HCI-Book? Perspectives on E-Book Research, 2006-2008
(Foundational to Implementing New Knowledge Environments)

Ray Siemens, Teresa Dobson, Stan Ruecker, Richard Cunningham, Alan Galey, Claire Warwick, and Lynne Siemens, with Karin Armstrong, Michael Best, Melanie Chernyk, Lynn Copeland, Wendy Duff, Julia Flanders, David Gants, Bertrand Gervais, Karon MacLean, Steve Ramsay, Susan Schreibman, Colin Swindels, Geoffrey Rockwell, Christian Vandendorpe, John Willinsky, Vika Zafrin, the HCI-Book consultative group and the INKE Research group.*

0. A Note On the Type of Essay This Is

In most all things that exist at the intersection of several domains, domain-specific cultures have potential to collide, in useful ways as well as others. The book is such a thing—especially so, one might observe, in an age witnessing the book's seeming redefinition in the midst of societal reconfiguration with respect to technologies and media of conveyance. As has been noted by many, Katherine Hales and Alan Galey among them, the advent of the electronic book, the *e-book*, has made the book itself visible to us as an object of study in new ways, ways that have, in quick succession, metaphorically and analogically fertilized and fomented our understanding of new forms of electronic-reader book-ishness and concomitant electronically-facilitated reading practices. Indeed, as a powerful metaphor for textual forms of communication, the notion of the book as knowledge

---

environment spurs development of e-readers in the direction of emerging universal electronic libraries; from perspectives of its physical artifactual nature, and its formal components, book elements and features are mimicked, augmented, and enhanced as they are prototyped and deployed in electronic reading environments. The rise of the e-book has taught us not only how much we still have to understand about books themselves, but also how much we have yet to understand about books and reading in the context of their emerging electronic correlatives; further, as we explore in this vein, we cannot help but realize the great number of domains that inform our understanding of the book.

This essay reflects an examination of the book in the context of these many pertinent domains, in the early days of the electronic book's formulation. It summarizes a consultative process associated with a research team that met 2006-2008, struck to foster further understanding of the significance of digital and analog books and their role in humanities scholarship, supported by the Social Sciences and Humanities Research Council of Canada (SSHRC) via a research cluster formation award, and called HCI-Book (Human-Computer Interaction/Interface with the Book). Drawn to reflect the expertise of a considerable number of pertinent domains, disciplines, and sub-disciplines, our exploration was marked by varied types of presentation, documentary narrative, exemplification, argumentation, and conflict- and consensus-oriented deliberations in the many ways accepted by all the domain-/disciplinary-cultures represented by members of our group—something that is maintained in the voice(s) of this essay. In our deliberations, cultures did collide, in the most useful ways, and aspects of those points of collision are maintained across a number of elements of this essay, including those which manifest in voice and style of expression.

Together, across several meetings, we asked ourselves what we really knew about the book, and about the ways in which we interact with new texts that replace the print artifact and re-present to us the knowledge and experience of the past, as well as those that deliver the direct-to-digital record of the present; we explored the ways in which we understand our interaction with these knowledge objects, and the information they contain, as well as the impact that the confluence of media formats in these digital objects has on our use of them, such that we may best facilitate interaction with the new digital artifact. Through such consultation, our group began to identify issues central to the digitization of the human record,

and to act on that identification with the goals of understanding and describing the basic principles of humanistic interaction with knowledge objects (digital and analog alike), of articulating core strategies for the design of humanistic knowledge objects, especially electronic books, based on this understanding, and of suggesting basic principles necessary for the evaluation and implementation of current technologies, and the exploration of future ones. Our final expression took shape around the key areas discussed below in this essay, beginning with the contexts we might most readily identify for electronic book research: those mediated by audience, by issues of interface and design, of form and content and, in relation to the readers of books and users of their analogous digital counterparts, methods of reader/user studies in work such as ours, their importance, and previous studies on which we could imagine informing book and e-book research. These larger contextual foundations supported more specific consideration of books and e-books in terms of their material, symbolic, and formal aspects, plus a use-based consideration of their features and, further, led to a wide-ranging consideration of issues pertinent to research prototyping of e-books that reflect identifiable patterns of book interaction, from the perspectives of interface, preservation, encoding, reading tools, manifest analytical and algorithmic matters, media and platforms, and issues related to reading culture. Our final consideration took the form of a reflection on the nature of the consultative process itself, the need for an interdisciplinary approach to understand the broad range of issues pertinent to books and e-books in the context of our work, the desire to achieve consensus even in domain-influenced collision, and the challenges of extending this type of work beyond consultation and into an active program of research.[1]

