

EXTENDED MEDIA RESEARCH AT THE FX PALO ALTO LABORATORY

Stephen W. Smoliar
James D. Baker
FX Palo Alto Laboratory, Inc.
3400 Hillview Avenue, Bldg. 4
Palo Alto, California 94304
Fax: +1-415-813-7081
Internet: smoliar@pal.xerox.com

Abstract

The fundamental objective of Extended Media research is the *empowering of documents through technology*. We see our goal as that of *inventing* documents which communicate more effectively. Furthermore, we need to make it easier for the *writer* to *record* what must be communicated and for the *reader* to *access* it. Thus, we anticipate that one cannot invent new documents without also inventing new processes for both writing and reading. Our primary research thrust is thus in the authoring of hypermedia documents, supplemented by a secondary thrust concerned with the problem of managing archives and libraries where more than text is involved.

The FX Palo Alto Laboratory (FXPAL) is the American research laboratory affiliated with the Corporate Research Laboratories (CRL) of the Fuji Xerox Corporation. Our mission is to become a significant "change agent" for software systems, services and capabilities supporting the next generation of Fuji Xerox digital products; and our objective is to pursue this mission by developing the vision of the next generation of software to support office work of the future. Two themes contribute to this vision: empowering documents to become dynamic media of communication for the future, and empowering organizations to become the offices of the future. FXPAL is currently divided into three groups, each of which is concerned with a mixture of these two themes:

1. **Office Innovation** involves operations in the office of the future and the support of those operations through local network (Intranet) services connected to the Internet. Specific projects are concerned with a corporate information server and document understanding as a tool for information management and work-flow support.

We are also investigating the use of adaptive software to contribute to these projects, and we are forming a Document Innovation Laboratory to address how technologies will be integrated to support the role of documents in the office of the future.

2. **Global Wireless Access** is concerned with supporting participation in office activities by remote parties equipped with wireless communication technology. On the one hand it addresses those technologies which will be necessary to implement its vision. At the same time it must also address the fact that the limitations of wireless hardware may require that documents may have to be delivered to remote users differently from their delivery within the office itself.

3. **Extended Media** is concerned with the role of multimedia in both the document of the future and office of the future, and this group will be discussed in greater detail in the sequel.

Finally, these three groups are supported by foundation research in the technologies of adaptive software and software agents. The Extended Media group will now be examined in greater detail.

The fundamental objective of the Extended Media group is the *empowering of documents through technology*. We see our goal as that of *inventing* documents which communicate more effectively. Recalling Buckminster Fuller's motto that he always wanted "to make more and more with less and less," we see effective communication as being achieved by capturing more content with fewer resources. Furthermore, we need to make it easier for the *writer to record* what must be communicated and for the *reader to access* it. Thus, we anticipate that one cannot invent new documents without also inventing new processes for both writing and reading.

Multimedia research is necessary by virtue of the very nature of the document, itself: Documents are *physical* media which explicitly describe knowledge. Because documents are physical, it is necessary to consider how the scope of their physical form can be expanded to accommodate new needs for knowledge. This is where multimedia enters the picture, but the *real role* of multimedia remains a question to be investigated.

Nevertheless, we feel that role is important since just about every form of work is a

multimedia experience where the processing of text must inevitably be supplemented by working with some combination of images, sounds, and video.

Because our facility is comparatively small, our progress will require that we establish fruitful collaborative efforts. Our major sources for collaboration will be the Xerox Palo Alto Research Center (PARC) and two universities in California, the University of Southern California and the Berkeley campus of the University of California. Plans for collaboration will now be reviewed, beginning with our work with PARC.

A major research effort at PARC has involved the capture and access of information in group meetings. This was recently documented in the paper "A Confederation of Tools for Capturing and Accessing Collaborative Activity," written by Scott Minneman and his colleagues and presented at ACM Multimedia '95. This effort has developed technologies for collecting video and audio recordings of group meetings, as well as interactions which take place through LiveBoard hardware which may have network connections to laptop computers controlled by individual participants in a meeting. Recent research has been concerned with extracting "units of knowledge" (also called "salvaging"), once these recordings have been made. Our collaboration is concerned with two key issues: First is the "authoring" problem of how, and under what circumstances, recorded material should and can be incorporated into hypermedia documents. Second is the problem of indexing and retrieval, through which the author of such a document will be able to identify and access material to be considered for incorporation.

