

The Future of Annotation in a Digital (Paper) World

Catherine C. Marshall
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304

*Presented at the 35th Annual GSLIS Clinic: Successes and Failures of Digital Libraries,
University of Illinois at Urbana-Champaign, March 24, 1998*

If order-making in the large is part of the institutional mission of libraries, then order-making in the small – the informal work of annotating and organizing materials collected in service of particular, day-to-day work or pleasure – is part of the business of library patrons. This paper focuses on just such activities, activities that stem from readers' engagements with texts, and possibly with each other, against a backdrop of real-world settings and practices.

I hesitate to call digital library patrons *users*, since that's the word we computer scientists tend to use to hide the characteristics of what we hope is a diverse population.¹ This paper describes three very different communities who might take advantage of new digital resources, including short accounts of technologies I and others have developed to support their activities as they use formal collections and more ad hoc resources. In fact, what I hope to do is to advocate annotation as a key function to what Patrick Bazin (1996) refers to as a reading machine. But first, I'd like to clarify what I mean when I say "annotation."

In Robert McCrum's (1994) account of the annotations Graham Greene's biographers found as they looked through the books in his personal library, he writes:

"Many writers have left much larger collections, but what is different about [Graham] Greene's library is the wealth of personal annotation, reflecting a long and crowded life of writing, politics, travel, and friendship. Scattered along the margins and jotted on the flyleaves and endpapers of these books are thousands of meticulous handwritten notes and comments." (p 46).

Four properties of this account of annotation stand out. First, Graham Greene's annotations are personal, with no expectation of an audience beyond himself. Second, they are literally on the pages of the book, and as such have become part of his library. Third, they've crossed from a

1. Jonathan Grudin (1989) presents a particularly compelling account of why we, as system developers, should be very aware of how we use the term "user."

private space – his library – into a more public space – the hands of his biographers. Finally, they have lasting value. This characterizes a very particular kind of annotation.

My second example of a kind of annotation comes from Vannevar Bush's (1945) prescient description of hypertext:

"The owner of the memex, let us say, is interested in the origin and properties of the bow and arrow... He has dozens of possibly pertinent books and articles in his memex. First he runs through an encyclopedia, finds an interesting but sketchy article, leaves it projected. Next, in a history, he finds another pertinent item, and ties the two together. Thus he goes, building a trail of many items. Occasionally he inserts a comment of his own... Thus he builds a trail of his interest through the maze of materials available to him." (p.107)

So again we're looking at personal annotation, but this time, it's digital and serves to connect documents; the trails Bush describes are not part of the documents themselves. Later in the article, they too will cross into a more public space, in this case, the hands of a friend researching a similar topic. However, by contrast to Greene's annotations, the move of the memex owner's annotations from a wholly private to a shared space is done intentionally. Again, the annotations have value beyond their original purpose.

The variety of types of annotations, and indeed range of interpretations of what an annotation is and how it functions in the world, provoked me to lay out an initial set of dimensions to characterize what I've encountered so far. These dimensions are not intended to suggest dichotomous classifications, but rather to gently investigate why all annotations are not created equal.

The first two, *formal/informal* and *explicit/implicit*, are distinctions of form. Informal annotations, like Graham Greene's jottings in the flyleaves, may be descriptive, but in a digital world, they aren't necessarily computationally tractable. On the other end of the spectrum, metadata created according to a standard, using attribute-value pairs and a source of authority may be computationally tractable, and a good way to promote interoperability; but it may also be costly to create. To reduce the overhead of description, we may use methods of extracting more formal description from informal annotations. The explicitness of an annotation is what allows us, as non-authors, to interpret it. An exclamation point in the margin of a technical manual may be cryptic; a note on the frontispiece of *Portrait of the Artist as a Young Man*, "Most important epiphany, p.47", makes a good deal more sense. Explicitness is thus crucially related to the ultimate intelligibility of an annotation.

The second three, *writing/reading*, *extensive/intensive*, and *permanent/transient*, have to do with the function of the annotation. In much of the literature about readers-as-writers, readers are variously a force that decenters authority, or they may play a far more traditional role as an engaged audience. This tension, explored by Moulthrop (1993), crucially dictates the ultimate value of the annotations in relation to the primary text. If we take extensive reading to be along the lines of what Bush's memex user was doing, and intensive reading to be deep engagement with a single text, as exhibited in Graham Greene's personal library, another characteristic of annotations comes to light: is the annotation across various works, or within them? This dimension follows from the distinction Levy (1997) makes when he describes types of reading and attention. The permanence of the marks really comes to the foreground when we talk about "going digital," since the marks are now readily separable from the document; they are transferable from one digital copy to another; and they can be easily removed. They can now take on a life of their own, or be removed at will.

