temporal indexing, and searching by automatically extracting metadata from the documents. Additionally, at that time, he feared that continual personal capture would involve too much variation in document types and he did not think databases could be flexible to the moving or modifying of files.

However, the actual realization of the CyberAll vision as the MyLifeBits project actually used an SQL server that supported full-text search. The schema consisted of a table for the resources, a table for the annotation links and a third table for collection links. The resource table stores information as blobs and has additional columns for standard properties such as type, size, creation date, last modified date and a short description.

To explore storage of information beyond personal and professional content, such as all of a user's reading, presentations, and music, Gemmell et al. [59] created SenseCam, a personal, mobile, passive capture device. SenseCam automatically captures images from a person's life without her having to operate the recording equipment. Gemmell et al. designed the device to be the size of a pager that could be worn on the front of the user's body via a neck strap. SenseCam includes accelerometers, a light sensor, a temperature sensor, and a passive infrared sensor to detect motion. A PIC microcontroller polls these various sensors every second. SenseCam automatically captures images after a certain amount of time elapses or when triggered by the various sensors. Additionally, the user carries a GPS unit to record her location information. An import program uploads the images and all sensor data into MyLifeBits. This program automatically relates all properties, including GPS coordinates to the image. This work identifies a challenging design decision as to whether to store metadata with the original data or on its own. Gemmell et al. made the decision for storing these data together with the assumption that photos can be shared between users and that the metadata should easily travel with the original data when sharing. A rapid serial visual presentation (RSVP) access application allows the user to later playback the captured photos and the sensed data. A separate access interface allows users to review information based on location information as well. To demonstrate the performance of MyLifeBits for SenseCam results, Gemmell et al. replicated one day's worth of data for a year. This resulted in 318,000 SenseCam samples and over 55,000 SenseCam images. This allowed them to measure processing time for various operations (such as query and sort) performed over the MyLifeBits storage.

SenseCam researchers have also conducted numerous studies to understand how SenseCam can be used by patients with brain injury [80, 161], memory impairment [109, 110, 172], to support reflective practice by teachers [55], for personal information management [108], to create visual diaries of daily life [140] and more.

The idea of continuous capture of images, audio and other media streams also has been investigated by others. Like the SenseCam, Aizawa et al.'s LifeLog system [7] and Ellis and Lee's audio-based personal audio archive system [49] both explore the concept of using sensed context to facilitate the retrieval of captured media streams. Ellis and Lee analyzed the capture audio stream itself to segment the recordings and enabled the user to label the data so that a spectral clustering algorithm can be used to classify similar information. Similarly, the LifeLog system performs voice, face, and conversation scene detection and uses data from sensors such as a brain wave analyzer, accelerometers, and a gyro to describe the captured information in addition to time and location.

Academic conferences often have multiple tracks of concurrent activities that include paper presentations, demonstrations, special interest group meetings, etc. Conference attendees typically move about and listen to the track they find most interesting. To help remember the presentations they have seen, attendees usually take notes. However, the abundance of potentially novel and interesting information means that attendees may struggle between attempting to take notes or to synthesize the presentations. Furthermore, the large amount of information, ranging over many different topics makes it difficult to organize the notes. Dey et al. [40] created the Conference Assistant as a mobile capture and access application that allows users to take notes that automatically integrate with the tracks they attended for later review. As attendees arrive at a conference, they each receive a handheld PDA for use during the conference. Rather than requiring attendees to take detailed notes, the instrumented environment automatically captures and tags each presentation with the fixed location. Conference attendees can take summary notes on their personal, mobile devices. Because attendees are often mobile during presentation sessions, the Conference Assistant logs the user's physical location at all times through the use of RF-ID positioning technology. When an attendee reviews a talk she attended, the application uses location information to integrate the personal notes with the actual presentation.

At the Advanced Telecommunications Research Institute International (ATR), Sumi et al. [170] explored providing useful information to visitors during exhibition tours of museums, trade shows, conferences, etc. as part of the Context-aware Mobile Assistant Project (C-MAP). Based on the user's profile, which includes age, gender, experience participation type and personal interests, as well as, the user's location, an animated agent character appears on the user's mobile computer to help guide her through the physical space. Later, Sumi et al. [171] added the ComicDiary application to automatically generate a comic strip that recounts the conference attendee's experience based on sensed and manually inputted content. This application, running on the handheld device, accumulates personal contextual information, such as the touring history or list of people encountered, and provides the user with an interface for inputting the level of personal preference (for the current situation), current interest, etc. A story generation engine creates a story from the captured information and presents it as a comic strip. More recently, Korhonen et al. [101] created the Mobile Fair Diary, a system that includes a mobile SmartPhone interface for capturing experiences at a fairground and providing access to them later over the web.

Similarly, the HP Remember system allows museum visitors to author automatically generated Web pages recounting their experiences through both sensed and manually added content [54]. In this system, upon entering the environment, users receive RFID tags that can be docked at readers placed throughout the exhibit to register her presence. Additionally, users also carry wirelessly networked Pocket PCs that they can point at and communicate with Cooltown infrared beacons mounted on an exhibit. This explicit action invokes cameras instrumented in the environment to take pictures of the user and the exhibit. A record of the user's museum visit is preserved as a set of web pages that can be reviewed afterwards.

The user's mobility in the situations discussed above requires the capture applications to be also context-aware. These applications use more information than simply time to integrate the appropriate information. In particular, the sensing of people present in any given location needs to be supported. By automatically sensing or allowing users to input this piece of context, these applications can dynamically integrate all the different streams of information that the user experiences. In contrast, systems such as NoteLook and StuPad, introduced in the education and workplace sections respectively, support the personalization of captured information in settings such as classrooms or seminars, but they assume fixed user location. The NotePals application has also been used at conferences [107]. NotePals integrates a collection of personal notes based on matches in context, such as location and time. However, it operates independent from any instrumented environments.

4.2 Personal Photograph Capture and Management

Many hobbyists carry GPS units when taking pictures tag their captured photos with location information. As GPS units have continued to become smaller and cheaper, manufacturers of digital cameras have begun to package this feature in their products as well.[1] For cameras

[1] http://www.ricoh.co.jp/dc/caplio/pro_g3/.

on cell phones, location information can be obtained from the GSM network's cell IDs. The MMM prototype uses a simpler form of this information to tag captured photos with crude location information, but also to roughly predict the people's location [156]. By collecting the user location, it is possible to suggest the list of people who may be captured within a photograph at any given location. Because a single GSM network cell ID can only provide crude location information, the predicted list of people present can be inaccurate. Additionally, as geo-tagged photographs have become increasingly more common, the large collections of images have become inherently more difficult to browse as well. Jaffe et al. [86] explored how to create representative summaries of the picture collections, and how to generate Tag Maps visualization of these datasets. Their system examines visual features in images and performs hierarchical clustering on the data set to determine representative photos. In so doing, the system also identifies important locations, which are visualized on a map.

Researchers have continued to investigate how to capture additional context information that can facilitate in the organization and retrieval of these photos. As previously described, using a galvanic skin response (GSR) sensor, the StartleCam detects changes in the user's emotional state and automatically triggers a camera worn by the user to capture a picture of the surroundings [76]. In addition to detecting the user's emotional state using GSR, users of the LAFCam system also wear a microphone that continuously monitors the audio [116]. An HMM trained to detect laughter processes the audio source and annotates captured video with points in the captured stream when the user laughed. This interface can simplify the video editing process by marking the most interesting parts, as detected by the sensors.

In contrast to the previously mentioned camera projects that capture context as a separate stream of information, Håkansson et al. [66] explored how to add *within* a photo the pieces of context that can be easily sensed, including movements, sound, temperature, pollution, humidity, smell, and electromagnetic fields. Inspired by lomography, the art of taking photographs that capture spontaneous moments that include unpredictable color and lighting effects using special cameras and a "don't think, just shoot" mentality, Håkansson et al. used sensor data to control the hue, saturation and value of an image, thereby encoding the sensed context into the captured photograph, itself.

Although a photograph or a video clip obviously can capture people within a scene, an access application can not easily extract this piece of context from the content to facilitate users in searching for the specific picture or video segments they wish to review. Patel and Abowd developed the ContextCam device to capture video that is automatically tagged with context information [142]. ContextCam detects people present within the field of view of the camera through an active tagging scheme that assumes people wear badges that transmit ultrasound and radio frequency signals. The ContextCam device includes a pair of RF and ultrasound readers along the front of the device to triangulate the distance and position of the signals it reads from the device. Based on the zoom level, the camera can determine the list of people present in its field of view. The camera can also determine the list of people nearby but not in the field of view. ContextCam encodes the list of people it detects as well as time and location information directly into the captured video. The camera modifies every sixtieth video frame. ContextCam encodes metadata information into the least significant bits of the eight-bit RGB value for each pixel of the video frame. As a result, how much metadata needs to be stored determines how much the camera adjusts the quality of these video frames. Patel and Abowd developed ContextCam as a point of capture device with the intention of creating video content that would eventually feed into the Family Video Archive application.

The Family Video Archive application provides the user with an interface for manually tagging captured video with high-level context that cannot be sensed easily, as well as, an interface for searching the archive [3]. Taking the notion of capturing and accessing family memories one step further, Stevens et al. [166] developed the Living Memory Box. Based on extensive fieldwork, the design of the Living Memory Box enables families to store and review a variety of multimedia documenting family experiences.

4.3 Managing Personal Captured Data

Although the works described in the preceding section focus on adding metadata information to captured content to facilitate access, with the exception of the Family Video Archive application, they do not actually investigate the storage and retrieval aspect of the problem. Many of these works simply store the information streams within a hierarchical file structure.

However, as Gemmell et al. observed after struggling with their initial choice of a file system, a database back-end provides the most flexibility for storing and organizing captured information. As information retrieval highly depends on what users can remember about a document, property-based storage systems have been proposed as an alternative solution to hierarchical storage systems. Property-based storage systems better support user interaction with documents through document attributes. A number of systems use (or can use) information's natural space and time attributes to manage, organize and visualize documents. Time-based systems, such as LifeStreams [56] and TimeScape [148] organize documents based on time and provide special visualizations. LifeStreams archives documents as a time-ordered "stream" of documents that are displayed stacked diagonally across the display. Users can scroll this stack and click on documents to interact with them. TimeScape takes snapshots of the desktop workspace, and the desktop can be "played back," meaning that individuals can visually examine their desktops at any past point in time. For example, a user could retrieve the document she placed "in the top right corner last year." It also has a horizontal timeline view of the documents and uses fading to handle clutter. Both LifeStreams and TimeScape allow users to retrieve retrospective content, but also enable users to create prospective content as well. By inserting information with time anchors from the future, users can insert reminders to act upon at a later time.

