Electronic Chronicles: Empowering Individuals, Groups, and Organizations

Gopal Pingali
IBM T.J. Watson Research Center
Hawthorne, NY 10532, USA
gpingali@us.ibm.com

Ramesh Jain
School of Information and Computer Sciences
University of California, Irvine, USA
jain@ics.uci.edu

Abstract

Continuing strides in processing, storage, sensing, and networking technologies are enabling people to capture their activities and experiences as greater volumes of ever-richer media. A big emerging challenge today is the organization, retrieval, and exploitation of such multimedia data surrounding the activities of individuals or enterprises. The field of multimedia electronic chronicles deals with the unified contextual organization, presentation, and analysis of temporal streams of multimedia data captured by individuals, groups, or organizations. The value of electronic chronicles is in converting activity and experience from the past into actionable intelligence in the present. Such multimedia electronic chronicles, with their associated techniques for search and navigation, analysis and reasoning, and prediction and alerting, will have enormous impact on various spheres of life spanning enhancement of personal life, business productivity, entertainment, and government operations. This paper, which serves as a companion to a keynote talk by the first author at ICME 2005, explains the notion of electronic chronicles and outlines the research challenges in this area.

1. Introduction

With the advances in computing, communication, and sensing technologies, the last few years have seen tremendous growth in the ability to capture and store information in ever-decreasing form factors. With the continuing strides made in semiconductor technology, it is clear that today's platforms – PCs, smart phones, and servers – will be dwarfed in a few years, just as the current platforms have rendered insignificant those from only a few years ago. Table 1 shows the projected processor speed, primary storage, and secondary storage of PC's, PDA's, and servers by 2010. A cell phone could have 500 gigabytes of storage while a server could have 5 billion gigabytes of storage. A big question asked today is whether this means that more information and experiences are captured in greater detail or if it only means that it will be even harder to find anything useful and that the enormous storage will mostly be for "write and forget".

So a big challenge before computer scientists today is how to facilitate the access and increase the utility of the voluminous capture of information and experiences. In particular, how can we better acquire or derive the context surrounding an information transaction or a captured experience? And how can we retrieve the information or experiences that are most relevant to the current context? People have traditionally written diaries or chronicles to capture important experiences and associated context. With the proliferation and dominance of electronic devices and media in various spheres of human activity, it is now possible to obtain electronic chronicles [10] that are created automatically or semi-automatically. We see a big emerging opportunity in tools for creation, navigation, and analysis of such electronic chronicles. The promise of such electronic chronicles is that they will be as effective, or potentially much more powerful, than written chronicles in associating context and providing insights.

2010 Computing Platforms			
	Desktop	Cell phone/PDA	Server
Processor	50 GHz	15 GHz	50 GHz
Memory	50 GB	15 GB	200 GB+
Storage	3 TB	500 GB	5 EB

Figure 1. Anticipated processor speed, primary memory, and secondary storage on computing platforms in 2010.

2. Scope of Electronic Chronicle Research

Historical accounts of important events in human society, minutes of a meeting, proceedings of a conference, wedding videos, surveillance videos, logs of visitors, warehouses of all sales activities of a retail store chain, and after-action reports by soldiers are all examples of keeping records of what happened. Such record-keeping or logging has thus far mostly been in paper and alphanumeric form. *The goal of electronic chronicling research is to go beyond mere diary-like partial reports of what happened to very rich multimedia records created in*

real time that are searchable, shareable, discoverable, visualizable, and instantly actionable.

We define an electronic chronicle as an extensive record of events obtained using multiple sensors and sources of information, which provides access to this data at multiple levels of granularity and abstractions along temporal and other contextual dimensions, using appropriate access mechanisms in representations and terminology familiar to application users.

