

# Digital Annotation of Printed Documents

Corsin Decurtins  
Institute for Information  
Systems, ETH Zurich  
CH-8092 Zurich, Switzerland  
decurtins@inf.ethz.ch

Moira C. Norrie  
Institute for Information  
Systems, ETH Zurich  
CH-8092 Zurich, Switzerland  
norrie@inf.ethz.ch

Beat Signer  
Institute for Information  
Systems, ETH Zurich  
CH-8092 Zurich, Switzerland  
signer@inf.ethz.ch

## ABSTRACT

We present a general model and information server for the digital annotation of printed documents. The resulting annotation framework supports both informal and structured annotations as well as context-dependent services. A demonstrator application for mammography that features both enhanced writing and reading activities is described.

## Categories and Subject Descriptors

H.5.4 [Hypertext/Hypermedia]: Architectures, User Issues

## General Terms

Design

## Keywords

Cross-Media Annotation, Augmented Paper, Mammography

## 1. INTRODUCTION

Annotations are heavily used in the world of paper to enhance both reading and writing activities [6, 11]. Further, they often provide a basis for communication and collaboration whether it be as part of a co-authoring activity or publishing annotations to aid the interpretation of data or comment on its quality. There have been a number of proposals therefore to provide annotation tools in the digital world either as part of general web infrastructure [3] or for specific domains such as the annotations of brain images [1].

Despite predictions of the paperless office, paper and printed documents are still prevalent in current work practices [13]. Our goal therefore is to develop a general annotation framework that can support all forms of annotation [7], not only within a digital world, but also across media. Here we present a general model and information server for the digital annotation of printed documents and, by means of a demonstrator application for mammography, discuss how various technologies for digitally augmented paper can assist in the collection, analysis and reporting of scientific data.

In section 2, we discuss the various forms of annotation and the role that paper can or does play to enhance the associated activities.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CIKM'03 November 3–8, 2003, New Orleans, Louisiana, USA.  
Copyright 2003 ACM 1-58113-723-0/03/0011 ...\$5.00.

Section 3 introduces the technologies that can be used to annotate printed materials. Details of the annotation model are given in section 4 and the mammography application is presented in section 5. Concluding remarks are given in section 6.

## 2. ANNOTATION OF DOCUMENTS

Annotation is some means of marking up a document to augment its content. Annotations come in many forms and have a variety of uses [7]. They may be private or public, permanent or transient, and formal or informal. Informal annotations are often free text but could also take other forms such as sketches, images or audio.

Formal annotations follow defined structures and conventions that enable them to be interpreted by other persons or computer programs. An example is typographic markup for the editing of documents. The scientific database and semantic web communities use formal annotation to markup digital data with metadata that describes its semantic content and later enables either human readers or programs to access and process that data [1].

Most annotation tools support one particular form of annotation designed for specific activities such as collaborative writing, data integration, search and so on. However, users typically have to work with lots of different forms of information and perform a wide variety of tasks that may require different forms of annotation. For example, in a study of university textbooks [7], it was shown that some kinds of annotations were used to support *reading*, while others were used to support *writing*.

Another study examined the task of writing document summaries and how users annotate documents and take notes [11]. Users had to perform the task using either only digital documents or only paper. There were problems with the digital annotation systems in terms of both the input of annotations and also having to work with various documents alongside each other.

Within the scientific domain, paper forms are heavily used for collecting data and also reporting on the analysis of data. For example, in breast screening clinics, mammograms are analysed by experts referred to as “readers” to determine whether or not the patient should be called for further tests. For each patient, four mammograms are taken — two views of each breast — then the resulting film is analysed by a reader and the results reported on a paper form. The form contains breast outlines corresponding to the four mammogram views and the reader annotates these to indicate any abnormalities or special features and, generally, any reasons for their decision whether or not to recall the patient. Further details of the practices of breast screening clinics and their use of annotations are given in [2].

One could propose that the whole breast screening process be digitised. However, some clinics see the generation of digital versions of the mammograms as simply creating an extra step in the

process, resulting in extra delays and costs. Further, it is much more convenient and quicker for a reader to enter data on a paper form, especially the annotations.

This is just one example of many where users are very reluctant to move away from paper. Sometimes it may be part of general resistance to change, but studies of the working environment and work practices often reveal that there are very good reasons for remaining with paper and other forms of non-digital media. On the other hand, if we consider the breast screening example, there are certain activities such as the reviewing of previous reports and annotations that could be improved with digital support. Summarising, we think it is important to provide a general annotation framework that can support all forms of annotations, including digital annotations of paper documents.

### 3. DIGITALLY AUGMENTED PAPER

The integration of printed and digital materials can be achieved through technologies for encoding active links on paper so that special reader devices can detect and activate these links. A basic linking mechanism is to encode unique IDs as special forms of printed symbols, e.g. standard barcodes. With the appropriate reader devices and server technologies, detection of these symbols can activate a mapping to a digital resource which is then displayed.