Haystack [6] takes a similarly flexible approach, relying on relational data structures, hierarchical file structures, and user-generated metadata. In this solution, a data model enables representation of interesting information that is inherently searchable and machine-readable. A user interface provides access to and visualization of those data, allowing individuals to browse, edit, and share the data. Furthermore, Haystack uses machine intelligence to attempt to understand and proactively act on user needs and preferences, adapting its behavior over time. Finally, the Haystack model enables sharing of data models between individuals in a collaborative workspace. In this way, Haystack provides both a relational and a user-centric approach to data storage.

In the Presto project, later renamed to Placeless and finally Harland, [43] investigated a more flexible model for interacting with documents that uses arbitrary properties instead of simply temporal context for storing and retrieving information. Users can retrieve a document (including those in mail boxes or mounted over network file systems) by specifying a query for known properties of that desired document (e.g., author, topic). The use of properties to store and search captured written notes has been explored in the NotePals application as well [82]. However, Presto also supports executable properties which embody autonomous behaviors that help the document space and applications reorganize the documents and the define ways to interact with them.

Focused less on documents and notes manually created and accessed by users, Rhodes and Starner [149] developed the Remembrance Agent to examine how all the information a user has previously seen and recorded can be related and made useful to her. This application demonstrated the potential benefit of having captured information available after long periods of time beyond the initial live experience. Rhodes and Starner intended for the Remembrance application to run on a wearable computer that the user always has on her body. As the user interacts with a document, the system determines the local context or keywords related to that file and attempts to automatically remind the user of other documents she has viewed in the past that are similar. Similarly, more recently, Starner developed a wearable capture application for documenting military field patrols that records the entire patrol and enables access to it later [165]. For more traditional computer users, Cowley et al. [33] created the Glass Box, which unobtrusively captures a history of all computer activity for security analysts. The captured data are used by these analysts collectively for research into the development of new tools for the intelligence community and can also provide glimpses into trends across analysts to address security threats.

A variety of research projects have focused on how people deal with small bits of information in everyday life. Many of these projects have focused on calendaring, an area slightly out of scope of capture and access and thus not covered in detail here. Others, however, are more concerned with capturing, storing, and providing access generally to snippets of personal information.

The Jourknow project focuses on "information scraps," which they define as "personal information where content has been scribbled on Post-it notes, scrawled on the corners of sheets of paper, stuck in our pockets, sent in e-mail messages to ourselves, and stashed in miscellaneous digital text files" [14]. They note that information scraps are commonly used for temporary storage, cognitive support, archiving, reminding, and as a catch-all for unusual information that might be important in someone's life. This group has now begun to develop tools to support these practices [14, 15, 181] based on these results and those found in other work [10, 25, 72], Lee et al. [108] capitalized in the concept of small and often ephemeral information to explore how a mobile automatic picture taking device, the Microsoft SenseCam, could be used as a personal information management system.

Taking a slightly different approach, other researchers have focused on history of digital documents already being used as the major repository for personal information management. For example, Dumais et al. [44] developed and evaluated the "Stuff I've Seen" system for reminding users of appointments, documents, web pages, email, and other electronic artifacts they have seen in the past. Based on this approach, Cutrell et al. [36] then argue that search, a cornerstone of capture and access, would significantly reduce many of the issues of personal information management.

4.4 Health and Wellness

Physicians have long used patient-recorded data to assist with diagnosis of often-mysterious symptoms. For example, when attempting to diagnose a potential food allergy, doctors may request that a patient keep a detailed food diary to be analyzed alongside a symptom diary in which patients record everything eaten and every symptom experienced. Despite continued growth in methods and technologies for documenting and managing personal heath information, people still experience a variety of challenges related to personal health information management [147]. Civan et al. [28] found that the "fragmentation of personal health information and reliance on human memory" are two of these significant challenges. The aggregation of information for ease in both recall and analysis is a major goal of capture and access technologies, as previously described. Thus, these technologies may represent a potential part of the solution of managing this information as it grows in scope and complexity.

As telehealth technologies have become more prominent, the tools used to gather this sort of everyday information have improved. As another example, in diagnosing heart abnormalities, cardiologists often employ mobile event recorders asking patients to push a button that allows for the recording of detailed medical data when the patient feels symptoms not just when the patient is in the office. The Point-of-Care Engineering Laboratory at Oregon Health Sciences University focuses on developing new technologies that can help monitor and diagnosis cognitive decline in the aging population early by inserting these technologies into everyday activities [75].

The Careview project [119] focuses on the capture and access of medical information to support work practices and information needs of homecare nurses. A field study of nurses led to the development of interface design guidelines for displaying clinical data including temporal visualization, integrated data gathering and data analysis, and hands-free speech-driven operation. Similarly, the LifeLines project [145] focused on visualizing patient data over long periods of time.

Xu et al. [191] developed a system to support the capture of the physical therapy sessions for stroke patients. They focused on assisting patients with reaching and grasping movements. It supported the real-time collaborative annotation and visualization of the therapy activities, producing an interactive record of the rehabilitation history.

Other researchers have examined the idea that data captured about a patient might need to be shared with all of the caregivers for that person. In the Computer Supported Cooperative Care and Proactive Health projects from Intel [32, 131], researchers utilized a variety of displays to reflect back information about a person as well as to share that information with her caregivers. Similarly, the Digital Family Portrait project focused on communication of behavioral data about an elder with one significant caregiver as part of a long term deployment of a working sensor and display system [155]. The CLever Project at the Coleman Institute at the University of Colorado[2] strives to develop computationally enhanced environments to assist not only people with a wide range of cognitive disabilities, but also their support community.

Researchers have also explored the way that information about an individual's health and behavior reflected back at them can alter the ways they behave or the ways they think about their own health. Mamykina et al. [120] explored the ways patients with diabetes think about their own health by providing them extra information about blood sugar levels and other data using mobile technologies. In addition to treating disorders, illnesses, and health challenges, capture and access technologies have been used to encourage positive activities related to general health and wellness, such as exercise or proper eating. For example, Lin et al. [113] examined the ways reflection of not only the individual's activity, but also the activity of a competitive group can affect the ways people think about and enact exercise routines. As another example, Consolvo et al. [30, 31] focused on tracking exercise behavior over weeks and even months and reporting progress back to users through a "glanceable display" on a mobile phone. In studies focused on deploying these devices and measuring their impact, the researchers found that behavior not only changed in the short term but was also maintained over longer periods when the mobile capture and access system was used by participants with the goal of improving their exercise routine compliance.

[2] http://l3d.cs.colorado.edu/clever/index.html.

Recently, in addition to research applications for health and wellness, much attention has been turned to commercial efforts to document personal health information. Websites for personal health records provide an online repository for health information, in some cases including data from physicians and labs, but in other cases simply including patient-entered subjective data.

Although there is a potentially innumerable set of these sites, many of them provide similar services. We describe six such sites here to give an overview of their potential.

WorldMedcard is a free service for end users that also includes a paid and branded service for medical professionals, hospitals, and so on. In addition to the online services, WorldMedcard encourages users to print a paper card that includes emergency contacts, medications, allergies, brief medical history, and contact information. Microsoft's health records product, the HealthVault, also allows end users to store health information for free and uses a Microsoft LiveID for login and security, enabling seamless integration into other Microsoft online products. HealthVault allows the import of data from a variety of external services and devices, such as blood glucometers and scales. RevolutionHealth acts more as a health portal, like WebMD or other wellness related sites, providing articles and information on a wide variety of health related concerns. The site also includes the ability to manage personal health information. Google Health allows importing health records from a variety of external sources including pharmacies and medical treatment centers as well as other smaller medical records web sites (e.g., MyMedicalRecords.com). Other sites are less focused on storing full personal health record for the general population, and instead focus on particular needs. For example, MyCycle.com specifically targets women tracking their menstrual cycles. The site provides personal information as well as chat forums and general information for users. Likewise, PatientsLikeMe.com only allows patients with specific diagnoses (e.g., HIV-positive or Multiple Sclerosis) to join the site but then supports the sharing of medically relevant information with other users for the greater knowledge of all.

Like with all capture and access technologies, personal health and wellness data storage technologies have the potential to disrupt and alter carefully negotiated boundaries around privacy, control, and autonomy [57]. Government regulations, such as HIPPA in the United States [79], work towards addressing some of these concerns, but the exact scope and control of personal health records (as opposed to the more traditional medical records from treatment centers, hospitals, and pharmacies) is not yet clear.

5

Techniques for Capture and Access

In the previous sections, we reviewed a large set of applications and uses of capture and access technologies. In these applications, recording can happen passively (e.g., through cameras and microphones) or through the active usage of tools augmented with capture capabilities. If enacted automatically, capture can be triggered by sensors or computational logic, and sometime manually, such as by a human pushing a button. A capture system may include a basic access interface that supports the searching, indexing, and playback of captured content. To facilitate retrieval of particular salient information from the large amount of content captured over time, an interface could also provide a visualization and automatic summarization of this collection of data as a way to help users locate high-level points of interest and then allows the user to drill down to specific points in the stream for playback. In this section, we focus on the specific trade-offs between these techniques for capture and access.

5.1 Always-On Passive Capture

To allow users to review content from their past, that information first must be recorded. One strategy for ensuring that the information exists during access is to always be capturing data. Applications, such as SenseCam, passively capture information from a person's life without requiring that user to operate the recording equipment.

Always-on recording has two specific disadvantages. First, it produces a large amount of data making it difficult for a person to retrieve previously captured content using only time information. As a result, these applications also may include a variety of sensors to capture key context information that can be used to filter the captured content during access. For example, the Ubiquitous Home [39] always records video from a collection of cameras embedded throughout the home. It detects the users' footsteps to determine their locations within the home. In doing so, the system can always produce a continuous video of any user in the home by properly segmenting the captured data and locating the video stream that has the user in it from the appropriate camera source. The system also performs data analysis and clustering to produce summarizations to facilitate review of the content. Second, always-on recording produces an environment in which the user is constantly monitored. The recording may produce unnecessary discomfort [11, 73], especially when it is unclear if all the captured data will be useful [83].