Thus, while our vision of an electronic chronicle emphasizes rich capture or logging of information and events, it goes well beyond such logging to include sensor processing, unification and assimilation of multiple structured and unstructured information sources, extraction of domain semantics at multiple levels of abstraction, and providing experiential [12] navigation mechanisms.
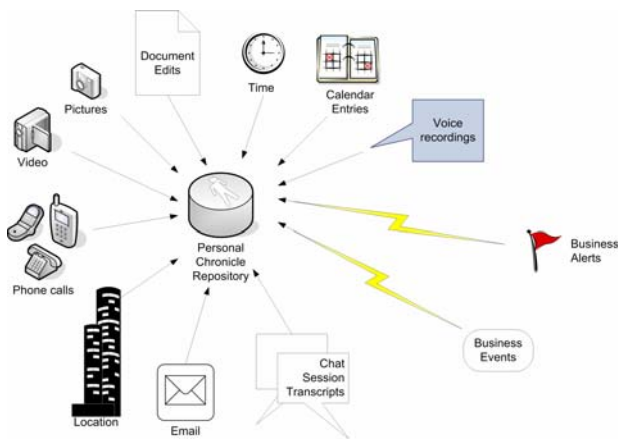

Figure 2. Data Sources for Electronic Chronicles

Figure 2 illustrates the various sources of information that are unified in a personal electronic chronicle. As seen here, these include traditional activities performed by an individual on a PC or PDA such as document edits, email, chat sessions, and calendar entries. In addition, these could include data from other devices such as digital still cameras or video cameras, voice recordings from microphones, logs or recordings of phone calls or phone messages, and estimates of location of the individual, obtained through sources such as GPS systems or wireless localization techniques. An individual in a business organization may also receive business events and business alerts that also get unified in the individual's chronicle.

3. Personal, Group, and Organizational Chronicles

In fact, we see electronic chronicles playing a significant role not only at the individual level, but across personal, group, and organizational levels. At all these levels, the ability to organize and retrieve information by context becomes very important. While chronicles have traditionally been seen as a means of summarizing and recalling experiences and events, and predominantly as a memory aid, we see electronic chronicles playing as significant a role in deriving crucial insights, providing critical alerts, and driving significant actions.

Electronic chronicles can be considered along two dimensions of complexity. One is people complexity, and the other is the order or complexity of search performed on the chronicle. This is illustrated in Figure 3. People complexity indicates the number of people chronicled, while the search complexity indicates the sophistication of the query performed on the chronicle. For simplicity, Figure 3 shows three levels of people complexity – person, (small) group, and (large) organization. The Figure also shows three orders of search complexity. The first order query indicates retrieval of facts already stored in the chronicle repository – such as who, what, when, or where. These are typically implemented as relational database queries. Second order queries involve a first level analysis of the stored facts to derive summaries, detect interactions etc. Third and higher order queries require detection of complex patterns involving multiple levels of interactions and analysis of the stored data. A few simple examples of order 1, order 2 and order 3 queries are shown in Figure 3.

Search Complexity

People Complexity	Order1	Order 2	Order 3
Individual	Where was I? When did I last see this?	What is the average time I take on this task?	I am most productive when I get up at 8, have lunch at 3, and listen to Jazz
Group	When was the last group meeting? Who attended?	Show me the highlights of last month's meetings	The experts first show price, then features, then a video
Organization	Who finished a Java project last month?	Which groups that work on Java have interacted in the last one year?	The most productive groups have leader with less than 4 yrs experience, have pizza lunches once a month, and spend 30% of time working from home

Direct queries   Mining/alerts

Figure 3. Search complexity (columns) and people complexity (rows) in electronic chronicles. Lower levels of people and search complexity are more suited for direct queries by humans while higher levels of search or people complexity are more suited for discovery/mining by machine, resulting in alerts to humans.

An important observation is that lower orders in search or people complexity imply that these queries can be easily stated and executed by people. Higher levels of complexity make it increasingly difficult for people to think of or formulate the queries. These are more amenable to mining by machine, or by cooperation

between man and machine in a visualization or experiential environment. These are the types of queries that a machine can perform in the background and alert people based on the significance of the result.

As the volume of data captured in electronic chronicles keeps growing relative to the time that people have to navigate the data, it becomes important to continuously mine the data and alert people when appropriate. We see the dominant mode of usage, and hence the dominant research activity in this field, to move towards sophisticated analysis, search, and mining of electronic chronicles.

