Reading and Interactivity in the Digital Library: Creating an experience that transcends paper

Catherine C. Marshall
Microsoft Corporation

Two articles in a recent issue of the *Communications of the ACM* put forth somewhat alarming claims about the role of the Internet in our lives. Columnist Peter Neumann observed that "we seem to have evolved into a mentality of 'If it's not on the Internet, it doesn't exist.'" Neumann also noted that the Web duplicates very little of the material we might find in traditional institutions like public libraries [18]. In the same issue, Graham and Mataxas argued that although most of today's students do research exclusively on the Internet, very few of them double-check the information that they find or locate multiple independent sources to corroborate what they've learned [6].

These observations underscore two central themes of digital literacy. First, and most fundamentally, much of what we read when we research a topic now arrives in digital form; how can we ensure that people are able to <u>read</u> it when and where they retrieve it? The second, and more provocative of the two questions is: How will people <u>interact</u> and work with these digital materials in a way that takes advantage of the power of computers? The Graham and Mataxas article implies that there is a need to go beyond simply reading a single digital resource; new forms of critical thinking must be taught and the tools to support critical thinking must be readily available.

Let's look at reading first before we move into interactivity and functionality that transcends paper. Although the answer to the question of how people will read digital materials seems simple, digital reading practices are evolving rapidly. These changes should be no surprise; reading practices have always varied according to the genre of the materials, the reader's purpose, and the introduction of new technologies [12].

Four ways that reading is changing

The reflections of experts in the field, combined with my own empirical observations over the last decade, reveal at least four profound ways that reading is changing. Some of the changes stem from the availability of an overwhelming number of digital resources; others arise from the altered circumstances of reading, for example the changing role of paper. Still others have to do with the introduction of new reading technologies like e-books, tablet computers, and accompanying advances in digital typography.

Reading practices themselves are evolving

First, many readers have developed new strategies for handling the volume of digital information that is now a few keystrokes away. Unfortunately, some of these strategies are not entirely positive. In the digital age, our attention spans are ever shortening and our reading is becoming increasingly shallow; we rely on encountering the same material – or portions of the same material – multiple times in multiple circumstances and forms. According to David Levy:

> "... Changes in the technologies and the character of modern life may be putting an end to reading in depth. That's the fear, at any rate, in some

quarters... It isn't that the book has gone away, but rather that the cultural conditions for [deep] reading... are fast disappearing." (p. 109) [9]

This change is evident in the way people talk about how they read, and indeed, in what counts as reading. In several recent studies, we have heard students and office workers characterizing their reading in terms of just-in-time skimming or scanning material very quickly in an effort to find the most salient portion of a longer document. Study participants have described taking a quick look through newspaper headlines and possibly returning to the periodical on multiple occasions, to read for more depth as time allows. Furthermore, we have found that people may print things or save things in lieu of reading them immediately; if they know where something is, and they know they can get back to it, they'll hold off on reading it until they need to.

The role of paper is changing

Second, the role of paper is changing. It is becoming less common to reflexively print electronic documents just to read them. About five years ago, in an article in *Online* magazine, Walt Crawford wrote:

> "Reading from digital devices ... suffers in several areas - among them light, resolution, speed, and impact on the reader - and there has been essentially no improvement in any of these areas in the last five years... It's just too hard to read from a computer, and it doesn't seem likely to get a lot easier." [2]

In this article, Crawford argued that physical collections still matter and that because people would invariably print digital materials to read them, digital collections may create a dystopian situation in which the same work is printed many times rather than read and returned to the collection. While there may be some of this going on, it instead appears that people are choosing between paper books and electronic texts according to their purpose, and are not printing unnecessarily.

Furthermore, as Sellen and Harper point out in their book, *The Myth of the Paperless Office*, paper offers clear advantages over digital technologies for certain cognitive tasks [21]. Computer scientists are beginning to recognize that the presence of paper in offices does not necessary mean that computer technology has not succeeded, which until recently has been the consistent interpretation of the paperless office's failure to materialize. There is now widespread recognition that there are ways in which paper is simply a convenient and malleable document technology.

For example, a paper document set on top of a colleague's keyboard serves as a tangible physical reminder that the document should be read. Sharing a printed document often fills many requirements including informally limiting the scope of the sharing and indicating the document's importance. Thus, a central theme across many of our recent studies is that paper is diminishing in its role as the sole viable reading technology, but provides vital support for other document-related activities.

People have begun to read on their screens

Third, people have begun to read on their screens. This is a new trend; a decade ago, there were clear results that people preferred to read on paper, even if their performance

didn't suffer [3]. Five years ago, it still seemed that people did not like to read on the screen and that they would come up with a variety of strategies to avoid it.