---

[1] In 2009, this group's research was supported via a SSHRC Major Collaborative Research Initiative award as the Implementing New Knowledge Environments project, which founded its research plan firmly in these formative consultations and published its next steps in Siemens, *et al.*, "Codex Ultor: Vers une base conceptuelle et theorique pour la nouvelle recherche sur les livres et les environnements documentaries" (2010), "Codex Ultor: Toward a Conceptual and Theoretical Foundation for New Research on Books and Knowledge Environments" (2010), and "Implementing New Knowledge Environments: Laying Research Foundations for Understanding Books and Reading in the Digital Age" (2009). The most recent publications from this group can be found via http://bit.ly/oLiY6P (Google Scholar, "INKE Research"), http://bit.ly/r6XVbn (Google, "INKE Research") and, annually updated beginning in 2012, http://inke.ca/.

1. Introduction: Contexts for Electronic Book Research

More than half of all people living in developed countries make use of computers and the internet to read newspaper, magazine, and journal articles; electronic copies of books; and similar materials. The next generation of adults already recognizes the electronic medium as its chief source of textual information. Our knowledge repositories increasingly favor digital products over the print resources that have been their mainstay for centuries. Moreover, a chief priority of those professionals who produce and convey textual information is to make such information available electronically in ways that meet the standards of quality, content, and functionality that have evolved over half a millennium of print publication. The movement toward the use of the digital medium is an inevitable one, with clear benefits associated with the production, dissemination, and reception of the record of human experience, as well as the ultimate impact of these processes on our knowledge-based society. However, for all the good we perceive, we also realize that there is much still to know about this new media form. We must gain this knowledge in order to make the best use of what the digital has to offer us. What do we really know about the ways in which we interact with these new texts that replace the print artifact, re-present to us the knowledge and experience of the past, and deliver the direct-to-digital record of the present? Do we understand the ways in which we interact with these knowledge objects, and the information they contain—and do we understand the impact that the confluence of media formats in these digital objects has on our use of them, such that we may best facilitate interaction with the new digital artifact?

In short, there is much that we still need to learn about the new knowledge machine presented to us by the computer. More specifically, there is at the moment a real need to understand the principles involved in dealing with digital artifacts, so that we may interact in an effective way with the digital representations of the objects of human experience. This paper reports on the efforts of a working group to identify issues central to the digitization of the human record and to act on that identification, with the aim of:

- understanding and describing the basic principles of humanistic interaction with knowledge objects (digital and analog alike),
- articulating core strategies for the design of humanistic knowledge objects, especially electronic books, based on this understanding, and

- suggesting basic principles for evaluating and implementing current technologies, and exploring future ones.
- Possibilities for human–computer interaction and the electronic book may be examined from a range of interrelated perspectives, which are approached in several essential ways:
- via a process that seeks to identify, quantify, and evaluate print and electronic books in terms of their features and uses;
- via a process that explores the material, symbolic, and formal aspects of the book, toward the end of computational modeling; and
- via a process of prototyping computational models and simulations of the book, both literal models and metaphoric.

Audience

Perspectives on the relationship of audience to the electronic book come from a number of disciplines, and debates about reading in hypermedia environments often become mired in philosophical disagreements among these disciplines regarding what constitutes a satisfactory reading experience—or, put another way, what constitutes a usable text. For instance, early research on reading hypertext, conducted by researchers concerned with interface design and software engineering, tended to focus on articulating best practices to prevent reader disorientation and cognitive overload in complex, highly networked multimedia environments. Navigational supports such as site overviews and hierarchical structures were promoted as a means by which to improve usability (Foss 1989; Leventhal, Teasley, Instone, Rohlman, and Farhat 1993; Astleitner and Leutner 1995; Kim and Hirtle 1995; Rouet and Levonen 1996).