Finally, the last two dimensions, *published/private* and *institutional/workgroup /individual* have to do with intentional and unintentional movement of annotated documents from one person's hands to another. Let's first take the published/private dimension. Graham Greene's annotations moved seamlessly from being private to being, in effect, published. It is this movement that's of the greatest interest to me as we move to digital media, since this kind of movement is now very clumsy compared to what we do with paper documents. I can mark on a paper I'm reading, and when you ask for it, unless I've written something dreadful in the margins and rush to erase it, you'll be able to see it – and make what you will of it – when I hand you my copy. What will this be like in a digital world, in which we can lift our own responses to a document off with ease before we pass it on? Institutional/workgroup/individual simply refers to the intended audience for the annotations. The original of visions of hypertext, in particular those of Engelbart and Nelson, were fundamentally additive; documents and their commentary make up the docuverse.

To explore some of these dimensions, and to introduce some related technologies, I'll be telling three stories about annotation. The first is about intelligence analysts, and the annotations and order they make in the course of interpreting document collections; the second concerns K-12 teachers and their students, and their use of Web materials in the classroom; finally, the third begins with college students and the marks they make in their textbooks and ends with some implications for future efforts to create a digital library reading machine.

Before I start though, I'd like to expose another, more hidden, agenda for this talk that is very

much in line with the “successes and failures” theme of this workshop. That is: most of the technologies I’m going to talk about are not first generation efforts. They are in each case *simplifications* of earlier technologies that were found to be unworkable, given the constraints of work in the world.

Intelligence analysts and their notes.

The first story is about intelligence analysts and the sense and order they make in the course of interpreting heterogeneous collections of materials. This story begins well over a decade ago, and even before that if I count my involvement with this user community as a technology developer. In 1989, I had the opportunity to conduct a series of work practice studies of intelligence analysts in their offices at various sites around the Washington, DC area.

Much of what I learned about the analysts came out of individual interviews, coupled with observations in and around their workplace. At the time I was engaged in this study, analysts often covered the same "beat" for long periods of time, and had considerable familiarity with specific geo-political regions and topics.

Some of what I observed was, in retrospect, unsurprising, but useful. The analysts use annotations the way most experts do; they are resourceful gatherers of materials from different forms and different places; and they organize working materials in ways suitable for immediate use and for personal archival storage. I will describe each of these annotation-related facets of their activities very briefly.

The analysts engaged in at least three kinds of annotation. The first was a product of an analyst’s engagement with a particular document. Analysts marked on reprints, cables, and other paper documents using highlighters and pens; these markings included marginalia, highlighting, and underlines – in short, much of what one would encounter in the office. Analytic work is crucially integrative; as such, analysts did what they could to capture and explore the relationships among documents, including ordering and reordering the documents in piles according to different criteria (for example, a chronology of action or a chronology of when the documents crossed their desks). Because the analysts used a variety of different on-line resources, their monitors were framed by a clutter of post-its to remind them of the "how to" details.

The documents they used were heterogeneous, both in source and in media. They would readily combine personal materials with workgroup files – some of the analysts in the study had

explicit shared files, others would rely on mediated access to their colleagues personal files –and institutional publications. They would also consult on-line news providers, such as Nexis, Dialog, Comline, and others. Because this study took place in 1989 and early 1990, Internet information resources were not yet a part of standard practice. Despite claims that collaboration was infrequent, the analysts consulted with each other freely, looking for corroboration, missing information, opinions, and so on. It was only when these consultations resulted in a co-authored analytic paper that they were institutionally acknowledged as collaboration.

The analysts used multiple means of organizing the materials that they would gather over the course of an analysis, and over the course of their careers in government service. Most important for this discussion are the transient, visible ways of organizing materials to task exigencies. These organizations were exploratory – for it matters the order in which documents are encountered, and which documents are in spatial juxtaposition – and highly fluid. The organization of materials for a particular analysis did not necessarily reflect a longer-term archiving strategy. Analysts cited simple archiving schemas based on “people, places, and things” or geographical regions as they way they would make their files accessible over the long term.

Other findings about annotations and use of materials were more provocative, especially from my perspective as a system-builder.

First of all, note-taking *per se* was uncommon. Brief annotations and manipulations of the physical documents mediated between reading and writing. An analyst might write "Wrong!" or “don’t believe this” in the margin of a document, but not elaborate on the interpretation until he or she was producing an analytic report. Our previous efforts had resulted in a system, NoteCards (Halasz, Trigg, and Moran, 1987), which assumed a real note-taking model of the sort we all learn in school.

Much of the pre-writing interpretive activity took place using paper and the phone. Cables and retrieval results were printed; borrowed materials were on paper, photographic paper, or involved physical media. Needless to say, this aspect of practice presents a real challenge for those of us interested in supporting a range of interpretive activities on-line.

Although many institutional initiatives have been aimed at automating up-stream aspects of analytic practice – i.e. extracting useful tidbits of information, or visualizing immense document collections – analysts’ sense-making relied crucially on communion with their source materials. That is, much of what they made of materials hinged on reading, skimming, and otherwise

manipulating individual documents. Anecdotally, this communion is illustrated in an analyst's story about "the dog that didn't bark," an important insight he had gained by observing what was *missing* from a document, rather than from what was *in* it. A second example comes from observations of analyst who printed out source materials, marked key passages with a highlighter, then typed them back into her computer. While it's tempting to dismiss this as yet another interoperability problem (why couldn't she just move the materials from one window/system to another?), it is more likely that the act of retyping the content was critically important to gaining purchase on what was said.