5.2 Manually Initiated Explicit Capture

Instead of always recording the user, a system can allow the user to initiate the recording manually. In fast, this mode of interacting with recording devices is the most familiar the average end user. Whether using a hand-held video or still camera or recording a conversation or lecture using a handheld audio recorder, most adults in industrialized nations have recorded multimedia at some time or another. When considering using this approach to recording in full scale capture and access systems, however, the end result is often to augment the tools already in use for a particular activity with capture capabilities. For example, the Audio Notebook records content when the user interacts with the device and does not capture otherwise [167]. Similarly, in systems such as eClass, the user starts the recording of a session only when capture is desired [1]. This necessitates a user remembering to initiate capture. Manual intervention then provides the benefit of protecting users against the threat of unwanted automatic recording. However, this model also requires users to foresee the value of recording an event before it occurs.

5.3 Buffering and Selective Archiving

Many researchers have explored buffering and near-term access to recorded data as a means of providing augmentation to human memory without the burden of long-term recording. These applications represent a wide variety of technological design choices as well as types of media recorded and uses of those recordings.

Hindus and Schmandt [78] created a near-term audio reminder service, known as Xcapture. This service provided a "digital tape loop" of audio for a single office. It was later augmented to provide a short-term audio memory of telephone conversations (5–15 min long). This application runs on a workstation that captures the phone line and allows the user to quickly review audio content as well as to mark important snippets in the audio loop to save permanently.

Although the Xcapture application buffers phone conversations, the user must interact with the captured audio at a workstation, away from the phone. The intended interaction seemingly is for the user to review important portions of the audio after the conversation ends. Dietz and Yerazunis [42] later provided a similar capability using the phone device itself as the interface for interacting with the captured audio. MERL's real-time audio buffering technique does rely on a computer on the back end to record the phone conversation. However, Dietz and Yerazunis modified the earpiece to a phone itself to include a capacitive proximity sensor to determine when the phone is near the ear. A change in capacitance indicates the user's desire to relisten to the previous few seconds of the discussion. The user can repeatedly tap the phone to move successively further back in time. This application continues to record the other participant in the phone conversation while the user is reviewing the audio. The application allows the user to catch up to the live conversation by speeding up the playback using audio processing techniques.

As the mobile phone continues to increase in processing power, the phone itself can be used to record audio conversations in the place of a workstation on the back end. Using the Motorola i730 phone, Hayes et al. explored the potential usability, usefulness and acceptability issues involved with the user always being able to review audio from her recent past [4, 71, 143]. The Personal Audio Loop application continuously records audio from the phone's microphone and stores the audio for a pre-defined amount of time. The user presses any button on the side of the phone to begin review of the recorded audio. The user can continue to press that button to jump further backwards. The user can press the other button on the side of the phone to jump forward in the audio stream. In a later study, the researchers examined people's reactions to the Personal Audio Loop if they encounter it during real situations from their daily life [84]. They learned that people want to be informed about recording but were not likely to ask that the recorded audio be deleted.

Video content also can be buffered, as demonstrated by the Where-Were-We application [126]. However, instead of acting as a short-term memory aid by preserving recent content that users can quickly review, the StartleCam application buffers video to capture interesting images of a user's surrounding after the system has detected a change in the user's emotional state [76]. The user wears a galvanic skin response reader that continuously monitors the user's affective mood; when the application detects a deviation, it triggers a camera worn by the user to capture images that caused the response. To compensate for latency caused by the sensor as well as the processing of the data points, StartleCam buffers a very short amount of video content to allow the application to grab images from seconds ago, when the trigger point occurred. Instead of buffering continuous video, the What-Was-I-Cooking application (also known as Cook's Collage) explores the use of collage displays to show recent activities in the kitchen [173, 174]. Tran and Mynatt envisioned this application would be useful to a parent with young children in a busy household.

Buffering means that recorded data only persist for a specified period of time, reducing storage issues. However, it does not allow users to review important content at a much later time — after the buffer has expired. As a result, Hayes et al. proposed extending buffering capabilities into a larger framework of *selective archiving*, in which services are always on and available for recording but require some explicit action to archive data. If no such action is taken, recorded data is deleted automatically after a specified time. A prototype instance of selective archiving, known as the Experience Buffers Architecture was then developed and evaluated in relation to three domains: behavior assessment for children with disabilities in schools Hayes et al. [69], impromptu meetings and social gatherings in a semi-public research space [73], and recording of developmental milestones in homes for newborn and young children [92]. These applications of the architecture focused primarily on audio and video recording, but the overall architecture is flexible enough to handle other sensor inputs, such as from ink and environmental sensors, all in individual modules connected by simple networking protocols [74].

5.4 End-User Specified Capture Behaviors

As the trend towards technology-enriched environments progresses, the need to enable users to create applications to suit their own lives increases. Formative user studies indicate that user descriptions of a service tend to focus on the function and not the devices [41, 175]. Truong et al. [178] performed a preliminary design and evaluation of a GUI interface, known as CAMP (Capture and Access Magnetic Poetry) that allows users to map their conceptual models onto the INCA architectural model [175]. The CAMP interface offers users a flexible way to specify desired applications through the use of a "magnetic poetry" metaphor. Users can combine home or capture-themed magnetic poetry pieces into statements that describe an application. Their early evaluation indicates that the magnetic poetry interface is simple to learn and to use, and allows users to specify the types of applications they want in the way that makes the sense to them. Other effective methods allowing end-user specification of ubiquitous computing applications may remain, and need to be developed and studied.

5.5 Accessing Captured Data

The metadata of time and date are simple to add to any captured information at the point of capture without any additional burden to a user. As a result, time-based access interfaces are often utilized in capture and access applications as the simplest to creat programmatically. Using time, any captured information can be replayed back to the user. As shown through a myriad of studies across multiple applications [20, 129, 150, 168], user playback behavior often involves skimming and non-linear jumps within data streams. Features such as the ability to index into any portion of the captured content and faster than real-time playback are useful ones to include.

To further facilitate the access of information, many systems also capture and tag information with additional context, such as user location and others present or nearby. With the additional context streams, an access interface can support more sophisticated ways to search and index the content. For example, the Family Video Archive allows the user to query for video that was captured in specific locations with different people present [142]. With the Remembrance Agent, Rhodes and Starner [149] demonstrated how an interface which automatically searches and indexes previously captured content can augment a person's memory.

Although playback of previously captured content supports remembering and recall of those experiences, it is a time-consuming process. Furthermore, as recording has become easier over the years, the amount of captured data has increased significantly. To facilitate retrieval over the large amount of content, interfaces can also provide visualization and automatic summarization of this collection of data as a way to help users locate high-level points of interest and then allows the user to drill down to specific points in the stream for playback. Using the meta-data often gathered through additional context sensing during the capture of the live experience, an interface can shown information such as when and where a stream was recorded and who was present. Further analysis and clustering of the captured streams can be performed to produce summaries of the recorded experiences [7, 86]. Other effective methods for summarizing and visualizing captured content may remain, and need to be developed and studied.

5.6 Support for Developing Capture and Access Applications

While Moran et al. explored the benefit of long-term capture and access with the Tivoli system, researchers at Xerox PARC also explored the potential benefits of near-term reminder systems. Minneman and Harrison [126] developed the Where-Were-We application as a service that captures video of the meeting activity. Although the content can be reviewed at a much later time, this service also allows users to index into the captured streams when they need to be reminded of certain pieces of information *during* the activity. Later, Tivoli was combined with the Where-Were-We application to form Coral, an infrastructure that explores coordinating a defined confederation of tools to capture collaborative activities for later access [127]. This integration was performed using the Inter-Language Unification (ILU) project [87], which provides a distributed-object programming facility. This mechanism allowed developers to create objects that essentially acted as proxies, connecting heterogeneous components together in the larger system.

The Infrastructure for Capture and Access (InCA) toolkit supports novel exploration of the capture and access design space by providing programmatic abstractions for the capture, storage, transduction, and access of multiple streams of data. These abstractions alleviate programmers from addressing several concerns:

- A network abstraction eliminates the need to locate distributed modules;
- All captured data are handled generically, eliminating the need to consider the specificities of types of data;
- Watchdog threads ensure reliability and self-maintenance, eliminating the need to develop code specifically related to checking network connections, and so on; and
- Built-in network filtering reduces the need to adjust bandwidth within the application itself based on the constraints of the environment [175].

Similar architectural support used to develop capture and access applications have also been investigated by others. For example, Kim et al. [97] developed a set of Personal Chronicling Tools (PCTs) which consists of four components for event monitoring, interactively annotating captured events, searching/browsing of information, and editing/publishing captured content.

A key to making capture systems extensible is flexible storage scheme. Obviously, to synchronize different captured streams, a repository must be time-based [56]. However, other ways of linking the information must be supported as well. Kim et al. [96] developed an event-centric storage and linkage model, which takes into consideration spatial, temporal, and semantic variations of the captured information. Alternatively, Kiss and Quinqueton [98] proposed that a model for persistent and incremental knowledge storage should consist of two parts: a knowledge layer and a resource layer. The knowledge layer would contain annotations about the resource base.

6

Summary and Open Challenges

Over the years, many have been inspired by the visions of those such as Vannevar Bush and Mark Weiser. In this paper, we provided an extensive review of this body of work. Systems in this application space can be characterized based on the following design dimensions:

1. *Length of time captured information persists*: Some applications, such as Xcapture [78] and the Personal Audio Loop [4, 71, 143] keep captured content available for only a short amount of time. These applications act as quick short-term reminders for the user. Other applications, such as eClass [1, 20] and the Cornell Lecture Browser [132], store information indefinitely, allowing the user to review the content at a much later time after the live experience. These applications allow the user to recall rich details from their past.
2. *When and where capture occurs*: Some applications capture experiences that occur within fixed spaces that occur at regularly scheduled times. For example, eClass and NoteLook [26] support the capture of lectures and meetings that typically occur in a classroom and meeting room at fixed times,

respectively. A few applications have begun to explore the capture of a user's experience in a larger space than a single room [40, 54] and her experience continuously across multiple environments [185].