In a study of second and third year law students we performed almost five years ago, one student reported that although he searched for legal cases in the Westlaw database from the networked PC in his dorm room, when he found a case that he needed for his class work, he wrote down the case citation, and went over to the law library to print it so he could read it carefully. The law students only seemed to read the results of their last-minute research on the screen. This was no surprise at the time; we expected the law students to print their source materials, and perhaps even to go out of their way to do so.

But all that is changing; people seem to be increasingly willing to read on the screen. In a recent study of how people read a variety of paper and electronic periodicals, we found that twelve out of twenty of the study's participants read some sort of periodical on the screen [15]. Many subscribed to the *New York Times* daily headline service and followed interesting headlines to the full online stories. Others regularly read trade magazines electronically.

To confound matters further, one important over-arching finding of a recent internal field study of office workers and their reading practices is that reading is not a well-defined activity when it is taken to the screen. When office workers read on paper, they have a clear sense of what reading is: it is a focused engagement with a single document, distinguished (and distinguishable) from related activities like writing, filing, or talking on the telephone. On the other hand, when office workers use electronic documents on the screen, activities blur, and much of their interaction with the documents falls under the general rubric of reading. For example, skimming subject lines, responding to, and categorizing email all count as reading.

Mobile devices provide a better form factor for reading

A technology factor motivating changes in reading is the widespread adoption of mobile computing devices: more and more readers are carrying screens around with them that are potential venues for reading. These screens can represent form factors as small Pocket PCs and Palm Pilots, as specialized as e-book hardware, or as powerful as general purpose computers with high resolution color displays like laptops and Tablet PCs. Furthermore, user interaction on these devices is becoming much more suitable for reading. Many employ styluses or pens for interacting with the screen, have hardware buttons that can be used for reading operations like page turning, and are sufficiently lightweight to make them convenient to carry around much of the work or school day.

The use of these mobile device technologies, especially those with smaller screens to do any sort of non-recreational reading, is somewhat surprising in the light of earlier findings. In the study reported in [16], the University of Virginia's EText Center had distributed Pocket PCs to students in two different courses, an undergraduate class on the Salem Witch Trials and a graduate English Literature course. I went out to the University of Virginia expecting to see the Pocket PCs tucked into desk drawers in dorm rooms or used to hold a small stash of illicit MP3s. Instead, almost a quarter of the students were using the handhelds to read and refer to secondary materials for their classes.

Furthermore, supporting technologies are improving. ClearType, a subpixel rendering technique may be used in conjunction with adaptive layout to create a more familiar, more booklike reading experience [1]. Later I'll discuss other technologies that support reading and some of the trade-offs they imply. But what I mean to emphasize here is that technologists are paying attention to reading as an activity that very well might take place on a computer.

The rise and fall of electronic books

E-books were, in general, hardware and software specifically intended for reading; published content – novels, business books, magazines, and newpapers – could be purchased in electronic form. E-book research projects were aimed at different kinds of reading material: legal resources, textbooks, office documents, and other everyday documents. Although e-books were not widely adopted, they give us a foil for looking at reading practice and provide a way for us to explore some of the requirements and tradeoffs involved with reading on a mobile device. In other words, e-books were just good enough to put into readers' hands and see what they did.

So why did e-books fail to be adopted?

An obvious reason that e-books stumbled stems from a nearly universal anxiety about the future of the book. E-books proposed to replace an important cultural artifact with an uncertain technology. Much has been written about this topic in scholarly and popular publications, so the argument can be put aside after acknowledging its significant emotional and intellectual force. Instead, I will focus on more pragmatic concerns about the adoption of e-books.

One real issue with e-books – and one of the leitmotifs of this paper – is that the original round of e-books represented a poor cost-benefit situation for both readers and publishers. This is an easy mistake to make when one is developing computer technology: technology is often developed with an implicit perspective on its use, a vague picture of who the user is, and who ultimately benefits from technology deployment and adoption [7].

It is common to adopt the perspective of the end user, in this case the reader, to evaluate new information technologies. Assuming this perspective, it is easy to see how critiques akin to this one arise:

> "Publishers need to add value. Current onscreen magazine systems ... simply load a bulky software program and copy-protected PDFs of the magazines onto your system. There's some rudimentary searching, zooming, and annotating, but that's all... There's ... no desire to do anything but replicate a flat print publication." [8]

The journalist's point is well made. Why read an electronic book if there's no advantage over paper books? Certainly the reader has lost something in the transition to a digital reading experience; what has he gained in return?