Finally, the one finding that to me amplified the need for observation in situ was difference between what I saw and the unified institutional story I encountered. Organizations that spend substantial time reflecting on their own failures and successes, and telling "how we work" stories, have well-crafted narratives describing their day-to-day practice. In this case, as in others, these stories aren't the whole picture.

VIKI (see Figure 1a) is the technology that we ended up developing, a workspace for gathering source documents and recording coarse-grained interpretations of them. This example illustrates four requirements we found central for the workspace. First, the workspace is an

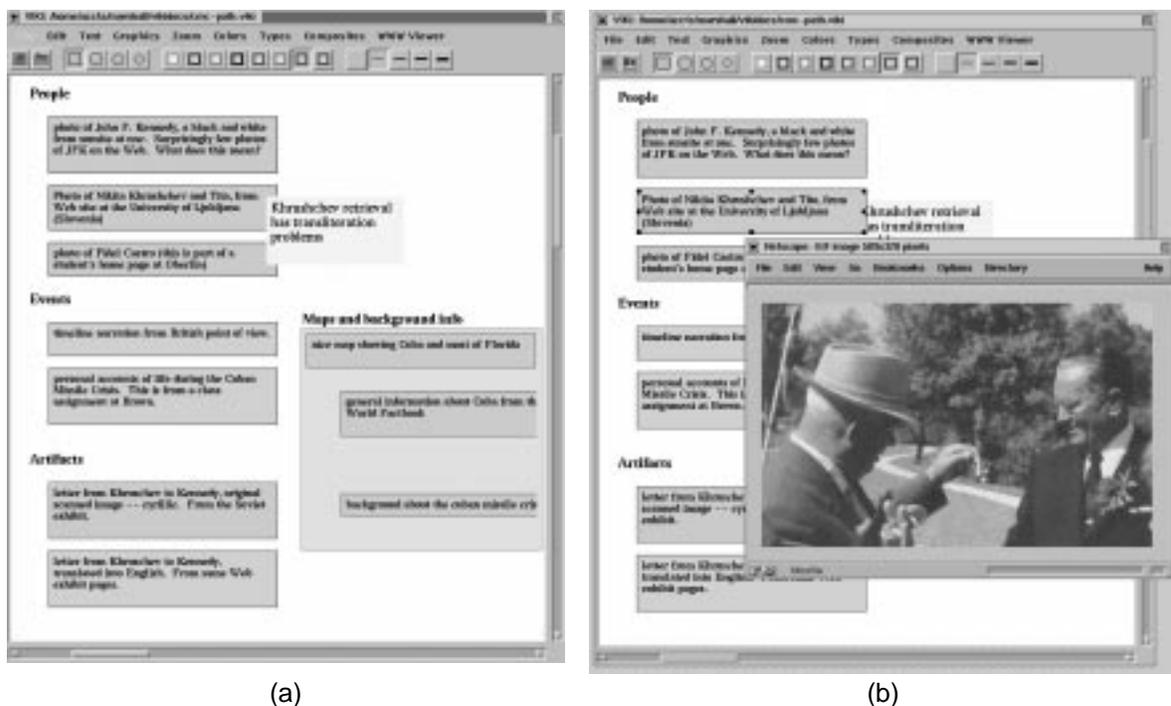


Figure 1. VIKI, a workspace for gathering and interpreting documents. Figure 1a shows the workspace itself; Figure 1b shows a Web browser that has been launched as a result of clicking on a visual reference in Figure 1a.

interpretive infrastructure; it doesn't hold the documents. In this case, the documents themselves are Web pages, and are simply referred to by URLs. This is not a radical idea given the way the Web implements URLs, but in pre-Web days, this kind of openness was considered an important – but not obvious – requirement.

Figure 1. VIKI, a workspace for gathering and interpreting documents. Figure 1a shows the workspace itself; Figure 1b shows a Web browser that has been launched as a result of clicking on a visual reference in Figure 1a.

The second aspect of the system that I'll note here is the ability to tailor a reduced representation of individual documents. In other words, each object in the interface represents a document, and I as a user can dictate what form that should take – a title, an abstract, an automatic summary, and so on. Although this version of the system doesn't have the capacity, in cases in which document types or genres are visually distinctive, a thumbnail is the ideal reduced representation.

Third, we use hierarchy to tame complexity. What we had in mind is a flexible means of either shifting from task to task, or task to subtask, while maintaining context. How is this realized? By providing subspaces, so the main workspace can be subdivided. The references to individual documents can appear in as many subspaces as need be.

The fourth aspect of this system that I want to stress in this short account is manipulability, and the ability to record ad hoc, partial, or ambiguous interpretations: it is what we most often lose when we go digital, and what I consider to be the truly annotative aspect of the system. I can change a document to red to signal to myself that it's important, or that I need to come back to it, or that it's about Greece. It is these lightweight classifications that we need to reclaim in our digital library reading machines.