A major drawback of such solutions is that they support only a uni-directional mapping between printed and digital material with no way to link back from annotations to the source material. There would therefore be no way to use annotations to support search as used for example in image searches [1]. What is missing is a notion of link context: It is not sufficient to have a unique ID, we also need contextual information to know where the link anchor is located within a document.

For a single page, the  $(x, y)$  coordinate within a page is sufficient to determine location. The mimio Xi and Seiko InkLink products use special pens and high-resolution ultrasonic position capture to create digital versions of handwritten material on whiteboards and A4 sheets of paper, respectively. They rely completely on pen movement and are therefore independent of the material being written on. However, they have the disadvantage of being based around individual pages.

To cater for documents of several pages, we need a mechanism to encode the page number as well as the position within the page. For sets of interrelated documents, we also need a means of identifying the document. The company Anoto licenses a special pattern of tiny visible dots which can encode exactly this information over a large virtual document space. The pattern results in a very light grey background to the printed document which creates almost no visual interference with the document content.

Both Sony Ericsson and Logitech have developed Anoto pens for *information capture*. These are writing devices with a camera alongside the stylus to capture the movement and written information. The actions of the user are stored within the pen and transmitted later. These devices would be suitable for the *authoring* of annotations, providing a convenient means of creating the annotation by drawing and writing on paper. We therefore categorise them as *enhanced writing* devices.

We also want to support *enhanced reading*, whereby users are supported in their reading activity. They should be able to navigate easily within an information space that spans not only different information sources, but also different media. This scenario requires a reader device with that has the basic capability of sensing and immediately transmitting position information. The emphasis here is on *interaction* rather than *capture* and the desire to turn paper into an interactive device.

We are participating in a European project Paper<sup>++</sup> [12] which is investigating concepts and technologies for enhanced reading. Within this project, printing and reader technologies have been developed based on conductive inks that provide a potential low cost solution to the reading of positions on paper.

It is beyond the scope of this paper to discuss the technologies for digitally augmented paper in detail. Further information about the various technologies and projects is given in [10].

We believe that the key to developing powerful cross-media annotation systems lies primarily with the information server rather than the specific client technologies. A suitable framework must be capable of dynamically mapping document positions to information objects and vice versa. It must manage not only the links, but also their semantics and have flexible means of classifying and associating both annotations and domain concepts. In addition, it must be able to deliver personalised digital information on a range of client devices.

### 4. ANNOTATION AND LAYER MODEL

Anchors within a printed document correspond to *active areas* defined in terms of geometrical shapes. We assume the reader device delivers  $(x, y)$  coordinates that specify a position within the page, along with the page ID. Standard barcodes or RFID tags can be used to provide the document ID. From the  $(x, y)$  coordinates, we can determine whether or not the device location lies within an active area of the page and, for example, display an associated annotation.

We may want overlapping active areas (similar to the overlapping links described in [8]). For example, complex images which are annotated may contain sub-images that we also want to be able to annotate. Similarly, it is often the case that we want to annotate individual words or phrases within a piece of text as well as the entire piece of text. To enable this, we have introduced a virtual page layering scheme as illustrated in figure 1.

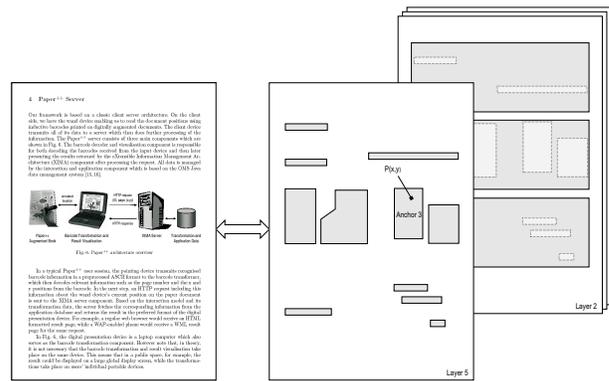
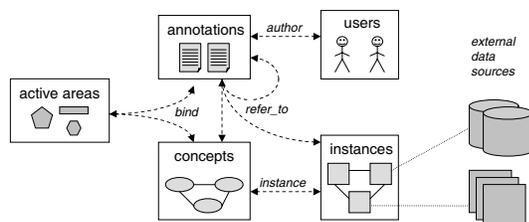


Figure 1: Virtual Page Layers

Each active area is associated with a specific virtual page layer and there can be no overlapping areas within a layer. From the  $(x, y)$  coordinates, we can identify *all* the active areas to which the point belongs across the layer set. If it belongs to more than one active area, these areas must be on different layers and the one from the uppermost layer is selected.