Different perspectives, like that of the publisher, reveal other cost-benefit conundrums. Implicit in the quoted critique is the idea that publishers are benefiting from e-books and not perceiving the needs and desires of readers. But this analysis falls short as well.

Although e-books circumvent some of the problems intrinsic to the printing and distribution of paper materials, they also demand that publishers disrupt the processes they already have in place – often based on high-end publication systems – to create new forms such as Open E-Book mark-up, which is HTML/XML based. Because there is no direct path from these high-end publishing systems to standardized markup languages, publishers and other electronic publication projects are often forced to have the material rekeyed and tagged. It difficult to resolve the cost/benefit trade-offs in a manner that is sensible for everyone in the picture.

There were also pragmatic concerns that stood in the way of e-book adoption. Two illustrative issues demonstrate why developers must pay attention to particular use settings and work practices to design technologies like this.

One of these issues stems from the movement of physical and digital materials in the world and the way e-books fit into public institutions like the library. I talked to the head librarian in a public library who decided to institute an e-book lending program in her institution. Because the public library's mission is to serve all segments of a community, not just those library patrons who own laptop computers, she decided to purchase e-book hardware (Gemstar's REB 1100s). These devices could be checked out with pre-loaded content, in the same way patrons would check out physical media like books, video tapes, and CDs. Right away, the head librarian had a problem: because the device required an analog phone line to download e-book content, and the library had upgraded to a digital phone system, the library had to pay $450 to re-install an analog line. Second, because the downloading process was so cumbersome, the librarians had to second-guess their patrons' reading tastes. This experience was not uncommon in the public libraries that decided to experiment with e-books; McKnight and Dearnley report similar findings in a formal survey [17]. They also discuss the reactions of public library patrons to e-books, which is relevant to this discussion as well.

Another practice-related issue is very pragmatic indeed. A purpose-built e-book device, one that is dedicated to reading, does not address the fact that many kinds of non-recreational reading are interleaved with related activities like writing or collaboration. Reading does not always exist in splendid isolation. Even the most dedicated technophiles tend to balk at the idea of carrying separate hardware for reading along with whatever other mobile communication/computing devices they normally use.

In other words, reading is a complicated affair made more difficult to understand by its basic invisibility.

Creating a paper-like reading experience

To promote the transition to reading on the screen, it is vital to make the experience as good as the experience of reading on paper. What are some of the most essential elements of reading on paper? First, reading is an incredibly mobile activity. Second, reading on paper is a material activity, involving physical artifacts that afford certain central kinds of feedback and experience. Third, reading on paper is more interactive than we'd think at first blush. Finally, there's sharing: in many situations, purposeful reading is done in a social setting. Good design can bring each of these paper-like advantages to reading on the screen.

Mobility

When we talk about mobility, we often picture travel. And indeed, people do read when they're en route from here to there.

But very local mobility is just as important in designing a paper-like computer technology for reading. As it stands, people print documents to read them away from their computers, even if "away from their computers" only means leaning back in their chairs or sitting elsewhere in their offices. This effect becomes more pronounced in the case of within-building mobility. In a recent internal survey about reading, one question asked respondents why they print; nearly 90% said they'd print a document to take it elsewhere to read it. In another question to determine the three most important reasons for reading on paper, nearly half of the survey respondents said that one of the key reasons that they printed documents is because of the "location where I will read the material." Field visits revealed that many office workers printed materials so they were able to read where it was quiet (for example, a first aid room), at home in the evening, in a comfortable place in their office, or to bring a document with them to another venue (for example, to a meeting).

Likewise, the students participating in the Pocket PC study at the University of Virginia found mobility to be a compelling aspect of the handheld computers. [16] And indeed, in many cases where standard secondary texts are heavy, and multiple resources must be brought to bear on classroom assignments, digital devices are a compelling way to access and use digital materials.

Materiality

One of the important losses people cite when they are given digital materials to read concerns the materiality of books and documents. In Geoff Nunberg's 1993 article in *Representations*, he creates an evocative image of the shift to digital materials:

> "Reading what people have had to say about the future of knowledge in an electronic world, you sometimes have the picture of somebody holding all the books in the library by their spines and shaking them until the sentences fall out loose in space..." [19]

Thus it is important when we design reading technologies to think about reclaiming some the material characteristics of physical documents and books, including the affordances of the page and how people perform the physical act of reading.

The materiality of the book

The materiality of the book speaks volumes about what we're reading – how long the work is, the degree of permanence of the publication, how old the particular edition is and how much it's been used by other readers, whether the book has sentimental value, and so on. For example, compare books bound by hand by a small high-end art press with mass-produced mechanically bound books and with books that are spiral bound at a consumer-oriented print shop like Kinko's.