Information Triage: an experiment

One important lesson we learned in this particular development effort is that work practices are dynamic, shifting with social changes. The study I described began before the fall of the Berlin Wall. By the time we'd gone through a couple of different prototypes and had reached the system state illustrated by Figure 1, major changes had taken place in our users' work environment.

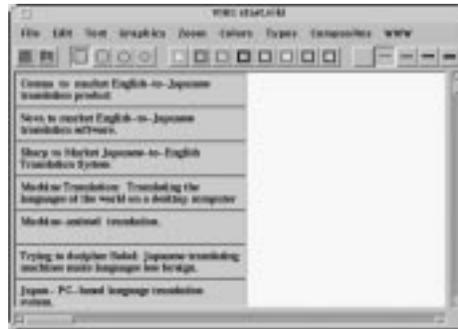


Figure 2. The starting out configuration of the VIKI workspace for the information triage

Although the analysts still performed long-term studies of technology, politics, or events in world regions, the institutional emphasis had shifted to shorter-term – oftentimes daily – results. Would our approach still make sense given this shift?

What we did to understand the effects of our technology on a very short-term analysis task is to conduct a controlled experiment, which was based on a real analysis². Fifteen subjects were given a small corpus of relevant documents (75 in all), and asked to perform an open-ended task over the course of 45 minutes. The documents were the same set that had been collected from an information service for the real task, and were incomplete, contradictory, and sometimes redundant. Some documents were frustratingly general; others were brief, and highly specific. One-third of our subjects performed the analysis using paper; the second third used VIKI without the subspace mechanism; and the third used a complete version of VIKI that allowed them to create and populate subspaces. Both versions of VIKI had a multi-term search capability, which turned out to be very popular for winnowing down the document set to specific documents of interest.

Figure 2 shows what the space looked like at the outset. Each purple rectangle represents a single article. The title of the article is shown on the rectangle. The subjects using paper were given a comparable stack of printouts of each article, and some appropriate office supplies like highlighters, pens, post-its, paper clips, and a stapler.³

Three results stood out:

-
2. A complete account of the experiment and its results is presented in (Marshall and Shipman, 1997).
 3. In some ways, the reduced document representation that VIKI showed, the titles, was not entirely comparable to the paper condition, in which subjects always saw the entire first page of the document.

First, in many ways, the kinds of things people did to cope with the surfeit of relevant documents were remarkably similar. In all cases, people sorted the articles into rough categories; these categories shifted as they began to understand the nature of the corpus and the nature of the task. They had remarkably little patience for the general articles, and figured out ways to get them “out of the way.” Figure 3 shows a side-by-side comparison of two subjects’ results. In Figure 3a, the categories are implemented as piles and “out of the way” meant on the floor. In Figure 3b, the categories are implemented as subspaces; documents of deemed of little value for the analysis are left in their original positions.

Second, the tool influenced the way people thought about the task. This variance is most evident in how they responded to the question, “What would you do if you had more time?” The subjects who completed the task using paper generally focused on reading; the subjects who tackled the analysis using the simplified version of VIKI without subspaces talked about organizing; and finally, the subjects who used the complete version of VIKI spoke even more intensely about their desire to create order. The following three responses to the “more time” question exemplify this effect:

“[I would] Read the info that I selected as critical more carefully and perhaps highlight some important text for my boss to help support my decision.” – a subject from the paper condition