Another significant challenge in electronic chronicles is building and deriving electronic chronicles at the group and organizational levels, i.e. increasing the people complexity of the chronicle. Most related work in this area (Section 5) has been limited to personal chronicles. Blogs have emerged recently as a dominant form of personal electronic chronicles while other forms such as wikis have appeared as initial examples of group chronicles. Besides becoming powerful social media, blogs and wikis are beginning to impact enterprises as well.
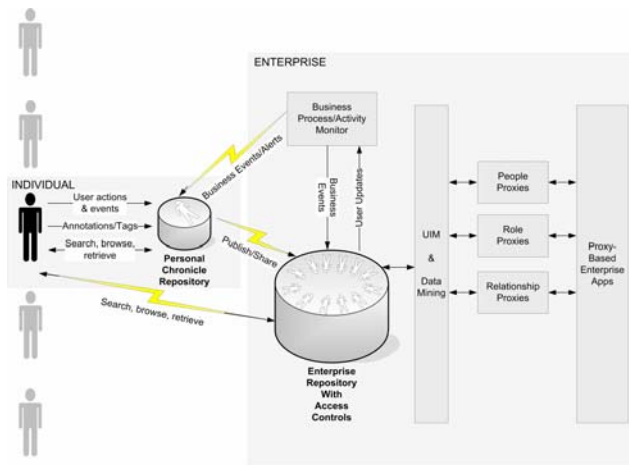

Figure 4. Example of going from personal or individual chronicles to organizational or enterprise chronicles

However, blogs, wikis, web pages etc. represent only a minute portion of the enterprise data that can potentially be exploited for improving enterprise productivity. Thus, another challenge is to develop approaches for deriving richer chronicles in the enterprise starting from individual level to enterprise level. A significant issue here is that of privacy and security while developing the chronicles. Any approach to electronic chronicles should have inherently supported notions of security and privacy. Figure 4 shows one example (from [13]) on how enterprise chronicles and related

applications are built from chronicles of individuals in an enterprise. Here, the chronicles provide easy-to-use mechanisms for sharing or publishing information with built-in access controls. Group and enterprise chronicles are built based on this shared information, while preserving the privacy of individual chronicles.

4. Chronicling in the Mobile Era

Among the many forces driving the emergence of electronic chronicles, perhaps the strongest is mobile computing and communication technology. Smart phones are emerging as true multimedia devices that combine communications, computing, and content [11]. These devices are ensuring that electronic chronicling can truly happen anytime and anywhere in a rich multimedia form. What can be more helpful for chronicling than a device that fits in your pocket, notes time and location, enables capture of images or video any time, allows you to make notes by simply speaking, facilitates instantaneous sharing of information with others, and provides alerts and notifications?

Cell phones are already available today that provide PDA, GPS, camera, microphone, and Bluetooth capabilities. With these devices, it is possible to capture and store both time and location information continuously, thus providing two vital contextual cues to any activity or experience. WiFi and 3G infrastructure available today has already set the stage for instantaneous sharing and publishing of chronicled data. Emerging wireless broadband technologies such as WiMax are enabling sharing of much richer media and content.

The sensing and communication capabilities of mobile devices, coupled with electronic chronicling technology, have the potential to change the everyday practices of people in both personal life and business activities. We have seen a significant change in the way numerous mobile professionals operate today, routinely using cell phone-PDA's while on the road or in meetings for calendaring, email, access, and chatting. In the near future, we can expect them to do much more of media capture and speech annotations for the purpose of chronicling. For example, few professionals make frequent notes to themselves by talking to themselves although the technology for doing this exists today. If it is extremely easy to retrieve these speech annotations (by when, where, what, who etc.) however, making such annotations can become much more commonplace. Sales people, for example, can make notes to themselves after a meeting, when things are still fresh in their mind, so long as they can retrieve these easily for reporting or recall.

The potential of mobile devices for more effective chronicling is being rapidly recognized in various domains. It is already commonplace in alphanumeric form at car-rental operations, shipping companies, inventory management at warehouses etc.

There are efforts under way to use wearable devices for richer situation awareness and decision making in military and law-enforcement operations (for example, [19]).

The multimedia field has been dominated and, indeed, created by the emergence of new multimedia input and output capabilities on the personal computer. The field has not yet made a shift towards multimedia for the cell phone although it is already the most ubiquitous and versatile multimedia platform available today. We see electronic chronicling technologies accelerating this shift of multimedia towards mobile devices, while continuing to enhance the information management and analytics on back-end PCs and servers.