It's difficult to get this kind of immediacy from a digital representation, no matter how clever the user interface is. An English Literature grad student, describing reading on the Pocket PC, observed:

> "You get this little screen, so you get no sense of even how long the work is... You have 600 pages, which means what? No-one knows. And so ... I definitely don't see it as a literary experience." [16]

Interestingly, the e-book displayed how long it was (the number of pages) and the student had a page in front of her that told her how far into the work she had read, but the literary experience that she was seeking was still missing. It was crucially tied to the material form of the book; it may be difficult to reclaim such visceral feedback.

This brings us to another aspect of the materiality of the book: it is still easier to navigate on paper than it is to navigate in an electronic document, despite our collective best efforts to provide people with within-document search capabilities and hyperlinks. Some of the most satisfying navigation is visual. People use a variety of cues to make their way through physical books, a combination of distance into the work, paratextual information like section and chapter headings, and – if they are available – visually distinctive page elements such as photos and graphics.

We can take advantage of this combination of cues to greatly facilitate navigation in digital material. Usually this requires a combination of user interface techniques like structural jumps, reduced visual representations like thumbnails, search, and riffling.

The materiality of the page

By far the most effort so far has gone into reclaiming the materiality of the printed page and studying how it affects reading. My colleagues have given considerable thought to how to make typefaces more readable and more aesthetically pleasing when they are displayed on LCD screens. Subsequent evaluation has revealed that people recognize words more quickly when the ClearType technology is used as the basis for presenting text on LCDs [4].

My colleagues have also worked on adaptive layout to recapture the look of a typeset page. Techniques for adaptive page-like layout have shown clear benefit for some genres of electronic materials (novels, for example) on regular size displays, but involve further trade-offs when the displays are small, such as on handheld computers or when fixed layout is an intrinsic part of the genre (such as verse). Sometimes a more compact, less "readable" layout is preferable when the need for context outweighs the advantages of reproducing what looks like a paper page. Furthermore, when rhyming words of poems were set apart on new lines (according to the adaptive layout algorithm), readers found them to be almost doggerel-like to read.

The physicality of reading

The final area of materiality that we need to consider is the physicality of reading. In some ways, this is closely tied to our observations about mobility. Reading is a physical activity, an activity with a setting and ergonomic considerations. None of this argues against reading on a computer, especially on a computer that is designed for a typical

reading setting. A student described how she read class assignments in her living room at home using her laptop:

> "And I heard things like, 'Oh I could never curl up with a computer.' And I always test that for myself. And in fact, I would lie back on my couch with my feet up on this end, my head propped against the pillow there." [16]

In short, from a material point of view, to make reading on the screen succeed, we must identify genres least tied to material form and tackle those first; we need to identify affordances that can be reclaimed in digital representations (for example, the ability to navigate effectively); and finally, we should be mindful of matching physical form factor of the reading devices to the physicality of reading.

Interactivity

At the beginning of this paper, I cited recent publications that demonstrated students' – and no doubt others' – lack of critical perspective on the information they find on the Web. Paper documents admit a surprising amount of interactivity and in so doing, support critical thinking.

By contrast, digital forms often strike people as "dead" or oddly under-reactive. In many cases, it's easy to change the format and content of digital documents, but it's difficult to interact with them in an informal, unplanned, or unselfconscious manner, as you would if you were writing in the margins of a page as you read it, if you were tearing an article out of a magazine, or if you were comparing several documents side by side as you sorted them. Thus annotation, clipping, and gathering are key kinds of interactivity, interactivity that is supported – at least to some extent – by paper.

If we look at them in a space, we might think of them this way:
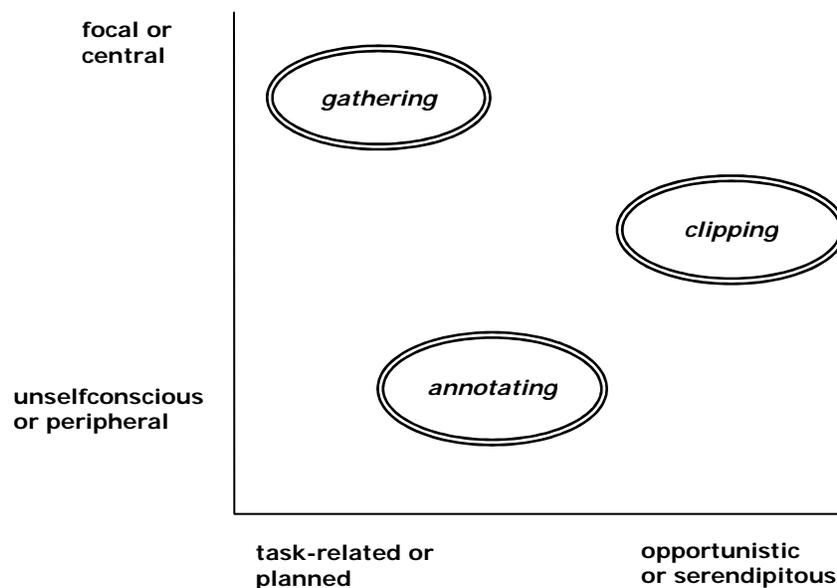
Personal annotations are usually done relative to a task, but it's interesting to note that they're relatively unselfconscious, peripheral to the reading activity. Unlike scholarly