“[I’d do a] better job of re-organizing the documents: I spent my time coming up

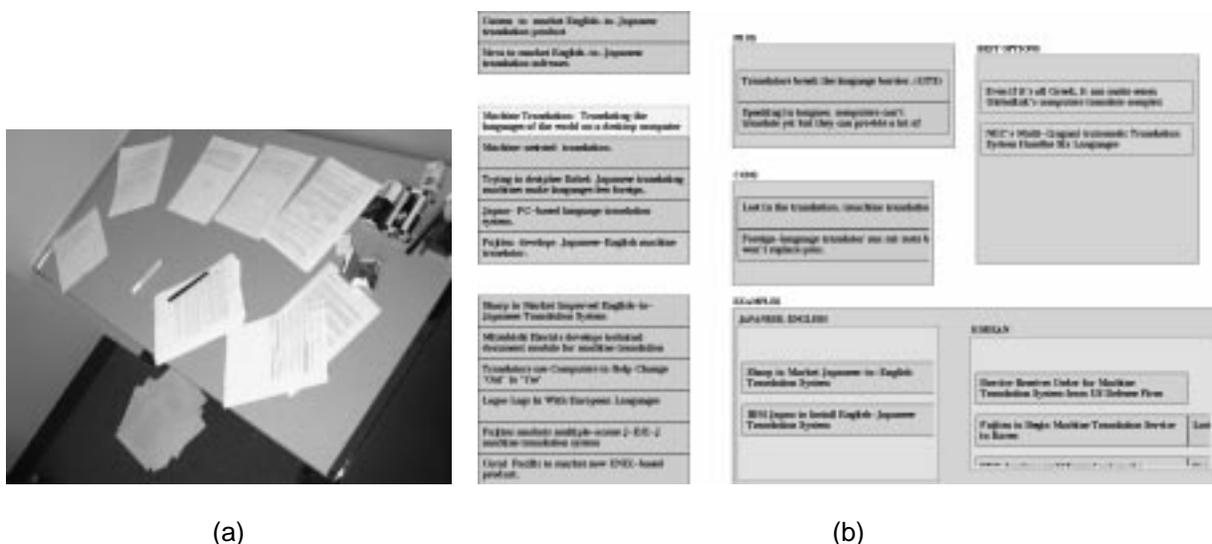


Figure 3. A comparison of two subjects’ results from the information triage experiment. Figure 3a shows a subject performing the analysis using paper copies of the article; Figure 3b shows the results of a subject’s analysis of the materials using the VIKI workspace.

with a recommendation, not organizing the documents.” – a subject from the VIKI-without-subspaces condition

“I would organize each big collection into smaller collections and possibly change some of the names... Also I would look a little more carefully at some of the articles as some might be misplaced.” – a subject from the VIKI-with-subspaces condition

Finally, the question that lingered in my mind after the experiment was over was what would have happened if we had just allowed the subjects to do their own research on Web? A well-constructed boolean query that had been used to gather the source documents from an information service turned up 6000 hits on the Web; sampling the hits showed many of them to be relevant. The type of information triage we investigated in our controlled experiment is bound to be a potential activity of digital library patrons when they find themselves gathering materials to answer open-ended questions.

K-12 teachers and students make their way through the Web

The second story I plan to tell is one of annotation as a means of ordering “found” Web materials for presentation in a classroom setting. This story begins at the close of 1994, with a DARPA-sponsored project in the CAETI (Computer-Aided Education and Training Initiative) Program. What we set out to do in this project was to find a way to take advantage of what the Web can bring to the classroom in terms of access to information, while still acknowledging that the Web is not a digital library – there are all kinds of materials on the Web that for one reason or another are over the kids’ heads or absolutely inappropriate for the classroom – and that the school setting introduces a variety of challenging technological and social constraints.

While we were in the early phases of our effort, we had the opportunity to observe classes of sixth graders from a local middle school using a standard browser to explore the Web. The sessions, which were part of a community enrichment program, took place in a Texas A&M engineering computer lab, so that each student who participated in the session could sit in front of his or her own PC. We observed several phenomena that are important here.

Paths were a natural way for students to approach the Web. If we listened to them, we heard things like, “How did you do that?” and “How did you get there?” Students helped each other navigate in a very literal sense: they would help another student retrace their own steps to get to a particular page on a Web site.

Some students would get stuck. They would reach a page that they didn't expect to see, or one that had no links out, and instead of using the "back" button to retrace their steps (admittedly we all know this to be a fallible means of navigation, given the models of backtracking in Web browsers), they would simply give up and wait for help. Intervention was necessary to put them back on track.

The unfocused exploration, coupled with the fact that learning is an essentially social activity, proved to be interesting as well. It became apparent that the classes near the end of the day were affected by the discoveries of the students in the earlier sessions. During a morning session, one student discovered MTV's *Beavis and Butthead* site. Soon other students were following his path there, and apparently by the end of the day, students had figured out how to pass the URL to one another to navigate there directly. There are two conclusions we came to as a result of this observation: one is that learning has a certain collective quality to it. But, more importantly, having a goal and guide to focus exploration is indeed necessary.

We developed a system called Walden's Paths to work in a K-12 classroom environment. The basic architecture of the system is constrained by existing technology, network bandwidth, and the exigencies of teaching and the classroom. Thus we designed the system as a proxy that would mediate between existing Web servers and clients, so no new software needs to be purchased, and upgrades can be performed as needed. We also paid attention to caching strategies, since network bandwidth to classrooms is sometimes lower than need be. A Path Authoring Tool allows teachers and students to assemble and annotate Web pages to form paths. The paths are linear in recognition of the time consuming nature of developing a rhetorical structure for a non-linear path.

Figures 4a and 4b show Web pages served by the path server. Note that the controls for moving along a path are readily available at the top of every page on the path. The arrows move a student forward or backward along the path; the numbers both show the number of stops that the tour has, and allow the student to jump to a specific one. Below the controls are annotations added by the path author; in Figure 4a, the path author has added some material about Kennedy's victory in 1960 to a digital image gathered from the Web. Figure 4b illustrates the approach that Walden's Paths takes to student explorations "off-the-path." A control is prepended to take the student back to his or her jumping off point.

My colleagues Frank Shipman and Rick Furuta and their graduate students are continuing



(a)



(b)

Figure 4. Web pages served by the Walden's Paths server. Figure 4a shows a page that is "on the path," Figure 4b shows a page that is "off the path."

this work. The Walden's Path system has been used in classrooms for Department of Defense dependents in Germany and Italy (Shipman et al., 1998).