5. Multimedia Research Challenges for Electronic Chronicles

Multimedia capture and analysis is clearly vital to realizing electronic chronicles. First, by definition, electronic chronicles unify various sources of information. These almost always involve multiple media even on a single device such as a PC or PDA-cell phone. Secondly, a variety of new sensors are becoming mainstream and integral to human life. These include digital cameras, microphones (especially in cell phones), GPS devices, radio frequency identification tags etc. The increasing usage of these devices has come to mean that an electronic chronicle has to be a multimedia electronic chronicle. Thirdly, the multimedia data is fast increasing to become the dominant and major part of stored data, reducing textual sources to a small fraction of chronicles. Besides, the unstructured nature of multimedia data poses one of the greatest challenges in realizing electronic chronicles.

The research issues in multimedia electronic chronicles span the gamut of capture, sensor analysis, storage, compression, transmission, indexing, unification, semantic modeling, search, retrieval, browsing, mining, visualization, notification, and presentation. In particular, the temporal nature of chronicles will require them to continuously process multiple data streams to segment the raw data at multiple levels of abstraction to include events, activities, and episodes. Another challenge in multimedia processing for electronic chronicles is to derive and associate rich context with the raw data by unified analysis of the multiple media. An important notion that has not been sufficiently addressed today is that of representation of activity and analytics and mining of activities. This notion is central to electronic chronicling technology.

We see speech, image, GPS, and text as the data sources that will initially dominate electronic chronicles, especially on mobile devices. It will be important to develop sensor analytics on these sources and techniques for detecting events based on combined analysis across these multiple modalities. It will be important to facilitate easy annotations by the user, particularly using speech.

Another challenge is to effectively combine existing alphanumeric data sources such as documents, email, chats, calendar entries etc. with sensory data such as images, video, and audio. The effective assimilation of these sources, combined with effective user input, holds the promise of bridging the signal to symbol gap for electronic chronicling applications.

Presentation tools that combine navigation, visualization, browsing, and notification will also be a dominant theme in the research in this area. Figure 5 shows an initial example of a chronicle navigation tool developed in [7]. Other examples of chronicles and navigation tools for sporting events can be found in [12,14]. This interface allows unified browsing of all the activities on a PC desktop or a PDA. This goes beyond a single user to show events shared with and shared by others.
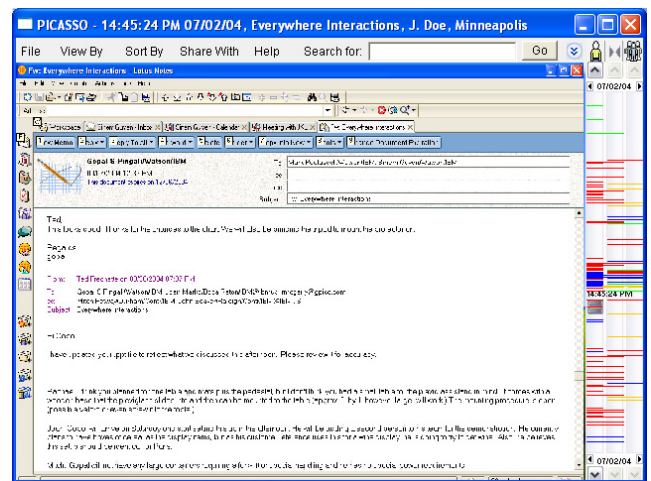

Figure 5. Example of a Chronicle Navigator running on a Desktop. The main window (View Panel) shows the selected event, a received e-mail in this case. Notice that the title bar displays the meta data about the selected event including the name and the location of the user. Slider Panel on the right shows all events for a selected timeframe. The events include (left to right): 1) Personal Events, 2) All Shared Events (both incoming and outgoing) 3) Incoming Events shared by other people. Each line represents an event and its color represents the user's location at that time.

5. Related Research

A number of previous and ongoing research activities relate to the field of electronic chronicles. The most popular research theme has been personal

information management and retrieval of personal information. This goes back at least to the 1945 vision of Vanevar Bush [1]. This work has inspired a number of efforts including MyLifeBits [4], as well as the workshop and web site on CARPE (Capture, Archival, and Retrieval of Personal Experiences) [5,6]. The advent of wearable devices and wearable computing [18] has had significant influence on the recent emergence of such efforts [e.g. 17] around personal chronicles that enable rich capture of an individual's life and retrieval based on context [8].

Other efforts have addressed the difficulties in organizing the regular information (even if not multimedia) stored on computers. A good example of this is [2] and the several commercial products that have recently been announced for "desktop search" such as products from Google Inc.

Several researchers have recognized the importance of time in the organization, visualization, and retrieval of stored personal information. These include efforts such as [3,15,16]. Such approaches are integral to electronic chronicles.