The University of Southern California has organized an Integrated Media Systems Center involving the participation of faculty beyond the usual scope of engineering for areas such as journalism and the performing arts. Our collaboration with this Center has begun with our working with Professor Carolyn Garrett Cline of the School of Journalism. Professor Cline has long been interested in the phenomenon that different cultures require documents to be presented visually in different ways in order for communication to be effective. Thus, there is more to the difference between an American

newspaper and a Japanese newspaper than the language of the text or even the fonts required to display that text. Given that our own business context will involve the exchange of documents between offices in the United States and in Japan, we feel it is important to understand how the translation of those documents may involve more than translating the language of the text and to identify software tools which might support "cultural translation" as a supplement to language translation.

The Berkeley Multimedia Research Center (BMRC) of the University of California is pursuing two areas of research where we shall be collaborating. Most important is their work in the development of tools to support multimedia authoring, particularly concerned with the access and presentation of *continuous* (i.e. time-dependent) media. BMRC is also supporting a major effort in digital libraries which we also feel is important to the authoring process. As is the case in dealing with the PARC recording technology, authoring always involves some degree of referring to external material. The hypermedia document allows that not only can such material be cited but also links can be installed to connect to it. Thus a system which will support the management of an extensive digital library of reference material to be incorporated into new documents is actually an essential authoring tool.

The primary research thrust within the Extended Media group, then, will be the authoring of hypermedia documents. We feel it is particularly important that multimedia documents be *truly* multimedia, rather than text with other media stuck on as auxiliary attachments, as is currently the case with multimedia available from, for example, the World Wide Web. However, since the Web is becoming equipped to deliver *software* as a medium ("applets"), we also feel it is important to be flexible and accommodating in regard to just what the *scope* of multimedia can be. Furthermore, in respect to the needs of humans who must interact with their documents in a variety of circumstances, authoring must accommodate not only the creation of extended documents, where all media are on equal footing, but also *paper* documents which effectively capture the content of those extended documents.

Our secondary research thrust is concerned with the problem of managing archives and libraries where more than text is involved. This is the general problem of content understanding which will relate our work to activities in document understanding being addressed by the Office Innovation group. We are particularly interested in *content-based storage* as an approach to effectively storing and indexing large quantities of information in multiple media.

We plan to take a case study approach to pursuing these research thrusts, concentrating our attention on the writing and reading of documents which address specific needs. The primary domain we shall investigate will be project reporting (beginning with activities within FXPAL); and we hope to extend our scope to work with some examples of academic and professional publication. What we expect to learn from these case studies can be summarized in three questions:

1. **How can we capture relationships among elements of content** (both within a single document and across a corpus of documents)?
2. **What are those relationships like when more than text is involved?** Here we shall concentrate our attention on time-varying media in order to cultivate our Berkeley collaboration. However, we shall also be interested in images and, most likely, other media.
3. **How will changing the nature of documents change the process of reading?**

This last question is, admittedly, rather philosophical in nature; but it is still important. There is no sense in inventing a document of the future if it turns out that such a document communicates far more poorly than the documents of the past.

While these are ambitious questions, we feel our case studies should be viable. Most important is that they provide us with limited semantic domains which should lead to content models which can be supported with practical resources and development effort. In addition a case study based on our own needs for documentation enables us to use our own facility as our environment for experimentation: Our own facilities provide an

abundance of sample material for analysis, as well as the first rank of users available to participate in trial studies.

Let us conclude, then, with a brief summary of our plan of attack which basically embodies our approach to Extended Media research:

- **The study of documents must begin with the study of communication.** The study of communication, however, must draw upon a variety of disciplines which, in the past, have not interacted very well. First we feel it is necessary to turn to artificial intelligence and the technologies of knowledge representation for an initial approach to *what* is being communicated. However, the reading of documents will always involve some level of subjective interpretation; so an understanding of literary criticism may be just as important as a command of knowledge representation. Finally, the study of writing and reading as *processes* is one of the concerns of semiology; so that discipline is likely to be as important as artificial intelligence and criticism.
- **Concentrate on case studies which focus on limited domains of communication.**
- **Develop tools which facilitate communication for those specific case studies.** Those tools need to address three areas:
 1. *Writing*, including both the recording of narrative (telling the right story) and representing those associations which will make the document "hyper."
 2. *Reading*, which must be concerned with both retrieving content and following the narrative.
 3. *Editing*, perhaps the most subtle area which requires "matching" the writer and the reader.