College students engaged with their texts

My third story is about annotation as a reflection of students' engagement with their course textbooks. What I plan to do in this story is first to take a close look at annotation of individual texts to look at annotation as an individual practice, then move from there to suggest a movement toward an ecology of annotation.

Personal annotations in books have been the inspiration for hypertext systems builders (and hypertext advocates) for some time now, especially for those working in educational settings.⁴ In his Hypertext '87 Keynote address, Andreas van Dam (1988) spoke about the roots of his early hypertext system, FRESS:

"The reason I encouraged such annotations [in FRESS] was that I remembered that when I was in college with Ted [Nelson], I would always grab the dirtiest copy of a book from the library, rather than the cleanest one, because the dirtiest ones had the most marginalia, which I found helpful." (pp.891-2).

4. See, for example, Landow's (1992) essay. Such annotations have also stirred considerable interest in the community of people developing computational support for on-line technical manuals and the like.

If such an assumption is true – that annotations have value beyond the immediate – and that these annotations are created through a particular set of practices and activities, I began to wonder how annotation would take place in a digital library.

Some of these musing were simply about the marks themselves and the means of making them. As it stands, annotating digital materials is not a straightforward activity. We have neither the practices nor the tools for fluidly marking on digital materials in all the ways we mark on paper. Yet we often desire to do so. People print out documents to mark on them (O’Hara and Sellen, 1997).

Furthermore, as we all know as annotators ourselves, the functions of the markings people make as they read are not so simple. Yes, some of them are the kind of useful commentary van Dam is addressing, but other functions are evident from even the most casual look at annotations in action. So then, putting on my developer’s hat, I felt obliged to ask, “How will the many functions of annotations inform implementation?”

Given that these annotations begin their life as a personal form, and not as public commentary, we can move into the realm of the boundaries between private and public forms – boundaries that are often far more explicit and pronounced in the digital world than they are with paper documents. Will the move to digital materials make this seamless transition from a personal form to a public, often anonymous, form impossible? Coupled with this question is the very real question of whether a typical annotation has any lasting value relative to the potential permanence of the work itself. If it doesn’t, then the transition between private and public is not particularly of interest to us.

Most generally then, I would like to pose the question: What does the activity of annotation on paper imply about reading and writing in the digital library?

To help answer some of these questions, I used the crowded textbook section of a large university’s bookstore as a source of both a community of annotators, and access to a large collection of annotated course materials. This bookstore, as well as many others, buys back used textbooks no matter how heavily annotated they are, as long as the books’ bindings and pages are intact. I’ve spent the first week or so of classes in the bookstore over the course of 4 terms to get a good sampling of both students and books.⁵ To date, I’ve examined about 410 used books,

5. In particular, since annotation is a learned practice, incoming freshmen and their used texts are quite different from upperclassmen and their used texts.

representing 39 titles (same edition) in 21 general course areas. What the bookstore gives me is access to comparable copies of the same edition, marked up under similar circumstances.

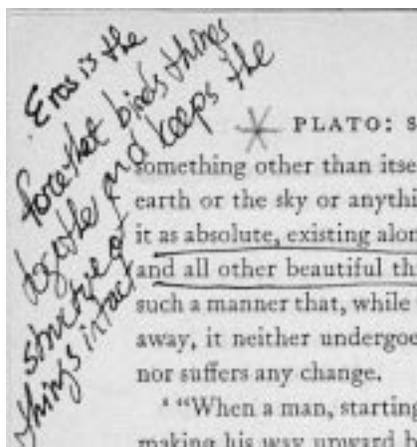
This setting, the university bookstore, has enabled me to not only look at the annotations in the textbooks themselves, but also to observe the students choosing their books and talking among themselves about, among other things, strategies for buying used books (Marshall, 1997). I've also conducted open-ended informal interviews of the textbook buyers, often using the annotations of other students to provoke comments and reactions, and to help the students describe their own annotation practices. Most recently, I've performed a detailed comparison of copies of a single edition to better understand how annotations made by many different members of a community add up (Marshall, 1998).

The form of these markings is, as one would expect, incredibly fluid. Annotations are made using all manner and all colors of specialized markers, pens, pencils, and – most generally – anything that can leave a record on paper. Likewise, any markable area of a book might have annotations on it.

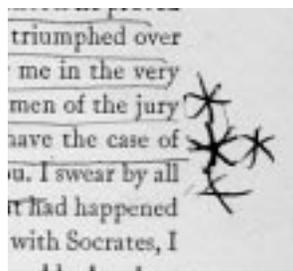
But what is more interesting is that there are some notable strategies people bring to bear in their annotations. First off, most often annotators use the writing implement that is “to hand.” Evidence for this is in the correlation between penned marginalia and underlining; I also found notes written in highlighter (an awkward writing implement at best – these notes are usually quite short). Second, there were a small number of complex (but implicit) coding schemes that annotators had developed. The fact that the number was small is probably of greater interest than that they existed at all; it is a great temptation to propose schemes for digital tools in which pen color means something, and is used computationally in some way. It was rare to find one of these schemes that lasted throughout a textbook. Finally, it seems that form follows textbook genre and expected disciplinary practices. This observation should come as no surprise, since works in different disciplines are “read” differently.⁶