marginalia that are part of an annotated edition, personal annotations are informal, and readers are often unaware that they've made a given mark. It is not unusual for them to be surprised by how much they've written in a given book or marked on a document.

Gathering is also performed relative to a task, usually research similar to that which motivated the Graham and Mataxas study. It is a focal activity, central to the interaction of the materials, and reading is relegated to a more secondary role.

A third common form of interaction is clipping, intentionally saving relevant portions of published material, such as an article from a magazine. Like annotation, clipping is usually not the focus of an activity, but rather an unselfconscious, yet productive, sideline to reading. Unlike both annotation and gathering, clipping tends to be opportunistic, occurring when a reader encounters an item of interest to save or share.

I'll look at each of these three forms of interaction to demonstrate their key characteristics, emphasizing what they all have in common and what it means for designing paper-like interfaces for reading.

Annotation

For the purpose of this discussion, I'm considering all of the informal marks that readers make on the printed (and someday, the electronic) page as annotations. Underlines, highlights, circles, marginalia, arcane symbols, calculations, drawings, and other marks all fall into this realm. Excluded are annotations intended as communication (as the marks a professor shares with her students) or authors' annotations that represent suggested changes to an unpublished manuscript. Thus, the focus of this discussion is on readers' marks on published materials of the sort you would find in a digital library.

Studies of how people annotate in these circumstances demonstrate the importance of direct pen-based interaction with an electronic document. If we want people to read on the screen, we're going to have to provide them with the facilities to annotate. The standard keyboard-mouse interaction paradigm is not sufficient support for the kind of unselfconscious interaction that we observe in practice.

But our studies also reveal several conundrums. First, what is immediately apparent from the empirical data is that most annotations are simply highlights, underlines, and circles, in spite of the fact that people idealize annotations as being significant marginalia and notes. In a recent study, we found that over 82 percent of the 1500-plus annotations we coded as part of our analysis were underlines, highlights, and circles. The ten percent that were comments on specified regions of the text were often cryptic, understandable only to the reader who wrote them. So we are left with a suspicion that readers may not find their own annotations particularly useful or intelligible later, especially outside the context of the immediate task.

Do readers realize this? Not always. They conceive of their own annotations as having some intrinsic long-term value and go off in search of them. One student described searching for her old annotations and being disappointed at what she found:

> "Some of them [the annotations] are absolutely ridiculous and I can't believe that I actually wrote this in pen in this book. Some of them are – I have no idea what I'm talking about. Some of them are really interesting,

and it's something I'd forgotten. It just depends on the notes. ... when I did Milton, we were doing the epithets about Satan or something, so I underlined all of them. And when I was going back through it, I'm like 'what on earth!'" [16]

The other aspect of personal annotations that this quote alludes to that seems to conform to a general pattern in data across studies is that personal annotations are private. In the long run, people may be embarrassed by them and what they reveal about their understanding of the text. In fact, it seems that the only way annotated books make their way into used book stores is by the apparent guarantee of anonymity.

Thus two paradoxes emerge: (1) the difference between the idealized memory of annotations and their actual value; and (2) the treatment of the annotated materials and the expectation of privacy. We'll see this again in the next kind of interaction, clipping.

Clipping

Readers clip items out of published materials – or print portions of longer electronic publications – for many reasons. Sometimes these clippings or printouts represent unfinished reading; sometimes they are reminders for later action (for example, a need to follow up on a URL); sometimes they seem useful in some anticipated situation, or seem to be good references (e.g. how-to descriptions); sometimes they will simply be evocative and call to mind a pleasant or historic event; and sometimes they represent something a reader wants to share with a colleague or friend. In fact, more than 40 percent fall into this category [15].

In this situation, the important thing about paper-like interaction is how easy it is to clip something out of published material and how little it interrupts reading. What people do with these clippings, and why they create such a paradox and an opportunity for us in the electronic realm, is that they are often forgotten and don't serve their original purpose. Readers rely on *re-encountering* their clippings, and thus leave them in places where they will see them again. Interestingly, computers provide relatively sparse venues for re-encounter: it's very difficult to "leave something out" on a computer.