3. *Number of devices comprising the application*: Applications that capture experiences within one fixed environment can afford to distribute the responsibility for capturing different streams of information across multiple machines embedded in that space. For these scenarios, the user typically would use yet another device to review the captured experience. However, applications that follow the user around as a personal service typically run on a single device. For example, the Personal Audio Loop [4, 71, 143] and the Audio Notebook [167] are mobile self-contained devices that support the capture, storage, as well as the later review of previous experiences.
4. *Methods for capturing and annotating the live experience*: Capture can be performed passively using cameras and microphones (as done in the Lecture Browser system) or actively by augmenting the tools that the user interacts with inside an environment (as done in the eClass system). Furthermore, active capture is sometimes enacted automatically, triggered by sensors or computational logic, and sometime manually, such as by a human pushing a button. Capture, however, is only the first challenge to preserving content in a way that allows users to review information at a later time. Additionally, applications must annotate the captured information in a meaningful manner to facilitate its recall. At the least, applications tag captured content with meta-data in order to distinguish one captured session from another and allow the various streams of information within that captured session to integrate. Many projects have demonstrated how time and/or location are minimally sufficient for this purpose. The Family Video Archive project [3] demonstrates a compelling reason for manually annotating captured data with a richer set of meta-data; such annotations can also

be performed automatically by the application using sensors and learning techniques [142].

5. *Techniques for reviewing captured information*: A person must use what she can remember about an experience to retrieve specific data from the collection of captured content. The Family Video Archive demonstrates one way an application can support the dynamic searching and grouping of relevant information from a large amount of captured data. In simpler scenarios, the user obtains all the information from a particular captured session for her perusal (meaning the application automatically performs the grouping of captured information at the session level). As shown with the eClass application, one stream of information (ink annotations) can act as an index into the other richer media stream (audio or video), which the user can play back for additional detail [1]. Beyond searching, indexing, and playback, the Manga application demonstrates a process for automatically summarizing captured video as a way to help users locate high-level points of interest and then allows the user to drill down to specific points in the stream for playback [180].

Despite the multitude of works in this area, there are still open challenges in capture and access. These include both computational-centric and human-centric concerns, and we overview some of these challenges here.

The value of automated capture applications is most apparent when users review the content at a later time. To avoid the problems of information overload and the general inability to find either specific information or general trends within the data, new, sophisticated access interfaces must be developed. To understand generally immense quantities of data, automatic summarization of dynamically captured content becomes increasingly important. As previously described, Whittaker et al. [189] proposed poor summarization as one of the primary drivers for limited use of captured information from meetings and developed a system for provision of the “gist” of meetings. Jaffe et al. [86] proposed a technique for creating a summarization of a large set of geo-referenced photographs. As another example, Lee and Dey have begun to explore the notion of "smart summaries." One of their exemplar applications is to provide memory cues into massive amount of recorded data for patients with cognitive disabilities and memory impairments [109].

Although the popularity of search engines has resulted in a general population better trained to search over browsing, there are some activities that still require browsing capabilities. Thus, improved browsing capabilities can support users in finding content of interest. For example, the Space–Time Browser enables visual representation of content using two naturally salient clues: time and location [27]. Several other research projects have begun to engage the immense and complex space of labeling and visualizing captured data. For example, with the enormity of digital photograph collections, new interest has emerged around how people organize their digital photographs [90, 153, 160]. Individual and group collections of multimedia are only continuing to grow, however, necessitating further research in these areas.

The majority of capture and access applications envisioned involve some degree of network connectivity, between various sensors and recording devices and data storage and access interfaces. Some applications have taken a "store and forward" model in which some storage is local, and when network connectivity is at its peak, the data are transferred off these local devices to a central storage space. Others have taken a more direct connection approach, in which data are streamed near real-time to storage elsewhere. Regardless of the capture approach taken, the mobility of the human users will continue to necessitate that both the recording and the review of data be allowed to occur in a multitude of locations. Thus, a significant open challenge remains in providing these services in locations in which power, network connectivity, and other infrastructure may be spotty.

As powerful recording technologies become more ubiquitous, the necessity for enabling end-users to control these technologies becomes more imperative. Already, people interact with a myriad of everyday recording technologies in public that they would like to but are unable to control [138]. When considering the use of these technologies in homes and other sensitive spaces, end-users may become even more interested in controlling what and how data are captured. Enabling end-users to appropriate these technologies for their own users is in many ways a substantial challenge for HCI researchers but also one of end-user programming. End-users tend to focus on the function, rather than the devices of recording as a developer might. Thus, new interfaces can allow users to map their more natural mental models of capture onto the technological infrastructure. One preliminary design for such an interface is the Capture and Access Magnetic Poetry (CAMP) application [175]. As described in Section 5, CAMP offers users a flexible way to specify desired applications through the user of a "magnetic poetry" metaphor. End-user specification remains a complicated and important area of research for capture and access applications, in particular as surveillance and recording technologies and the capabilities and desire of the average end-user to document important experiences in everyday life become more common place.

Following along these lines, understanding of and support for socially appropriate recording are significant challenges moving forward. Manual recording, even that augmented by computational support, take the user out of the moment and are less likely to garner valuable information. Recording automatically, in the background, can address this concern by removing the load of recording from the end-user. However, this model also removes the ability to control and often to be aware o the recording taking place. Although users tend to object to substantial recording, particularly of images, when simple sensors would be sufficient for the task at hand [11], people are often willing to accept intrusive technologies that offer useful services [73, 125, 138]. This trade-off falls at least partly in line with Hong et al.'s [81] notion of risk assessment and Iachello and Abowd's [83] concept of proportionality. One solution for balancing the concerns of removing the load of recording while keeping users in control is the idea of selective archiving, in which recording devices are always on and available in an environment, but without explicit user action, data are automatically deleted [74]. This solution worked well to handle the concerns in a particularly sensitive environment — classrooms for children with special needs Hayes et al. [69], but was challenged in more nuanced ways in an open office area [73]. In the latter work, people reacted to the technology based on a variety of social, physical, and experiential cues to the purpose, safety, and capabilities of the capture and access application. Research to support the social, technical, and practical concerns of capture applications warrants further investigation.

Finally, with the ubiquity of small mobile recording devices — such as camera phones or new tiny spy cameras and audio recorders — and the multitude of aggregating technologies — such as sharing sites like flickr® and mining technologies like face recognition — new concerns have arisen about the myriad of "little brothers" who may be recording. Although a variety of human legal and social processes can thwart much-unwanted recording, there are still cases in which it may not be practical or desirable to confiscate or regulate recording through these means. Thus, developing new techniques for blocking recording without the cooperation or sometimes even knowledge of the entity doing the recording remains an open area for research in this domain. As one example, the Capture Resistant Environment prevented CCD and CMOS cameras from recording still or moving images by identifying and shining light into the lens of these cameras [179]. This approach, although interesting and promising, only works to prevent consumer-level CCS and CMOS cameras, thus leaving open a wide space for development of new technological innovations.

In conclusion, the history of capturing and accessing records of human experience is long and filled with varied motivations, technological innovations, and social and political phenomena. As recording technologies become more powerful and ubiquitous, the research designs of the past have become everyday realities even while new research areas have been opened. This article serves as both a review of what has already been accomplished and a jumping off point for the capture and access research of the future.

References

[1] G. D. Abowd, "Classroom 2000: An experiment with the instrumentation of a living educational environment," *IBM Systems Journal*, vol. 38, no. 4, pp. 508–530, 1999.

[2] G. D. Abowd, "Using computing technologies to face the challenges of autism," in *Social Impact Award Talk at CHI 2007: The ACM Conference on Human Factors in Computing Systems*, San Jose, CA: San Jose McEnery Convention Center, May 3, 2007.

[3] G. D. Abowd, M. Gauger, and A. Lachenmann, "The family video archive: An annotation and browsing environment for home movies," in *Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval*, pp. 1–8, 2003.

[4] G. D. Abowd, G. R. Hayes, G. Iachello, J. A. Kientz, S. N. Patel, M. Stevens, and K. N. Truong, "Prototypes and paratypes: Designing mobile and ubiquitous computing applications," *IEEE Pervasive Computing*, vol. 4, no. 4, pp. 67–73, 2005.

[5] G. D. Abowd and E. D. Mynatt, "Charting past, present, and future research in ubiquitous computing," *ACM Transactions on Computer–Human Interaction*, vol. 7, no. 1, pp. 29–58, 2000.

[6] E. Adar, D. Karger, and L. A. Stein, "Haystack: Per-user information environments," in *Proceedings of CIKM 1999: The 8th International Conference on Information and Knowledge Management*, pp. 413–422, 1999.

[7] K. Aizawa, D. Tancharoen, S. Kawasaki, and T. Yamasaki, "Efficient retrieval of life log based on content and context," in *Proceedings of CARPE 2004: The First ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, pp. 22–31, 2004.

[8] C. Bacher, R. Müller, T. Ottmann, and M. Will, "Authoring on the fly: A new way of integrating telepresentation and courseware production," in *Proceedings of the 5th International Conference on Computers in Education*, 1997.

[9] D. Barreau, A. Crystal, J. Greenberg, A. Sharma, M. Conway, J. Oberlin, M. Shoffner, and S. Seiberling, "Augmenting memory for student learning: Designing a context-aware capture system for biology education," in *Proceedings of the American Society for Information Science and Technology*, vol. 43, no. 1, pp. 251–251, 2007.

[10] D. Barreau and B. A. Nardi, "Finding and reminding: File organization from the desktop," *ACM SIGCHI Bulletin 27*, pp. 39–43, 1995.

[11] C. Beckmann, S. Consolvo, and A. LaMarca, "Some assembly required: Supporting end-user sensor installation in domestic ubiquitous computing environments," in *Proceedings of Ubicomp 2004: The 6th International Conference on Ubiquitous Computing*, pp. 107–124, 2004.

[12] B. B. Bederson, "Interfaces for staying in the flow," *Ubiquity*, vol. 5, no. 27, p. 1, 2004.

[13] G. Bell, "A personal digital store," *Communications of the ACM*, vol. 44, no. 1, pp. 86–91, 2001.

[14] M. Bernstein, M. Van Kleek, D. Karger, and M. C. Schraefel, "Information scraps: How and why information eludes our personal information management tools," *ACM Transactions on Information Systems*, vol. 26, no. 4, pp. 1–46, 2008.