An example of a technology that might be a good basis for Patrick Bazin's digital library reading machine has been developed at Fuji Xerox's Palo Alto Laboratory. It is called XLibris (Schilit, Price, and Golovchinsky, 1998), and features a pen-based interface, a portrait page orientation, and is about the size of a laptop computer. This device is document-centered, rather

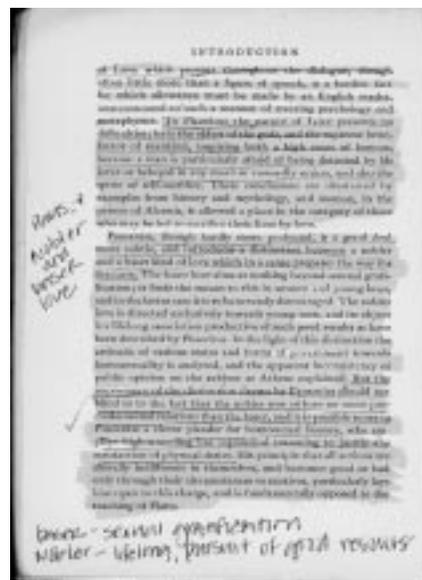
6. For example, heavy use of a yellow highlighter in the first few chapters of a math textbook does not bode well for the student; by convention, math textbooks begin with review, and memorization is usually less important than the ability to reason with and from the material.



(a)



(b)



(c)

Figure 5. Contrasting three different functions of annotations. Figure 5a shows an interpretation written in the margins; Figure 5b shows emphasis markings; and Figure 5c shows dense, non-interpretive highlighting.

than application-centered, and supports the fluid kinds of markings I have encountered in paper books.

The function of these annotations appears to vary a great deal, from clearly interpretive – marginalia that adds to the source text, as in Figure 5a – to asterisks that signal importance – but not why the passage is important, as in Figure 5b – to page after page of highlighted text, as if the reader is marking his or her attendance to difficult or particularly dense writing, as in Figure 5c. From an implementor’s point of view, it is clear that at least some of these markings should be considered as transient evidence of a reader’s engagement with the text.

What is more at issue here than the form annotations take and the functions they serve is their ultimate value. As you might guess, observations bear out all intuitions that some of these annotations are valuable, and others merely annoying. Some students looked for books that were as pristine as a used book ever is, and others took up strategies like the one I heard discussed, “Look for writing in the margins and no highlighting. Sentences not just phrases.” There was clear evidence that some found all annotations distracting, and other students tolerated some kinds of writing, say yellow highlighter, more than others, for example, black ballpoint underlines.

It is also clear that whatever we conceive of doing with these annotations in a digital world must take strict account of the fact that they are a private form of writing, made public only in though assumptions of anonymity. Expectations of privacy manifested themselves in telling ways: stray, signed credit card slips tucked in between the pages and names and social security numbers written inside front covers.

Taking privacy into account, and considering the kinds of marks that were the most common throughout the textbook sample – highlights, and emphasis marks like stars and asterisks – I tried to imagine what sorts of things one could do with these annotations in a digital library environment. This environment might be a place in which one could computationally harvest transient annotations (given a reading machine) in a wholly anonymous and transparent way. What sort of consensus to readers reach? How can it be used in ways that respect both a reader's sense of order and an author's sense of the original structure of the work?

To answer this question for myself, I performed a study (Marshall, 1998) in which I did a sentence-by-sentence analysis of 7 copies of a particular textbook, a computer science textbook in which the presentation was in a narrative form. The results, although preliminary, were convincing. Readers achieve some level of consensus about where the key passages are in each chapter of the book. These passages aren't necessarily the rhetorically predictable ones (the topic sentences at the beginning of sections, or even the topic sentences of paragraphs), but rather occur in the midst of loci of annotative activity and do have some significance.

Where to next?

What might we do if we took the desire for a digital library reading machine seriously?

First, there are some serious questions to answer about how and where marking should take place. If it is to take place digitally, how can we truly capture the fluidity of form we see on paper? Should we posit a digital library reading machine that's stylus-driven and document-centered like the XLibris device?

Given the right device for personal annotation, we've only answered part of the question. Annotation spans a huge range of activities, activities that may include proofreading and writing commentary to an audience. Can all these forms of annotative activity be supported by a single type of reading machine? Can they be subsumed by a general architecture? Phelps and Wilensky's (1997) work on Multivalent Documents as the basis for layering on annotations begins to

investigate some of these questions. The older ComMentor facility at Stanford, which was based on the NCSA Mosaic annotation function, also assumed that indeed a single general architecture for annotation is an appropriate goal (Roescheisen et al., 1994).

From the results at the bookstore, it looks like it's important for the reading machine to support non-interpretive marking as well as interpretive markings. Given that one might feed into a consensus mechanism like the one I've been discussing, and the other is probably not useful beyond the current reading, is there a way to tell the difference between the two?