Gathering

Gathering is the counterpart to annotation. Annotation, at its most useful, represents in-document interpretation; gathering represents a relative interpretation of related documents brought together for a single purpose.

In many ways, paper practices don't support gathering that well. In past studies, we have found that there is a tension between organization and actual use, and a need for spatial persistence that acts in opposition to a reader's desire for mobility. For example, we found that law students tended to organize printouts of legal cases into three categories: precedents that supported their side; precedents that ran counter to their side; and precedents that were a close match on the basis of legal facts. But they had two common ways of maintaining that organization. One was to file the printouts in a three-ring notebook. That way, the organization stayed in place, but to use any two documents together the student had to take them out of the notebook. The other was to spread the documents out around them into piles, but in that case, the students seldom had the luxury of leaving them that way (and the piles could become messy). More general

studies of office workers have found that piles are an effective way of keeping documents at hand and using them as reminders for action [11]. For the last decade, my colleagues and I have been working applications for gathering and representing this kind of lightweight interpretation by using spatial persistence to preserve evolving structure [22].

Sharing

Reading is, at its essence, a social activity. Of course, people may still read alone, but what they choose to read and their capacity for interpreting the material is circumscribed by the communities to which they belong.

Even more narrowly construed, there are many circumstances in which people read together. Some are institutionally supported, for example, students reading together in a classroom situation, researchers reading published articles together in a reading group, or analysts reading the news together in a morning meeting. Others are opportunistic. When WebTV first came out, I observed people congregating in public spaces to cruise the Web; this gathering was not intentional, yet it fundamentally changed the character of Web browsing as they negotiated which links to follow and where to go next. Similarly, when the Web was new and still a novelty, we observed a middle school class (students who were about 12 or 13 years old) using an on-campus lab to explore the Web. Instead of sitting alone at individual terminals – there were sufficient terminals for each student to have his or her own – the students gathered in small groups of two or three and explored the Web together [23].

Thus, reading together is a relatively common and important phenomenon. Some elements of reading together on paper in this fashion include the ability to look together at a single focus of attention and the ability to pass a book or other document hand-to-hand.

In some cases, reading together means looking over another person's shoulder or sitting side by side in front of a book or computer screen. In others, it means being able to readily get to the same point in multiple copies of the same document, as for example when a professor asks his class to "turn to page 47, top of the page." Each poses distinct challenges for a paper-like interface (and sometimes for paper itself).

In the first case, LCD technology places some distinct restrictions on viewing angle; it is often difficult for two people to see the same screen at the same time. Even if the display is legible, colors vary at different angles, and the reading experience differs significantly.

In the second case, if navigation is harder in electronic documents than in paper documents, then co-navigation – getting everyone on the same page – is harder still. For one thing, getting everyone on the same page is a very inexact notion in a dynamic social situation. Five members of a reading group can be looking at the first page of a journal article, and a sixth can be looking at the last page to examine a reference, and they can still be productively engaged in the same discussion.

In a classroom situation, it is unlikely that everyone will be using the same edition of a book, making absolute reference difficult. When we observed an English Literature class session, shared reference was often by line number or structural elements in the texts to get around differences in edition or differences in anthology contents. This is a problem

that is relatively easy to solve on a computer, but it is essential that we solve it. One class abandoned the in-class use of e-books on the Pocket PCs because navigation by word search was just too slow for the class to synchronize at a single place a long work [16].

The ability to pass electronic works hand-to-hand seems to be at its most paper-like when people are able to use the infrared ports to "beam" the material to each other. The undergraduates at the University of Virginia saw this capability – being able to transfer the documents to the professor's own handheld – as the most like physically handing in their homework assignments. In a digital library setting, we can envision this kind of hand-to-hand transfer only if Digital Rights Management software permits it. As Samuelson has pointed out, DRM relies on explicit specification of permissions, rather than on a common understanding of fair use rights [20]. In this case, it might mean that mode of transfer (infrared) should be taken into account for certain kinds of copyright-protected works, rather than making it impossible for the same work to reside on computers owned by two different people, or that infrared transfer might be more literally like transfer of paper materials, where it leaves one computer and moves to another.

Transcending paper

Designing a paper-like interface for reading digital library materials is a valuable and indeed necessary exercise. But besides addressing ecological concerns about wasting paper and other resources, it is not enough to simply echo paper's capabilities; in some situations and for some activities, paper is too malleable and too useful to replace. On the other hand, it is possible to create a compelling on-screen reading experience that takes advantage of maturing computer technology to go beyond paper.