[15] M. Bernstein, M. Van Kleek, D. R. Karger, and M. C. Schraefel, "Management of personal information scraps," in *Extended Abstracts of CHI 2007:. The ACM Conference on Human Factors in Computing Systems*, pp. 2285–2290, 2007.

[16] D. Berque, A. Hutcheson, D. Johnson, L. Jovanovic, K. Moore, C. Singer, and K. Slattery, "Using a variation of the WYSIWIS shared drawing surface paradigm to support electronic classrooms," in *Proceedings of HCI'99: The 8th International Conference on Human Computer Interaction*, 1999.

[17] M. Bianchi, "AutoAuditorium: A fully automatic, multi-camera system to televise auditorium presentations," Paper Presented at the Joint DARPA/NIST Smart Spaces Workshop, Gaithersburg, Maryland, July 30–31, 1998.

[18] M. Black, F. Berard, A. Jepson, W. Newman, E. Saund, G. Socher, and M. Taylor, "Digital office: Overview," in *Proceedings of the AAAI Spring Symposium on Intelligent Environments*, pp. 1–6, 1998.

[19] J. A. Brotherton, "Enriching everyday activities through the automated capture and access of live experiences — eClass: Building, observing and understanding the impact of capture and access in an educational domain," Ph.D. Dissertation: 270 pages, Georgia Institute of Technology, 2001.

[20] J. A. Brotherton and G. D. Abowd, "Lessons learned from eClass: Assessing automated capture and access in the classroom," *ACM Transactions on Computer–Human Interaction*, vol. 11, no. 2, pp. 121–155, 2004.

[21] J. A. Brotherton, G. D. Abowd, and K. N. Truong, "Supporting capture and access interfaces for informal and opportunistic meetings," Technical Report GIT-GVU-99-06, Georgia Institute of Technology, 1999.

[22] J. A. Brotherton, J. R. Bhalodia, and G. D. Abowd, "Automated capture, integration, and visualization of multiple media streams," in *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, pp. 54–63, 1998.

[23] V. Bush, "As we may think," *The Atlantic Monthly*, vol. 176, no. 1, pp. 101–108, 1945.

[24] W. Buxton and T. P. Moran, "EuroPARC's integrated interactive intermedia facility (IIIF): Early experiences," in *Proceedings of the IFIP WG8.4 Conference on Multi-User Interfaces and Applications*, pp. 11–34, 1990.

[25] C. Campbell and P. Maglio, "Supporting notable information in office work," in *Extended Abstracts of CHI 2003: The ACM Conference on Human Factors in Computing Systems*, pp. 902–903, 2003.

[26] P. Chiu, A. Kapuskar, S. Reitmeier, and L. Wilcox, "NoteLook: Taking notes in meetings with digital video and ink," in *Proceedings of the 7th ACM International Conference on Multimedia*, pp. 149–158, 1999.

[27] P. Chiu and K. N. Truong, "Interactive space–time planes for document visualization," in *Compendium of IEEE Symposium on Information Visualization*, pp. 10–11, 2002.

[28] A. Civan, M. M. Skeels, A. Stolyar, and W. Pratt, "Personal health information management: Consumers' perspectives," in *Proceedings of AMIA 2006 Annual Symposium*, pp. 156–160, 2006.

[29] B. P. Clarkson, "Life patterns: Structure from wearable sensors," Ph.D. Dissertation: 130 pages, Massachusetts Institute of Technology, 2002.

[30] S. Consolvo, P. V. Klasnja, D. W. McDonald, D. Avrahami, J. Froehlich, L. LeGrand, R. Libby, K. Mosher, and J. A. Landay, "Flowers or a robot army?: Encouraging awareness and activity with personal, mobile displays," in *Proceedings of Ubicomp 2008: The 10th International Conference on Ubiquitous Computing*, pp. 54–63, 2008.

[31] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, and I. S. and J. A. Landay, "Activity sensing in the wild: A field trial of ubifit garden," in *Proceeding of CHI 2008: The ACM Conference on Human Factors in Computing Systems*, pp. 1797–1806, 2008.

[32] S. Consolvo, P. Roessler, B. E. Shelton, A. LaMarca, B. Schilit, and S. Bly, "Technology for care networks of elders," *IEEE Pervasive Computing*, vol. 3, no. 2, pp. 22–29, 2004.

[33] P. Cowley, J. Haack, R. Littlefield, and E. Hampson, "Glass box: Capturing, archiving, and retrieving workstation activities," in *Proceedings of CARPE 2006: The 3rd ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, pp. 13–18, 2006.

[34] G. Cruz and R. Hill, "Capturing and playing multimedia events with STREAMS," in *Proceedings of the Second ACM International Conference on Multimedia*, pp. 193–200, 1994.

[35] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. HarperCollins, 1991.

[36] E. Cutrell, S. T. Dumais, and J. Teevan, "Searching to reduce the need for personal information management," *Communications of the ACM*

(*Special Issue: Personal Information Management*), vol. 49, no. 1, pp. 58–64, 2006.

[37] M. Czerwinski, D. W. Gage, J. Gemmell, C. C. Marshall, M. A. Pérez-Quiñones, M. M. Skeels, and T. Catarci, "Digital memories in an era of ubiquitous computing and abundant storage," *Communications of the ACM*, vol. 49, no. 1, pp. 44–50, 2006.

[38] R. C. Davis, J. A. Landay, V. Chen, J. Huang, R. B. Lee, F. C. Li, J. Lin, C. B. Morrey III, B. Schleimer, M. N. Price, and B. N. Schilit, "NotePals: Lightweight note sharing by the group, For the group," in *Proceedings of CHI 1999: The ACM Conference on Human Factors in Computing Systems*, pp. 338–345, 1999.

[39] G. C. de Silva, T. Yamasaki, and K. Aizawa, "Experience retrieval in a ubiquitous home," in *Proccesings of CARPE 2005: The ACM Multimedia Workshop on Continuous Archival of Personal Experience*, pp. 35–44, 2005.

[40] A. K. Dey, D. Salber, G. D. Abowd, and M. Futakawa, "The conference assistant: Combining context-awareness with wearable computing," in *Proceedings of the 3rd IEEE International Symposium on Wearable Computers*, pp. 21–28, 1999.

[41] A. K. Dey, T. Sohn, S. Streng, and J. Kodama, "iCAP: Interactive prototyping of context-aware applications," in *Proceedings of Pervasive 2005: The International Conference on Pervasive Computing*, pp. 254–271, 2006.

[42] P. H. Dietz and W. S. Yerazunis, "Real-time audio buffering for telephone applications," in *Proceedings of UIST 2001: The ACM Symposium on User Interface Software and Technology*, pp. 193–194, 2001.

[43] P. Dourish, W. K. Edwards, A. LaMarca, and M. Salisbury, "Presto: An experimental architecture for fluid interactive document spaces," *ACM Transactions on Computer–Human Interaction*, vol. 6, no. 2, pp. 133–161, 1996.

[44] S. Dumais, E. Cutrell, J. Cadiz, G. Jancke, R. Sarin, and D. C. Robbins, "Stuff I've seen: A system for personal information retrieval and re-use," in *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 72–79, 2003.

[45] P. Ehlen, R. Fernandez, and M. Frampton, "Designing and evaluating meeting assistants, keeping humans in mind," in *Proceedings of the 5th International Workshop on Machine Learning for Multimodal Interaction*, pp. 309–314, 2008.

[46] P. Ehlen, M. Purver, and J. Niekrasz, "A meeting browser that learns," in *Interaction Challenges for Intelligent Assistants: Papers from the 2007 AAAI Spring Symposium*, Stanford, CA, March 2007.

[47] J. H. Eichhorn, "Anesthesia record keeping," *Journal of Clinical Monitoring and Computing*, vol. 10, no. 2, pp. 109–115, 1993.

[48] M. Eldridge, M. Lamming, and M. Flynn "Does a video diary help recall?" In *Proceedings of HCI'92: The Conference on People and Computers VII*, pp. 257–269, 1992.

[49] D. P. W. Ellis and K. Lee, "Minimal-impact audio-based personal archives," in *Proceedings of CARPE 2004: The 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, pp. 39–47, 2004.

[50] S. Elrod, R. Bruce, R. Gold, D. Goldberg, F. Halasz, W. Janssen, D. Lee, K. McCall, E. Pedersen, K. Pier, J. Tang, and B. Welch, "Liveboard: A large interactive display supporting group meetings, presentations, and remote collaboration," in *Proceedings of CHI 1992: The ACM Conference on Human Factors in Computing Systems*, pp. 599–607, 1992.

[51] D. C. Engelbart, "Augmenting human intellect: A conceptual framework," Summary Report, AFOSR-3223 — Contract AF49(638)-1024, SRI Project 3578: Air Force Office of Scientific Research, 1962.

[52] European Union Directive 95/46/EC. http://ec.europa.eu/justice_home/fsj/privacy/index_en.htm.

[53] A. Fernandez, "Record-keeping and other troublemaking: Thomas Lechford and law reform in colonial Massachusetts," *Law and History Review*, vol. 23, no. 2, 2005.

[54] M. Fleck, M. Frid, T. Kindberg, E. O'Brien-Strain, R. Rajani, and M. Spasojevic, "Rememberer: A tool for capturing museum visits," in *Proceedings of Ubicomp 2002: The 4th International Conference on Ubiquitous Computing*, pp. 48–55, 2002.

[55] R. Fleck and G. Fitzpatrick, "Supporting reflection with passive image capture," in *Supplementary Proceedings of COOP'06*, pp. 41–48, 2006.

[56] E. T. Freeman, "The lifestreams software architecture," Ph.D. Dissertation: 185 pages, Yale University, 1997.

[57] R. Gellman, "Personal health records: Why many PHRs threaten privacy," A report prepared for the *World Privacy Forum*, February 20, 2008.

[58] J. Gemmell, G. Bell, R. Lueder, S. Drucker, and C. Wong, "MyLifeBits: Fulfilling the memex vision," in *Proceedings of the 10th ACM International Conference on Multimedia*, pp. 235–238, 2002.

[59] J. Gemmell, L. Williams, K. Wood, R. Lueder, and B. Gordon, "Passive capture and ensuing issues for a personal lifetime store," in *Proceedings of the CARPE 2004: The 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, pp. 48–55, 2004.

[60] J. Glass, T. J. Hazen, S. Cyphers, I. Malioutov, D. Huynh, and R. Barzilay, "Recent progress in the MIT spoken lecture processing project," in *Proceedings of Interspeech*, 2007.