A sustained conversation between the fields of scholarly editing and empirical reader studies is long overdue, especially since scholarly editing inevitably involves a process of informed speculation about who will read an edition (and for what purpose), and about the reader's use of paratextual elements such as cross-references, foot- or end-notes, etc. The reader studies component of our project is in part concerned with how scholars come to trust or distrust digital resources—a question of major and perhaps unexpected importance to digital scholarly editors (Best 2004), just as the socially constructed trustworthiness of print has been an important research question for scholars in book history (Johns 1998). Nevertheless, the figure of the reader in much editorial theory often remains abstract and even mystified, even though readers' habits are often invoked to justify

certain ways of presenting the text. Decisions such as modernizing spelling versus retaining the original orthography, conflating texts into a single ideal form versus presenting multiple versions, providing one kind of annotation but not another, and silently emending an apparent error versus leaving the text unchanged all serve to configure the relationship between reader, editor, and text.[2] Reader studies provides a formal vocabulary to describe these relationships, an important task for textual scholars given the capability of digital editions to reach new audiences, as well as the need for digital scholarly editors to engage with interface design and usability.

Interface and Design

The design of computer interfaces for researchers working with electronic texts requires a combination of specialist areas of inquiry, including the ethnographic study of information-seeking behaviors, diagnostic performance evaluation of existing interfaces, and iterative design and usability study of new design prototypes (Ruecker 2003). Some interfaces are intended to provide researchers with access to collections of materials. Others aim to facilitate research tasks once an appropriate subset of materials has been selected by the user. One of our goals is to bring together expertise in the areas listed above in order to inform the design of new "affordances" (Gibson 1979) for people working with digital texts.

Following Gibson's definition, an affordance is an opportunity for action, and the design of a new digital affordance provides people with a tool that was not previously available. An example of a widely successful digital affordance is the cut-and-paste function, which had not been available to writers using typewriters, but was adopted wholesale by those using word processors. As soon as the technology was able to support it, people learned that it was available, and a function that had been reserved for the editorial process became indispensable as part of the writing process. On the other hand, a digital affordance that is not yet widely available is the dynamic table of contents (Ruecker 2005), which would allow the reader to perform a variety of research tasks by interactively adding or subtracting content.

---

2   See the articles in Best 1998, especially those by Werstine, Anne Lancashire, Ian Lancashire, and Siemens.

Form and Content

A key question for those interested in the representation and re-presentation of texts in any form has been whether content exists abstractly, independent of its representation by an interface, or whether it exists only concretely, as the sum of its instantiations. One of the problems currently evident in the field is the document-mindedness of ideas inherited from (mostly literary) hypertext theorists, who tend to speak of content in abstract terms, as links and lexias, but almost never as material objects, identifiable classes, and specific instances. On the other hand, the field also contains many examples of idealistic approaches to content such as can be seen in the guidelines of the Text Encoding Initiative (TEI). The TEI is an international project developing guidelines for encoding machine-readable electronic texts, with the aim of facilitating activities such as text analysis and sophisticated searching, and ensuring long-term accessibility. While the TEI is well-suited to encoding print objects (for example, a novel), it is less useful when dealing with ontologically complex texts such as plays, which are both documents and events. As F.W. Bateson famously phrased the problem, "if the Mona Lisa is in the Louvre, where [is] Hamlet?" (quoted in Greetham, 342) And if an edition of Hamlet is an interface between the readership and the content, where then does the content begin?

2. Readers and Users

To begin, we think it is important to consider that "user-centered" approaches to information systems need to be "user-in-context" approaches. That is, we must take into account variations in use and interaction patterns that depend upon specific scenarios, tasks, roles, subject domains, and socio-cultural frames. To cite Andrew Dillon, "it makes no sense to describe a tool or technology as usable or unusable in itself. Any tool is made for use by certain users, performing particular tasks in specific environments. Its usability can only be meaningfully evaluated in relation to such contextual variables" (Dillon, 17). Examples of this approach to system design are contextual design and cognitive work analysis, but the approach itself has also become a focus of research, as evidenced by the recent *Information Seeking in Context* and *Information Interaction in Context* conferences (Ingwersen and Jarvelin 2005).

In this light, one of the goals of our investigation should be to identify and test the impact of contextual variables of interaction with electronic texts. This would include conducting task analyses and developing taxonomies of interaction types, use scenarios, and user roles. The starting point for this can be the two scenarios (comprehension and engagement) and the two user groups (scholars and readers), but we aim to develop a more robust model of user testing that is specifically relevant to the e-book. The development of these taxonomies would provide a framework for an ongoing discussion of how "use" and "interaction" can be interpreted with respect to the e-book.