Toward a personal digital library and persistent records of reading

First of all, it must be easy to get materials onto a personal reading device, both in terms of wireless access and in terms of a good interface for retrieving or moving relevant materials to the device. It is also wise to factor place-based displays into this equation. If a reader walks into a room with better displays available than the one she has brought with her, she should be able to take advantage of them.

Second, it's important to pay attention to the genre of materials we choose to prepare for digital libraries, especially in the early days of incorporating computers into our reading practices. In our study at the University of Virginia, the materials that were the most valued as digital resources were the extensive secondary sources that the students needed for their classes, especially the materials that currently have limited access, such as those in special collections. Obviously materials that incorporate multimedia are also good candidates for digital presentation, but because these materials are more difficult and expensive to prepare, I won't focus on them. Materials readily available as print editions and materials sensitive to form and format are less likely to be embraced as digital offerings, although shortly I'll talk about exceptions to this rule of thumb.

Materials preparation is also where the publishers and digital librarians play an important role in transcending paper. Earlier I discussed the essential nature of navigation in supporting many kinds of reading. Besides the direct access offered by search functionality, much of navigation within texts is performed using their structural characteristics, for example headings, captioned figures, and photos and other

illustrations. As it stands, much of the mark-up I see on texts does not anticipate navigation (a notable exception being when standards such as TEI are applied). Publishers familiar with high-end tools like Quark are oriented toward the visual appearance of the page, which is how navigation is supported on paper.

To further transcend the possibilities offered by paper, it may be better to not even talk about the notion of electronic books and electronic book readers, but rather of portable personal library with everything you've read and a record of what you've done with it. Several years ago, Clifford Lynch wrote:

> "Given the historic price-performance trajectories for storage, in a few years at least some high-end appliances will house hundreds, if not thousands, of books simultaneously, and certainly laptops with software book readers will house thousands or tens of thousands of books at once. Think of portable personal digital libraries, not portable electronic books, as the future role of these appliances." [10]

Facilities that institutional digital libraries offer in conjunction with their collections, like search and collection management interfaces, are just as relevant to these personal libraries. Besides being useful for navigation and analysis, search can help readers locate materials they've read before. Techniques such as those used in Susan Dumais's Stuff I've Seen might be brought to bear on this problem [5].

Earlier I talked about ways in which paper supports interaction. In personal digital libraries, we can think of these interactions as forming a persistent record of reading. The empirical data support the idea that each record – each annotation, each clipping, each pile of papers, each book finished and put away – doesn't have much value on its own. But there is also considerable evidence that taken together as an accumulation, *all of these records form a personal geography of one's own collection of reading materials.* Combined with the notion of a portable personal digital library, this is a powerful idea indeed. It's not so much that a reader would want to see the precise annotation she had made on any of the technical papers she has read over the years, it's rather that she had read them at all and paid attention to particular passages that is interesting later on.

Practice-specific capabilities

Additional capabilities that meet readers' needs can be developed using a rich understanding of specific work practices. However, it is important to realize that this sort of functionality may continue to be domain- setting- or situation-specific. For example, active links to precedent cases are a distinct advantage in the legal domain. However when we gave the same functionality to computer science researchers, they had their doubts about its utility and even thought the active links might be a significant source of distraction: unknown references were likely to be unknown for a reason, and even if the reference was useful, following a link would interrupt reading. In this case, active links might be better replaced by an active "To Read" list.

Another good example of situated beyond-paper capabilities comes from the Perseus project. Marchionini found that students may prefer to read ancient Greek texts on the screen over reading on paper because the immediacy of dictionary lookup increases the speed and comprehension with which they can read [13].

One other situation-specific functionality for transcending the limitations of paper comes from a possible use of annotations in the classroom. It is clear that people annotate for a variety of reasons. But if we look at a group of students all annotating for roughly the same purpose, to assimilate a reading assignment, we can see some interesting phenomena. There is greater consensus on what they annotate (remembering that most annotations are highlights and underlines) than what we would predict based on a probabilistic calculation of overlap. Furthermore, annotators converge on passages and sentences that are different from the text designated as important by the authors and publishers. Thus, if we were mindful of privacy issues, we could potentially use annotations and other records of reading in a variety of ways – summarization, guiding class discussion, and facilitating skimming [14].

The more ways in which we can transcend paper – and I've just offered a few examples that have been prominent in my own studies – the more likely we are to accelerate changes in reading practice.

Constellations of reading technologies

The changes in reading and reading technologies don't imply that there's a single way the future of reading will play out. In practice, reading is a heterogeneous activity and reading technologies are better for some things than they are for others. Choosing a single platform to support reading and critical thinking is not only unnecessary, it seems unlikely.