[61] A. A. Glatthorn, "Outcome based education: Reform and the curriculum process," *Journal of Curriculum and Supervision*, vol. 8, no. 4, pp. 354–363, 1993.

[62] M. L. Gleicher, R. M. Heck, and M. N. Wallick, "A framework for virtual videography," in *Proceedings of the 2nd International Symposium on Smart Graphics*, pp. 9–16, 2002.

[63] C. Goodwin, "Professional vision," in *American Anthropologist*, vol. 96, no. 3, pp. 606–633, 1994.

[64] C. Goodwin, "Transparent vision," in *Interaction and Grammar*, (E. Ochs, E. A. Schegloff, and S. Thompson, eds.), pp. 370–404, 1996.

[65] G. D. Haertel, *A Primer on Teacher Self-Evaluation*. Publication Series No. 93-3, 1993.

[66] M. Hâkansson, S. Ljungblad, and L. E. Holmquist, "Capturing the invisible: Designing context-aware photography," in *Proceedings of DIS 2003: The ACM Conference on Designing for User Experiences*, pp. 1–4, 2003.

[67] A. Hargreaves and A. Moore, "Educational outcomes, modern and postmodern interpretations: Response to Smyth and Dow," *British Journal of Sociology of Education*, vol. 21, no. 1, pp. 27–42, 2000.

[68] G. R. Hayes and G. D. Abowd, "Tensions in designing capture technologies for an evidence-based care community," in *Proceedings of CHI 2006: The ACM Conference on Human Factors in Computing Systems*, pp. 937–946, 2006.

[69] G. R. Hayes, L. M. Gardere, G. D. Abowd, and K. N. Truong, "CareLog: A selective archiving tool for behavior management in schools," in *Proceedings of CHI 2008: The ACM Conference on Human Factors in Computing Systems*, pp. 685–694, 2008.

[70] G. R. Hayes, J. A. Kientz, K. N. Truong, D. R. White, G. D. Abowd, and T. Pering, "Designing capture applications to support the education of children with autism," in *Proceedings of Ubicomp 2004: The 6th International Conference on Ubiquitous Computing*, pp. 161–178, 2004a.

[71] G. R. Hayes, S. N. Patel, K. N. Truong, G. Iachello, J. A. Kientz, R. Farmer, and G. D. Abowd, "The personal audio loop: Designing a ubiquitous audio-based memory aid," in *Proceedings of the 6th International Symposium on Mobile Human–Computer Interaction*, pp. 168–179, 2004b.

[72] G. R. Hayes, J. S. Pierce, and G. D. Abowd, "Practices for capturing short important thoughts," in *Extended Abstracts of CHI 2003: The ACM Conference on Human Factors in Computing Systems*, pp. 904–905, 2003.

[73] G. R. Hayes, E. S. Poole, G. Iachello, S. N. Patel, A. Grimes, G. D. Abowd, and K. N. Truong, "Physical, social, and experiential knowledge in pervasive computing environments," *IEEE Pervasive Computing*, vol. 6, no. 4, pp. 56–63, 2007.

[74] G. R. Hayes, K. N. Truong, G. D. Abowd, and T. Pering, "Experience buffers: A socially appropriate, selective archiving tool for evidence-based care," in *Extended Abstracts of CHI 2005: The ACM Conference on Human Factors and Computing Systems*, pp. 1435–1438, 2005.

[75] T. L. Hayes, M. Pavel, P. K. Schallau, and A. M. Adami, "Unobtrusive monitoring of health status in an aging population," in *UbiHealth 2003: The Second International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications*, Seattle, Washington, 2003.

[76] J. Healey and R. W. Picard, "StartleCam: A cybernetic wearable camera," in *Proceedings of the 2nd IEEE International Symposium on Wearable Computers*, pp. 42–49, 1998.

[77] L. J. Heflin and R. L. Simpson, "Interventions for children and youth with autism: Prudent choices in a world of exaggerated claims and empty promises," *Intervention and Treatment Option Review. Focus on Autism and Other Developmental Disabilities*, vol. 13, no. 4, pp. 194–211, 1998.

[78] D. Hindus and C. Schmandt, "Ubiquitous audio: Capturing spontaneous collaboration," in *Proceedings of CSCW 1992: The ACM Conference on Computer-Supported Cooperative Work*, pp. 210–217, 1992.

[79] HIPAA: Health Insurance Portability and Accountability Act of 1996 (HIPAA, Title II), 1996.

[80] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood, "SenseCam: A retrospective memory aid," in *Proceedings of Ubicomp 2006: The International Conference on Ubiquitous Computing*, pp. 177–193, 2006.

[81] J. I. Hong, J. D. Ng, S. Lederer, and J. A. Landay, "Privacy risk models for designing privacy-sensitive ubiquitous computing systems," in *Proceedings of DIS 2004: The ACM Conference on Designing Interactive Systems*, pp. 91–100, 2004.

[82] J. Huang, "A collaborative property-based note management system," Master's Dissertation: 48 pages, University of California at Berkeley, 2000.

[83] G. Iachello and G. D. Abowd, "Privacy and proportionality: Adapting legal evaluation techniques to inform design in ubiquitous computing," in *Proceedings of CHI 2005: The ACM Conference on Human Factors in Computing Systems*, pp. 91–100, 2005.

[84] G. Iachello, K. N. Truong, G. D. Abowd, G. R. Hayes, and M. Stevens, "Prototyping and sampling experience to evaluate ubiquitous computing privacy in the real world," in *Proceedings of CHI'06: The ACM Conference on Human Factors in Computing Systems*, pp. 1009–1018, 2006.

[85] W. Ives, "Using computer-controlled audio-video feedback to integrate evaluation and training," Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA, March, 1989.

[86] A. Jaffe, M. Naaman, T. Tassa, and M. Davis, "Generating summaries for large collections of geo-referenced photographs," in *Proceedings of WWW 2006: The 15th International Conference on World Wide Web*, pp. 853–854, 2006.

[87] B. Janssen ILU Manual. Xerox Technical Report ISTL-CSA-94-01-02, Xerox PARC, Palo Alto, CA, 1994.

[88] W. Ju, A. Ionescu, L. Neeley, and T. Winograd, "Where the wild things work: Capturing shared physical design workspaces," in *Proceedings of CSCW 2004: The ACM Conference on Computer Supported Cooperative Work*, pp. 533–541, 2004.

[89] V. Kalnikaité and S. Whittaker, "Software or wetware?: Discovering when and why people use digital prosthetic memory," in *Proceedings of CHI 2007: The ACM Conference on Human Factors in Computing Systems*, pp. 71–80, 2007.

[90] H. Kang and B. Shneiderman, "Visualization methods for personal photo collections: Browsing and searching in the photofinder," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 1539–1542, 2000.

[91] S. Karas, "Enhancing the privacy discourse: Consumer information gathering as surveillance," *Journal of Technology Law and Policy*, vol. 7, no. 1, 2002.

[92] J. A. Kientz, R. I. Arriaga, M. Chetty, G. R. Hayes, J. Richardson, S. N. Patel, and G. D. Abowd, "Grow and know: understanding record-keeping needs for tracking the development of young children," in *Proceedings of CHI 2007: The*

ACM Conference on Human Factors in Computing Systems, pp. 1351–1360, 2007a.

[93] J. A. Kientz, S. Boring, G. D. Abowd, and G. R. Hayes, "Abaris: Evaluating automated capture applied to structured autism interventions," in *Proceedings of Ubicomp 2005: The International Conference on Ubiquitous Computing*, pp. 323–339, 2005.

[94] J. A. Kientz, G. R. Hayes, G. D. Abowd, and R. E. Grinter, "From the war room to the living room: Decision support for home-based therapy teams," in *Proceedings of CSCW 2006: The ACM Conference on Computer Supported Cooperative Work*, pp. 209–218, 2006.

[95] J. A. Kientz, G. R. Hayes, T. L. Westeyn, T. Starner, and G. D. Abowd, "Pervasive computing and autism: Assisting caregivers of children with special needs," *IEEE Pervasive Computing*, vol. 6, no. 1, pp. 28–35, 2007b.

[96] P. Kim, U. Gargi, and R. Jain, "Event-based multimedia chronicling systems," in *Proceedings of CARPE 2005: The 2nd ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, pp. 1–12, 2005.

[97] P. Kim, M. Podlaseck, and G. Pingali, "Personal chronicling tools for enhancing information archival and collaboration in enterprises," in *Proceedings of CARPE 2004: The 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, pp. 56–65, 2004.

[98] A. Kiss and J. Quinqeton, "Uniscript: A model for persistent and incremental knowledge storage," in *Proceedings of CARPE 2004: The 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, pp. 66–73, 2004.

[99] S. R. Klemmer, M. W. Newman, R. Farrell, M. Bilezikjian, and J. A. Landay, "The designers' outpost: A tangible interface for collaborative web site," in *Proceedings of UIST 2001: The ACM Symposium on User Interface Software and Technology*, pp. 1–10, 2001.

[100] S. R. Klemmer, M. Thomsen, E. Phelps-Goodman, R. Lee, and J. A. Landay, "Where do web sites come from? Capturing and interacting with design history," in *Proceedings of CHI 2002: The ACM Conference on Human Factors in Computing Systems*, pp. 1–8, 2002.

[101] J. Korhonen, T. Ojala, A. Ristola, M. Kesti, V. Kilpelänaho, M. Koskinen, and E. Viippola, "Mobile fair diary: Hybrid interface for taking, browsing and sharing context-aware notes," *Personal and Ubiquitous Computing*, vol. 11, no. 7, pp. 577–589, 2007.

[102] M. Lamming, P. Brown, K. Carter, M. Eldridge, M. Flynn, G. Louie, P. Robinson, and A. Sellen, "The design of a human memory prosthesis," *The Computer Journal*, vol. 37, no. 3, pp. 153–163, 1994.

[103] M. Lamming and M. Flynn, "Forget-Me-Not: Intimate computing in support of human memory," in *Proceedings of Friend 21: International Symposium on Next Generation Human Interface*, pp. 125–128, 1994.

[104] M. G. Lamming, "NoTime — A tool for notetakers," *EuroPARC Technical Report*, 1991.