User Studies and Usability Assessment

The wide adoption of "use" as a descriptor for engagement with hypermedia reflects the challenges inherent in understanding and facilitating interaction with complex multimedia artifacts. It also points to a potential problem with research in this area: when we attempt to accommodate the digital artifact's complexity by devising terms like "use," which synthesize the range of processes involved in human-computer interaction, does that deter us from realizing the distinctiveness of those processes?

One component of our research will entail close examination of how literary structures and reading processes are extended and modified in digital environments. As many theorists have observed, a close look at modern print and electronic texts reveals a shift away from conventional narrative logic, and toward indeterminacy, fragmentation, and open-endedness (e.g., Landow 1997; Bolter 2001; Van Peer and Chatman 2001). At the same time, current theories concerning how readers engage with texts tend to be derived from studies of readers working with "normal" prose or other conventional narratives (e.g., Chatman 1978; Rabinowitz 1987; Kintcsh 1988; Zwaan, Magliano, and Graesser 1995).

Many contemporary models of reading are built on the premise that the act of reading relies in a large part on distinguishing between significant and insignificant narrative details, so that we may generate a workable mental model of the situation described in the text. The Construction-Integration Model proposed by Kintsch (1988), for example, describes a process whereby readers construct meaning by identifying potentially relevant elements; they then develop an integration system in which appropriate elements are strengthened and inappropriate elements are weakened or discarded. Chatman

(1978, 53 ff.) likewise proposes that readers of fiction distinguish kernels (major events) from satellites (minor events). Finally, Rabinowitz (1987) suggests that in making sense of narrative, readers follow rules of notice; he describes this as a process of identifying more significant details and separating them from less significant details.

Miall and Kuiken (1994) have critiqued models such as this, particularly Kintsch's. They point out that such models fail to take into account the extent to which readers' experiences of literary texts are modified by their emotional response to stylistically emphasized language, such as metaphor and alliteration, which engages the reader's feelings and evokes "less prototypic, more personal meanings" (339). According to Miall and Kuiken, understanding literary response requires a different mode of analysis from the one implicit in text theories that have been developed based on studies of "normal" (that is, informational) prose. These models, they note, generally describe "a resource-limited system in which cognitive structures (e.g., story grammars) or procedures (e.g., integrating processes) economize comprehension by deleting irrelevant propositions, inferring relevant propositions, and building macropropositions" (344). In other words, they focus on ways in which comprehension is facilitated or economized. In this respect, Miall and Kuiken argue that the current models are too limited for the purpose of understanding response to literature, because the essence of literary text dwells at least in part in its stylistic features, and these features are less likely to economize comprehension than to complicate it "by challenging the familiar, prototypic concepts that readers initially apply to the text" (344). Even theories of the reading process based on studies of readers' engagement with literature, such as Chatman's and Rabinowitz's, tend to be biased toward a focus on economizing comprehension. Rabinowitz's notion that readers engage in a process of distinguishing significant from less significant elements, for example, privileges plot over other features of narrative; Rabinowitz presumes that anything in the text not immediately relevant to developing the situation is marginal to the reader's understanding of the text.

This view of literary reading is problematic, particularly when we consider that much contemporary fiction, both print and digital, is increasingly complex and fails to conform to conventional expectations respecting narrative logic. Van Peer and Chatman (2001) observe that the diverse narratives of the twentieth and twenty-first centuries are incompatible with existing narrative models. This is in part because most of these models reflect a monolithically western

perspective, and in part because the models do not take into account new media genres. To develop appropriate models of reading complex print or digital narratives, we must examine how people engage these texts, and revise our perspectives of narrative structure and literary reading processes.

In discussing empirical research related to how people engage books and e-books, then, we envisage two main types of reading: 1) the type in which the economization of comprehension is the aim; and 2) the type in which engagement with an aesthetic artifact is the aim. With respect to the first, information-seeking in the humanities is a well-researched area (e.g., Dalton and Charnigo, 2004). Nevertheless, our understanding of how humanities scholars engage computer-based information resources is in continual need of refinement as new research resources are developed. An overview of research in this area is provided below. With respect to the second scenario mentioned above, models of reading based on observing readers of complex print narratives, or emerging hypermedia genres, do not yet exist. Further, Douglas (2000) has questioned whether there are even a dozen "studies or considerations of how hypertext may transform the way we read or write texts, and, indeed, our whole conception of a satisfactory reading experience" (73). In this assessment, Douglas disregards empirical studies of informational hypertext from the perspectives of interface design and software engineering. But her point is well made, and is still valid in spite of the intervening years since she first made this observation: there exist few examinations of how reading processes may change when readers interact with complex digital genres such as hyperfiction. Members of our research team have worked to fill this significant lacuna in the literature (e.g., Dobson and Miall 1998; Miall and Dobson 2001; Dobson and Luce-Kapler 2005; Luce-Kapler and Dobson 2005; Dobson 2006), but much research remains to be done, and appropriate models of reading are still in need of development.