The area of group and enterprise level chronicles remains largely unexploited, especially relative to the ongoing efforts on personal chronicles. One effort moving in this direction is [9]. Other recently initiated efforts include [13,7].

A rich area for research is mining of chronicle data and generation of associated alerts and notifications. Overall, most current research in electronic chronicles is dominated by efforts limited to order 1 and order 2 queries at the personal level (in Figure 3) . While there is opportunity for deriving ever richer useful data even at this level, the remaining parts of Figure 3 represent large open areas for research in this field. Such efforts have thus far been centered on data warehouses and enterprise data mining. The challenge is to bring these approaches into end-to-end electronic chronicling solutions.

6. Conclusions

A big research opportunity in electronic chronicles is emerging today, driven by the ever-growing capacity of storage, sensing, and computational power. Unstructured multimedia data is already the dominant form of data being captured today, and is hence resulting in the development of multimedia electronic chronicles.

Such multimedia electronic chronicles, with their associated techniques for search and navigation, analysis and reasoning, and prediction and alerting, will have enormous impact on various spheres of life spanning enhancement of personal life, business productivity, entertainment, and government operations.

7. Acknowledgements

References

1. V. Bush As we may think. Atlantic Monthly, 176 (1), pp. 641 - 649, January 1945.
2. S. Dumais, E. Cutrell, JJ Cadiz, G. Hancke, R. Sarin, D.C. Robbins. Stuff I've Seen: A System for Personal Information Retrieval and Re-Use. Proceedings of ACM SIGIR '03, Toronto, Canada, August 2003.
3. E. T. Freeman and S. J. Fertig. Lifestreams: Organizing your electronic life, AAAI Fall Symposium; AI Applications in Knowledge Navigation and Retrieval, 1995.
4. J. Gemmell, G. Bell, R. Lueder, S. Drucker, and C. Wong. MyLifeBits: Fulfilling the Memex Vision. ACM Multimedia '02, pp. 235 - 238.
5. J. Gemmell and H. Sundaram. Proceedings of the First ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE 2004), New York, October 2004.
6. J. Gemmell. ACM SIGMM CARPE home page. http://www.sigmm.org/jgemmell/CARPE.
7. S. Guven, M. Podlaseck, and G. Pingali. PICASSO: Pervasive Chronicling, Access, Search, and Sharing for Organizations. To appear in IEEE International Conference on Pervasive Computing, Hawaii, March 2005.
8. T. Hori and K. Aizawa. Capturing Life Log and Retrieval based on Context. In Proc. IEEE ICME 2004, June 2004.
9. D. Huynh, D. Karger, and D. Quan. Haystack: A platform for creating, organizing and visualizing information using RDF.
10. R. Jain. Media Vision: Multimedia Electronic Chronicles. IEEE Multimedia, July 2003.
11. R. Jain. Media Vision: A True Multimedia Client, IEEE Multimedia, April 2005.
12. Ramesh Jain, Experiential Computing. Communications of ACM, July 2, 2003, pp. 48-55.
13. P. Kim, M. Podlaseck, and G. Pingali. Personal Chronicling Tools for Enhancing Information Archival and Collaboration in Enterprises. ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE 2004), New York, October 2004.
14. G. S. Pingali, Y. Jean, A. Opalach, and I. Carlbom. LucentVision: Converting Real World Events into Multimedia Experiences. In the Proceedings of the IEEE International Conference on Multimedia and Expo, New York, July 2000
15. C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman. Lifeline: Visualizing Personal Histories. Proceedings of ACM Conference on Human Computer Interaction (CHI'96), Vancouver, USA, ACM Press 221-227, 1996.
16. J. Rekimoto. Timescape: A time machine for the desktop environment.
17. B. Rhodes and T. Starner. Remembrance Agent: A continuously running automated information retrieval system. Proceedings of the First International Conference on the Practical Application of Intelligent Systems and Multi Agent Technology. London, April 1996.
18. Thad Starner. Wearable Computing and Context Awareness. Ph. D. Thesis, MIT Media Lab, April 30, 1999.
19. Advanced Soldier Sensory Information System and Technology. Broad Agency Announcement from Defense Advanced Research Projects Agency (DARPA) http://www.darpa.mil/ipto/solicitations/open/04-38_PIP.htm. August 2004.