[105] M. Lampert 'Why use interactive media in teacher preparation?' *Innovator* vol. 25, no. 1, 1994.

[106] M. Lampert, R. Heaton, and D. Ball, "Using technology to support a new pedagogy of mathematics education," *Journal of Special Education Technology* (*Special Issue: Technology and teaching*), vol. 12, no. 3, 1994.

[107] J. A. Landay and R. C. Davis, "Making sharing pervasive: Ubiquitous computing for shared note taking," *IBM Systems Journal*, vol. 38, no. 4, pp. 531–550, 1999.

[108] H. Lee, A. F. Smeaton, N. E. O'Connor, and G. J. F. Jones, "Adaptive visual summary of lifelog photos for personal information management," in *AIR 2006 — First International Workshop on Adaptive Information Retrieval*, Glasgow, UK, October 14, 2006.

[109] M. L. Lee and A. K. Dey, "Providing good memory cues for people with episodic memory impairment," in *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 131–138, 2007.

[110] M. L. Lee and A. K. Dey, "Lifelogging memory appliance for people with episodic memory impairment," in *Proceedings of UbiComp 2008: The International Conference on Ubiquitous Computing*, pp. 44–53, 2008.

[111] G. Leinhardt, "Expertise in mathematics teaching," *Educational Leadership*, vol. 43, pp. 23–27, 1986.

[112] J. C. R. Licklider, "Man machine symbiosis," *IRE Transactions on Human Factors in Electronics*, vol. 1, pp. 4–11, 1960.

[113] J. Lin, L. Mamykina, S. Lindtner, G. Delajooux, and H. Strub, "Fish'n'Steps: Encouraging physical activity with an interactive computer game," in *Proceedings of Ubicomp 2006: The International Conference on Ubiquitous Computing*, pp. 261–278, 2006.

[114] M. Lin, M. W. Newman, J. I. Hong, and J. A. Landay, "DENIM: Finding a tighter fit between tools and practice for Web site design," in *Proceedings of CHI 2000: The ACM Conference on Human Factors in Computing Systems*, pp. 510–517, 2000.

[115] Q. Liu, Y. Rui, A. Gupta, and J. J. Cadiz, "Automating camera management for lecture room environments," in *Proceedings of CHI 2001: The ACM Conference on Human Factors in Computing Systems*, pp. 442–449, 2001.

[116] A. Lockerd and F. M. Mueller, "LAFCam: Leveraging affective feedback camcorder," in *Extended Abstracts of CHI 2002: The ACM Conference on Human Factors in Computing Systems*, pp. 574–575, 2002.

[117] B. MacIntyre, E. D. Mynatt, S. Voida, K. M. Hansen, J. Tullio, and G. M. Corso, "Support for multitasking and background awareness using interactive peripheral displays," in *Proceedings of UIST 2001: The ACM Symposium on User Interface Software and Technology*, pp. 41–50, 2001.

[118] W. E. Mackay, G. Pothier, C. Letondal, K. Bøegh, and H. E. Sørensen, "The missing link: Augmenting biology laboratory notebooks," in *Proceedings of UIST 2002: The ACM Symposium on User Interface Software and Technology*, pp. 41–50, 2002.

[119] L. Mamykina, S. Goose, D. Hedqvist, and D. V. Beard, "CareView: Analyzing nursing narratives for temporal trends," in *Extended Abstracts of CHI 2004: The ACM Conference on Human Factors in Computing Systems*, pp. 1147–1150.

[120] L. Mamykina, E. D. Mynatt, and D. R. Kaufman, "Investigating health management practices of individuals with diabetes," in *Proceedings of CHI 2006: The ACM Conference on Human Factors in Computing Systems*, pp. 927–936, 2006.

[121] R. W. Marx, P. C. Blumenfeld, J. S. Krajcik, and E. Soloway, "New technologies for teacher professional development," *Teaching and Teacher Education*, vol. 14, no. 1, pp. 33–52, 1998.

[122] D. R. McGee, P. R. Cohen, and L. Wu, "Something from nothing: Augmenting a paper-based work practice via multimodal interaction," in *Proceedings of DARE 2000 on Designing Augmented Reality Environments*, pp. 71–80, 2000.

[123] J. McKernan, "Some limitations of outcome-based education," *Journal of Curriculum and Supervision*, vol. 8, no. 4, pp. 343–53, 1993.

[124] G. McNeirr, "Outcomes based education: An overview *ERIC Difest* 85," Education Commission of the United States, November, 1993.

[125] A. S. Melenhorst, A. D. Fisk, E. D. Mynatt, and W. A. Rogers, "Potential intrusiveness of aware home technology: Perceptions of older adults," in *Proceedings of Human Factors and Ergonomics Society 48th Annual Meeting*, pp. 266–270, 2004.

[126] S. L. Minneman and S. R. Harrison, "Where were we: Making and using near-synchronous, pre-narrative video," in *Proceedings of the First ACM International Conference on Multimedia*, pp. 207–214, 1993.

[127] S. L. Minneman, S. R. Harrison, B. Janssen, G. Kurtenbach, T. P. Moran, I. E. Smith, and W. V. Melle, "A confederation of tools for capturing and accessing collaborative activity," in *Proceedings of the 3rd ACM International Conference on Multimedia*, pp. 523–534, 1995.

[128] T. P. Moran, W. V. Melle, and P. Chiu, "Spatial interpretation of domain objects integrated into a freeform electronic," in *Proceedings of UIST 1998: The ACM Symposium on User Interface Software and Technology*, pp. 175–184, 1998.

[129] T. P. Moran, L. Palen, S. Harrison, P. Chiu, D. Kimber, S. Minneman, W. V. Melle, and P. Zellweger, "I'll get that off the audio: A case study of salvaging multimedia meeting records," in *Proceedings of CHI 1997: The ACM Conference on Human Factors in Computing Systems*, pp. 202–209, 1997.

[130] M. Morris, S. S. Intille, and J. S. Beaudin, "Embedded assessment: Overcoming barriers to early detection with pervasive computing," in *Proceedings of Pervasive 2005: The International Conference on Pervasive Computing*, pp. 333–346, 2005.

[131] M. Morris, J. Lundell, E. Dishman, and B. Needham, "New perspectives on ubiquitous computing from ethnographic study of elders with cognitive decline," in *Proceedings of Ubicomp 2003: The International Conference on Ubiquitous Computing*, pp. 227–242, 2003.

[132] S. Mukhopadhyay and B. Smith, "Passive capture and structuring of lectures," in *Proceedings of the 7th ACM International Conference on Multimedia*, pp. 477–487, 1999.

[133] B. A. Myers, H. Stiel, and R. Gargiulo, "Collaboration using multiple PDAs connected to a PC," in *Proceedings of CSCW 1998: The ACM Conference on Computer Supported Cooperative Work*, pp. 285–294, 1998.

[134] E. D. Mynatt, T. Igarashi, W. K. Edwards, and A. LaMarca, "Flatland: New dimensions in office whiteboards," in *Proceedings of CHI 1999: The ACM Conference on Human Factors in Computing Systems*, pp. 346–353, 1999.

[135] National Evaluation Systems, "Advertisement," *Educational Researcher*, vol. 25, no. 8, p. 64, 1996.

[136] W. Newman and P. Wellner, "A desk supporting computer-based interaction with paper documents," in *Proceedings of CHI 1992: The ACM Conference on Human Factors in Computing Systems*, pp. 587–592, 1992.

[137] W. M. Newman, M. A. Eldridge, and M. G. Lamming, "PEPYS: Generating autobiographies by automatic tracking," in *Proceedings of the 2nd European Conference on Computer-Supported Cooperative Work*, pp. 175–188, 1991.

[138] D. H. Nguyen, G. R. Hayes, and A. Kobsa, "An empirical investigation of concerns of everyday tracking and recording technologies," in *Proceedings of Ubicomp 2008: The International Conference on Ubiquitous Computing*, pp. 182–191, 2008.

[139] E. Ochs, "Transcription as theory," in *Developmental pragmatics*, (E. Ochs and B. B. Schieffelin, eds.), pp. 43–72, Academic Press, 1979.

[140] C. O'Conaire, N. O'Connor, A. F. Smeaton, and G. Jones, "Organising a daily visual diary using multi-feature clustering," *SPIE Electronic Imaging — Multimedia Content Access: Algorithms and Systems* (*EI121*), San Jose, CA, 28 January–1 February, 2007.

[141] A. W. Pailliotet, "I never saw that before: A deeper view of video analysis in teacher education," *Teacher Educator*, vol. 31, no. 2, pp. 138–56, 1995.

[142] S. N. Patel and G. D. Abowd, "The contextCam: Automated point of capture video annotation," in *Proceedings Ubicomp 2004: The International Conference on Ubiquitous Computing*, pp. 301–318, 2004.

[143] S. N. Patel, K. N. Truong, G. R. Hayes, G. Iachello, J. A. Kientz, and G. D. Abowd, "The personal audio loop: A ubiquitous audio-based memory aid," in *Handbook of Research on User Interface Design and Evaluation for Mobile Technology*, (J. Lumsden, ed.), 2007.

[144] E. R. Pedersen, K. McCall, T. P. Moran, and F. G. Halasz, "Tivoli: An electronic whiteboard for informal workgroup meetings," in *Proceedings of CHI 1993: The ACM Conference on Human Factors in Computing Systems*, pp. 391–398, 1993.

[145] C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman, "LifeLines: Visualizing personal histories," in *Proceedings of CHI 1996: The ACM Conference on Human Factors in Computing Systems*, pp. 221–227, 1996.

[146] W. M. Post, M. A. A. Huis inít Veld, and S. A. A. van den Boogaard, "Evaluating meeting support tools," *Personal and Ubiquitous Computing*, vol. 12, no. 3, pp. 223–235, 2008.

[147] W. Pratt, K. Unruh, A. Civan, and M. M. Skeels, "Personal health information management," *Communications of the ACM*, vol. 49, no. 1, pp. 51–55, 2006.

[148] J. Rekimoto, "Time-Machine computing: A time-centric approach for the information environment," in *Proceedings of UIST 1999: The ACM Symposium on User Interface Software and Technology*, pp. 45–54, 1999.

[149] B. J. Rhodes and T. Starner, "Remembrance agent: A continuously running automated information retrieval system," in *Proceedings of the 1st International Conference on the Practical Application of Intelligent Agents and Multi Agent Technology*, pp. 487–495, 1996.