And what of interpretive markings? Can we in some way assess their value and intelligibility? How do they function as shared forms? Their intelligibility is always at stake, yet it is clear that sharing interpretive annotations is one of the benefits of working in a mixed digital-physical environment, in which at least some document surrogate exists on-line.

Finally, and most essentially, how can we move toward smooth integration of annotation with the various kinds of reading (intensive, extensive, hyperextensive) we see today? In an intensive reading situation, annotation must represent a deep, unselfconscious engagement with the text. In an extensive reading situation, documents or their surrogates must be readily manipulable and easy to juxtapose or informally link. In a hyperextensive reading situation, annotation can and should be a means of easing the fragmentation.

Finally, considering annotation in the large, how can we support the seamless kinds of transitions between private and anonymous public annotations that we see in the bookstore? How can they be used collectively to form an ecology as we begin to understand their status and value? Most importantly, how can we encourage the establishment and growth of these ecologies, without ignoring questions of scope and community?

The development of the reading machine of tomorrow, one that acknowledges the depth and variability of the reader's engagement with the texts in a mixed digital-physical library, is indeed a tall order.

References

Bazin, Patrick. (1996). *Toward Metareading*. In G. Nunberg (Ed.), *The Future of the Book* (pp. 153-168). Berkeley, CA: University of California Press.

Bush, Vannevar. (1945). As We May Think. Atlantic Monthly, Aug. 1945, 101-108.

Furuta, Richard, Shipman, Frank M. III, Marshall, Catherine C., Brenner, Donald, and Hsieh, Hao-wei. (1997). Hypertext Paths and the World-Wide Web: Experiences with Walden's Paths. In Proceedings of Hypertext '97 (pp. 167-176). New York: ACM Press.

Grudin, Jonathan. (1989). The case against user interface consistency. Communications of the ACM, 32(10), 1164-1173.

Halasz, Frank G., Moran, Thomas P., and Trigg, Randall H. (1987). NoteCards in a Nutshell. In Proceedings of the ACM CHI + GI Conference (pp. 45-52). New York: ACM Press.

Landow, George P. (1992). Hypertext: The Convergence of Contemporary Critical Theory and Technology. Baltimore, MD: Johns Hopkins University Press.

Levy, David M. (1995). Cataloging in the Digital Order. In Proceedings of Digital Libraries '95 (pp. 31-37). College Station, TX: Texas A&M University, Center for the Study of Digital Libraries.

Levy, David M. (1997). I read the news today oh boy: reading and attention in the digital library. In Proceedings of Digital Libraries '97 (pp. 202-211). New York: ACM Press.

Marshall, Catherine C. (1990). Work Practice Study: Analysts and Notetaking. Technical Report, Xerox Palo Alto Research Center.

Marshall, Catherine C. (1997). Annotation: from paper books to the digital library. In Proceedings of Digital Libraries '97 (pp. 131-140). New York: ACM Press.

Marshall, Catherine C. (1998). Toward an ecology of hypertext annotation. To appear in Proceedings of Hypertext '98. New York: ACM Press.

Marshall, Catherine C. and Shipman, Frank M. III. (1995). Spatial Hypertext: Designing for Change. Communications of the ACM, 38(8), 88-97.

Marshall Catherine C. and Shipman, Frank M. III. (1997). Spatial Hypertext and the Practice of Information Triage. In Proceedings of Hypertext '97 (pp. 124-133). New York: ACM Press.

McCrum, Robert. (1994). A Life in the Margins. New Yorker, April 11, 46-55.

Moulthrop, Stuart. (1993). you say you want a revolution: hypertext and the laws of media. In E. Amiran and J. Unsworth (Eds.), Essays in Postmodern Culture (pp. 69-97). New York: Oxford University Press.

O'Hara, Kenton and Sellen, Abigail. (1997). A Comparison of Reading Paper and On-Line Documents. In Proceedings of CHI'97 (pp. 335-342). New York: ACM Press.

Phelps, Tom A. and Wilensky, Robert. (1997). Multivalent Annotations. In Proceedings of the First European Conference on Research and Advanced Technology for Digital Libraries (Papers presented 1-3 September 1997, Pisa, Italy).

Roescheisen, M., Mogensen, C., and Winograd, T. (1994). Shared Web Annotations as a Platform for Third-Party Value-Added Information Providers: Architecture, Protocols, and Usage Examples. Technical Report STAN-CS-TR-97-1582, Stanford Integrated Digital Library Project, Computer Science Dept., Stanford University. November 1994, updated April 1995.

Schilit, Bill, Price, Morgan, and Golovchinsky, Gene. (1998). The Digital Library Information Appliance. To appear in Proceedings of ACM Digital Libraries '98. New York: ACM Press.

Shipman, Frank M. III, Furuta, Richard, Brenner, Donald, Chung, Chung-Chi, and Hsieh, Hao-wei. (1998). Using Paths in the Classroom: Experiences and Adaptations. To appear in Proceedings of Hypertext '98. New York: ACM Press.

van Dam, Andreas. (1988). Hypertext '87 Keynote Address. Communications of the ACM, 31(7), 887-895.