Specific layers may be activated and deactivated, enabling us to generate context dependent results by binding a position to different annotations depending on the current active layer set. A detailed description of the concept and implementation of virtual page layers can be found in [10].

The information model of our server can be divided into five main components as indicated in figure 2. The *active area* objects can be bound either to *annotation* objects or directly to *concept* objects. An annotation object represents the content of an annotation such as a piece of text, an image, a video clip or even an RDF document. Metadata such as source and date of creation may also be included.



**Figure 2: Information Model Overview**

Annotations may refer to instances or to concepts. A *concept* object represents a concept of the domain of discourse and is essentially a metadata (schema) object of an application database. An *instance* is an object of the application database that represents an individual domain entity e.g. a particular mammogram or a particular patient in the case of breast screening. When a person performs the screening of a mammogram, they will annotate a breast outline on the paper form. At the end of the screening process, the information is transferred and processed. The processing of the input data will include the creation of an annotation object with a digital version of the annotated breast outline as content, together with an association to the corresponding mammogram instance giving details such as the place and time it was taken.

A *user* object will manage various forms of information about a user inclusive of preferences, roles and access rights. An author can bind different annotations to the same active area based on user profile. For example, in a training system, active areas might be bound to detailed explanations for trainees, and result summaries for instructors.

Having presented the main ideas behind our annotation framework, we now describe briefly its implementation. In contrast to a number of other annotation architectures, our general approach is based on extending database technologies through the integration of new concepts, rather than on a middleware approach based around existing database technologies. Specifically, we have extended the OMS Java data management framework [4] to support open hypermedia systems by integrating concepts to support both multi-channel access and cross-media linking. Multi-channel access is supported through a clear separation of notions of information, structure and presentation and a general access layer based on XML technologies [5].

Cross-media linking is supported by means of a general link server called *iServer*. Our link model enables anchors to be bound to multiple links and allows for overlapping links and also links over links. This is achieved through a very general link model implemented in OMS Java, the schema of which is shown in figure 3.

The upper portion of the schema describes the general link server in terms of the OM data model [9]. Rectangles represent collections of objects of the same type and ovals associations over collections. By modelling *Links* as a subcollection of *Entities*, we achieve full generality of allowing links over any type of data item, including also links. Further, *Selectors* and *Resources* are included as subcollections of *Entities* to represent any form of media selector and data resource, respectively.

The *iServer* framework supports very flexible link management by introducing user management, layering etc. as core components. Every link, and therefore also every annotation, has an explicit owner and can be selectively shared with other users. Links are always bidirectional and it is possible to find all link anchors for a specific link target (backlinks).

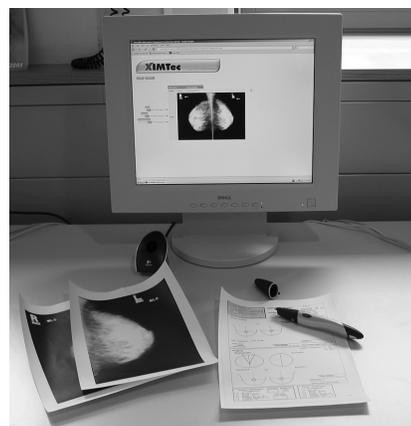
We extended this general schema with a Paper<sup>++</sup> component to integrate digitally augmented paper into the *iServer* framework. In the case of paper, shapes that define active areas act as selectors. The target resources bound to these areas may be any form of media file or paper document, enabling linking between printed documents. By extending the *iServer* schema with further subcollections of *Resources* that represent information objects of a particular application domain, we can link into semantic entities of a domain of discourse rather than simple media files. Details of the Paper<sup>++</sup> component of the model and how activation points on digitally augmented paper can be mapped to application databases are reported in [10].

Using the *iServer* framework, with the Paper<sup>++</sup> extension, we were quickly able to implement an annotation framework for the digital annotation of printed documents. The generality of the link model, inclusive of the extensive link metadata, provided for very flexible referencing schemes that enable annotations to reference not only concepts and instances, but also other annotations.

## 5. MAMMOGRAPHY APPLICATION

To better understand the needs and possible benefits of cross-media annotation, we developed a prototype system to support the mammography screening process. In our prototype shown in figure 4, we investigated the use of the Logitech pen for the authoring of links and content.

Since we did not have access to the actual X-ray images on film, we took paper copies as substitutes allowing us to use the Logitech pen for both the mammograms and the report form. However, we note that the mimio Xi product described in section 3 could be used for annotating actual film, allowing positions to be detected while leaving the film unchanged.



**Figure 4: Mammography Reading Workplace**

For each patient, the reader has multiple sheets of paper, covered with the Anoto pattern: the mammograms of the patient and a screening report form. The report form contains an active area linking to a patient's medical record.

To fill in the report form, the reader marks checkboxes and writes free text in the corresponding areas. After placing the pen in the