[150] H. A. Richter, "Designing and evaluating meeting capture and access services," Ph.D. Dissertation: 194 pages, Georgia Institute of Technology, 2005.

[151] H. A. Richter, G. D. Abowd, W. Fuchs, S. Daijavad, and S. E. Poltrock, "Integrating meeting capture within a collaborative team environment," in *Proceedings of Ubicomp 2001: The International Conference on Ubiquitous Computing*, pp. 123–138, 2001.

[152] H. A. Richter, P. Schuchhard, and G. D. Abowd, "Automated capture and retrieval of architectural rationale," Technical Report GIT-GVU-98-37, Georgia Institute of Technology, Atlanta, GA, 1999.

[153] K. Rodden and K. R. Wood "How do people manage their digital photographs?" in *Proceedings of CHI 2003: The ACM Conference on Human Factors in Computing Systems*, pp. 409–416, 2003.

[154] B. Rosensteing, "Video use in social science research and program evaluation," *International Journal of Qualitative Methods*, vol. 1, no. 3, p. 2, 2002.

[155] J. Rowan and E. D. Mynatt, "Digital family portrait field trial: Support for aging in place," in *Proceedings of CHI 2005: The ACM Conference on Human Factors in Computing Systems*, pp. 521–530, 2005.

[156] R. Sarvas, E. Herrarte, A. Wilhelm, and M. Davis, "Metadata creation system for mobile images," in *Proceedings of the 2nd International Conference on Mobile Systems, Applications, and Services*, pp. 36–48, 2004.

[157] J. M. Scaife and M. Pomerantz, "A survey of the record-keeping practices of clinical psychologists," *Clinical Psychology and Psychotherapy*, vol. 6, no. 3, pp. 210–226, 1999.

[158] P. Schafly, "What's wrong with outcome-based education?' *The Phyllis Schlafly Report*, vol. 26, no. 10, pp. 1–4, 1993.

[159] B. N. Schilit, N. Adams, R. Gold, M. M. Tso, and R. Want, "The PARCTAB mobile computing system," in *Proceedings of Workshop on Workstation Operating Systems*, pp. 34–39, 1993.

[160] B. Schneiderman and H. Kang, "Direct annotation: A drag-and-drop strategy for labeling photos," vol. 4, pp. 88–95, 2000.

[161] A. Sellen, A. Fogg, S. Hodges, and K. Wood, "Do life-logging technologies support memory for the past? An experimental study using SenseCam," in *Proceedings of CHI 2007: The ACM Conference on Human Factors in Computing Systems*, pp. 81–90, 2007.

[162] J. Smyth and A. Dow, "What's wrong with outcomes? Spotter planes, action plans and steerage of the educational workplace," *British Journal of Sociology of Education*, vol. 19, no. 3, pp. 291–303, 1998.

[163] G. D. Spindler and L. S. Spindler, *Fifty Years of Anthropology and Education: 1950–2000.* Lawrence Earlbaum Associates, 2000.

[164] Q. Stafford-Fraser and P. Robinson, "BrightBoard: A video-augmented environment," in *Proceedings of CHI 1996: The ACM Conference on Human Factors in Computing Systems*, pp. 134–141, 1996.

[165] T. Starner, "The virtual patrol: Capturing and accessing information for the soldier in the field," in *Proceedings of CARPE 2006: The 3rd ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, p. 2, 2006.

[166] M. M. Stevens, G. D. Abowd, K. N. Truong, and F. Vollmer, "Getting into the living memory box: Family archives and holistic design," *Personal Ubiquitous Computing*, vol. 7, no. 3, pp. 210–216, 2003.

[167] L. J. Stifelman, "The audio notebook," Ph.D. Dissertation: 150 pages, Massachusetts Institute of Technology, 1997.

[168] L. J. Stifelman, B. Arons, and C. Schmandt, "The audio notebook: Paper and pen interaction with structured speech," in *Proceedings of CHI 2001: The ACM Conference on Human Factors in Computing Systems*, pp. 182–189, 2001.

[169] N. A. Streitz, J. Geißler, J. M. Haake, and J. Hol, "DOLPHIN: Integrated meeting support across local and remote desktop environments and liveboards," in *Proceedings of CSCW 1994: The ACM Conference on Computer Supported Cooperative Work*, pp. 345–358, 1994.

[170] Y. Sumi, T. Etani, S. Fels, N. Simonet, and K. Mase, "C-MAP: Building a context-aware mobile assistant for exhibition tours," in *Community Computing and Support Systems.* LCNS vol. 1519, pp. 137–154, 1998.

[171] Y. Sumi, R. Sakamoto, K. Nakao, and K. Mase, "ComicDiary: Representing individual experiences in a comics style," in *Proceedings of Ubicomp 2002: The International Conference on Ubiquitous Computing*, pp. 16–32, 2002.

[172] D. Tan, A. Sellen, M. Czerwinski, G. Bell, J. Gemmell, K. Wood, G. Robertson, N. Oliver, E. Berry, S. Hodges, and N. Kapur, "Supporting human memory with a personal digital lifetime store," in *Personal Information Management: Challenges and Opportunities*, (J. Teevan and W. Jones, eds.), University of Washington Press, 2007.

[173] Q. T. Tran, G. Calcaterra, and E. D. Mynatt, "Cook's Collage: Déjà vu Display for a Home Kitchen," in *Proceedings of HOIT 2005* (*Home Oriented Informatics and Telematics*), pp. 5–32, 2005.

[174] Q. T. Tran and E. D. Mynatt, "What was I cooking? Towards Déjà Vu displays of everyday memory," Technical Report, GIT-GVU-03-33: Georgia Institute of Technology, Atlanta, GA, 2003.

[175] K. N. Truong and G. D. Abowd, "INCA: A software infrastructure to facilitate the construction and evolution of ubiquitous capture and access applications," in *Proceedings Pervasive 2004: The International Conference on Pervasive Computing*, pp. 140–157, 2004.

[176] K. N. Truong, G. D. Abowd, and J. A. Brotherton, "Personalizing the capture of public experiences," in *Proceedings of UIST 1999: The ACM Symposium on User Interface Software and Technology*, pp. 121–130, 1999.

[177] K. N. Truong, G. D. Abowd, and J. A. Brotherton, "Who, what, when, where, how: Design issues of capture and access applications," in *Proceedings*

of Ubicomp 2001: The International Conference on Ubiquitous Computing, pp. 209–224, 2001.

[178] K. N. Truong, E. M. Huang, and G. D. Abowd, "CAMP: A magnetic poetry interface for end-user programming of capture applications for the home," in *Proceedings of Ubicomp 2004: The International Conference on Ubiquitous Computing*, pp. 143–160, 2004.

[179] K. N. Truong, S. N. Patel, J. W. Summet, and G. D. Abowd, "Preventing camera recording by designing a capture-resistant environment," in *Proceedings of Ubicomp 2005: The International Conference on Ubiquitous Computing*, pp. 73–86, 2005.

[180] S. Uchihashi, J. Foote, A. Girgensohn, and J. Boreczky, "Video manga: Generating semantically meaningful video summaries," in *Proceedings of the 7th ACM International Conference on Multimedia*, pp. 383–392, 1999.

[181] M. Van Kleek, M. Bernstein, D. R. Karger, and M. C. Schraefel, "GUI — Phooey!: The case for text input," in *Proceedings of UIST 2007: The ACM Symposium on User Interface Software and Technology*, pp. 193–202, 2007.

[182] S. Vemuri, C. Schmandt, W. Bender, S. Tellex, and B. Lassey, "An audio-based personal memory aid," in *Proceedings of Ubicomp 2004: The International Conference on Ubiquitous Computing*, pp. 400–417, 2004.

[183] T. Walmsley and S. Neilsen, The VIDEOSHARE project. Missoula, Montana, USA: Montana University Division of Educational Research and Services, 1991.

[184] R. Want, A. Hopper, V. Falcão, and J. Gibbons, "The active badge location system," *ACM Transactions on Information Systems*, vol. 10, no. 1, pp. 91–102, 1992.

[185] R. Want, T. Pering, G. Danneels, M. Kumar, M. Sundar, and J. Light, "The personal server: Changing the way we think about ubiquitous computing," in *Proceedings of Ubicomp 2002: The International Conference on Ubiquitous Computing*, pp. 194–209, 2002.

[186] M. Weiser, "The Computer for the 21st Century," *Scientific American*, vol. 265, no. 30, pp. 94–104, 1991.

[187] D. R. White, J. A. Camacho-Guerrero, K. N. Truong, G. D. Abowd, M. J. Morrier, P. C. Vekaria, and D. Gromala, "Mobile capture and access for assessing language and social development in children with autism," in *Extended Abstracts of Ubicomp 2003: The International Conference on Ubiquitous Computing*, pp. 137–140, 2003.

[188] S. Whittaker, P. Hyland, and M. Wiley, "FILOCHAT: Handwritten notes provide access to recorded conversations," in *Proceedings of CHI 1994: The ACM Conference on Human Factors in Computing Systems*, pp. 271–276, 1994.

[189] S. Whittaker, S. Tucker, K. Swampilla, and R. Laban, "Design and evaluation of systems to support interaction capture and retrieval," *Personal and Ubiquitous Computing*, vol. 12, no. 3, pp. 197–221, 2008.

[190] L. D. Wilcox, B. N. Schilit, and N. Sawhney, "Dynomite: A dynamically organized ink and audio notebook," in *Proceedings of CHI 1997: The ACM Conference on Human Factors in Computing Systems*, pp. 186–193, 1997.

[191] W. Xu, Y. Chen, H. Sundaram, and T. Rikakis, "Multimodal archiving, real-time annotation and information visualization in a biofeedback system for stroke patient rehabilitation," in *Proceedings of CARPE 2006: The Third ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, pp. 3–12, 2006.

[192] E. Yakel, "The social construction of accountability: Radiologists and their record-keeping practices," *The Information Society*, vol. 17, no. 4, pp. 233–245, 2001.

[193] R. Yeh, C. Liao, S. Klemmer, F. Guimbretière, B. Lee, B. Kakaradov, J. Stamberger, and A. Paepcke, "ButterflyNet: A mobile capture and access system for field biology research," in *Proceedings of CHI 2006: The ACM Conference on Human Factors in Computing Systems*, pp. 571–580, 2006.