What we do see is new constellations of technology and new kinds of reading practices emerging to take advantage of them. A student writing a term paper might use a handheld with his class readings in conjunction with electronic materials he has retrieved from a digital library on his desktop computer, a paper text he is consulting, and the notes he took earlier that day in class. Or people may read different versions of the same material on the screen and on paper at different times and for different reasons.

So contrary to the predictions of just a few years ago, as well as somewhat older experimental evidence that says that people simply prefer to read on paper, reading is and will continue to change, and computer screens are becoming part of a richer picture of the future of reading.

References

[1] Betrisey, C., Blinn, J. F., Dresevic, B., Hill, B., Hitchcock, G., Keely, B., Mitchell, D. P., Platt, J. C., Whitted, T. Displaced Filtering for Patterned Displays. *Digest of Society for Information Display Symposium 31*, 1 (May 2000), 296-299.

[2] Crawford, W. Paper Persists: Why Physical Library Collections Still Matter. *Online Magazine* (January/February 1998), 42-48.

[3] Dillon, A. Reading from paper versus screens: a critical review of the empirical literature. *Ergonomics 35*, 10 (1992), 1297-1326.

[4] Dillon, A., Kleinman, L., Bias, R., Choi, G., and Turnbull, D. Reading and Searching Digital Documents. *Proceedings of the American Society for Information Science and Technology (ASIS&T 2004)*. Information Today, Inc. (2004).

[5] Dumais, S., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R., and Robbins, D. Stuff I've Seen: a system for personal information retrieval and re-use. *Proceedings of SIGIR'03.* ACM Press (2003), 72-79.

[6] Graham, L. and Metaxas, P. T. 'Of Course It's True; I Saw It on the Internet!': Critical Thinking in the Internet Era. *Communications of the ACM 46*, 5 (May 2003), 71-75.

[7] Grudin J. Why CSCW applications fail: problems in the design and evaluation of organization of organizational interfaces. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW88)*. ACM Press (1988), 85-93.

[8] Guterman, J. Making E-Books Safe for the Toilet. *Business 2.0*, November 14, 2002.

[9] Levy, D.M. *Scrolling Forward: Making Sense of Documents in the Digital Age*. New York: Arcade Publishing (2001).

[10] Lynch, C. The Battle to Define the Future of the Book in the Digital World. *First Monday: A Peer-Reviewed Journal on the Internet 6*, 6 (June 2001). Online at http://www.firstmonday.dk/issues/issue6_6/lynch/index.html

[11] Malone, T.W. How do People Organize their Desks? Implications for the Design of Office Information Systems. *ACM Transactions on Office Information Systems 1*, 1 (1983), 99-112.

[12] Manguel, A. *A History of Reading*. New York: Viking (1996).

[13] Marchionini, G. Evaluating Digital Libraries: A Longitudinal and Multifaceted View. *Library Trends 49*, 2 (Fall 2000), 304-333.

[14] Marshall, C.C. Toward an ecology of hypertext annotation. *Proceedings of ACM Hypertext '98*. ACM Press (1998), 40-49.

[15] Marshall, C.C. and Bly, S. Sharing Encountered Information: Digital Libraries Get a Social Life. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL04)*. ACM Press (2004), 218-227.

[16] Marshall, C.C. and Ruotolo, C. Reading-in-the-Small: a study of reading on small form factor devices. *Proceedings of the Joint IEEE and ACM Conference on Digital Libraries (JCDL02)*. ACM Press (2002), 56-64.

[17] McKnight, C. and Dearnley, J. Electronic Book Use in a Public Library. *Journal of Librarianship and Information Science 35*, 4 (2003), 235-242.

[18] Neumann, P.G. E-Epistemology and Misinformation. *Communications of the ACM 46*, 5 (May 2003), 104.

[19] Nunberg, G. The Places of Books in the Age of Electronic Reproduction. *Representations 24* (Spring, 1993), 13-37.

[20] Samuelson, P. DRM {and, or, vs.} the Law. *Communications of the ACM 46*, 4 (April, 2003), 41-45.

[21] Sellen, A. and Harper, R. *The Myth of the Paperless Office*. Cambridge, MA: MIT Press (2001).

[22] Shipman, F., and Marshall, C. C. Spatial Hypertext: An alternative to navigational and semantic links. *ACM Computing Surveys 31*, 4 (Dec. 1999), 14.

[23] Shipman, F., Marshall, C.C., Furuta, R., Brenner, D., Hsieh, H., and Kumar, V. Using Networked Information to Create Educational Guided Paths. *International Journal of Educational Telecommunications 3*, 4 (1997), 383-400.