The Importance of User Studies in the Humanities

The technical world has been slow to realize that users matter, not just in the field of digital humanities, but in broader areas of system design. As long ago as 1971 Hansen called for software engineers to know their users (Hansen 1971), and for the last thirty years the advantages of software projects whose systems are designed with an eye to the user have been well documented (Shneiderman and Plaisant 2005). Yet some are still inclined to assume that users might not know

what they want, and thus it is better not to ask them, in case they answer the wrong question.

If users are integrated into system design at all, it happens late in the process. Typically the user is presented with a late-stage prototype that the designers hope she will like. This can be of limited use, since at this point system designers may be unwilling to make significant alterations. Thus, late-stage user input can only have a relatively minor impact, since the radical redesign of a system costs time, money, and enthusiasm, all of which are often in short supply. It is much easier to make minor adjustments, and get the system into production. In scenarios of this sort, if users lose their enthusiasm for the product, it is they and not the designers who tend to be blamed. As a result, the failure rate of software in the commercial world is still staggeringly high (Dalcher and Genus 2003; Flowers 1996).

In the humanities, as Warwick notes (2004), scholars have too often been branded as digitally unskilled or even backward looking, because they have been slower to adopt digital tools than scientists. This is a fundamentally flawed assessment of the situation: when technologies fit well with what scholars do, they will use those technologies (Bates 2002). Several recent studies of humanities users and digital resources in the UK have found many humanities scholars reporting that they are enthusiastic users of digital resources. However, what they define as digital resources tend overwhelmingly to be generic informational resources, such as library and archive websites, or large online reference collections such as the O-DNB or Literature Online, rather than the kind of digital object that might be compared to a scholarly book (Warwick et al. 2007). At present general information resources are better suited to humanities researchers' needs. If we would like those researchers to use future electronic books, they must be equally fit for the purpose. To produce such a resource we must understand what users in the humanities do, what they like, and what they might like in future.

Previous Studies of Humanities Users

The study of digital resource usage in the humanities is well-researched. As Dalton and Charnigo (2004) show, in recent years there has been a flood of literature about scholars' information needs and seeking behaviours. Although useful recent work on humanities scholars has been done by Green (2000), Talja and Maula (2003), and Ellis and Oldman (2005), much of the literature tends to conflate information seeking and information needs in relation to humanities scholars.

The earliest work on humanities users was on their information needs and patterns of use, and it is only very recently that research has been conducted on their actual behavior in digital environments. Seminal work done by Stone (1982) and later by Watson-Boone (1994) showed that humanities users need a wide range of resources of different ages and types. This is still true in a digital environment, where humanities users continue to need printed materials and manuscripts, the latter implying older materials than those generally used by scientists (British Academy 2005). Humanities scholars also rely on face-to-face information gathering, from colleagues and at conferences. They may also use personal collections of knowledge built up over years of study. They do not necessarily expect to create new data or discover new facts, but reinterpret and re-express ideas, where the expression itself is as important as the discovery (Barrett 2005).

A major theme of the literature about humanities users is that they are not like those in the sciences or social sciences, although many systems designers of electronic resources have assumed that they are (Bates 2002). Humanities scholars are much more likely to use what Ellis has called "chaining;" that is, they proceed by following references that they have found in other literature (Ellis and Oldman 2005; Green 2000). Despite the hypertextual nature of the web, however, such activity is seldom well supported in online environments (Bates 2002). It is also at odds with keyword queries that tend to be the norm for information systems. Oddly, this has been seen as evidence that humanities researchers' techniques are somehow impoverished (Chu 1999). Yet as long ago as the mid-1980s Wiberley showed that humanities scholars constructed searches using well defined terms, but these terms were different from those used by scientists: for example, humanities searches were more likely to include names of places or people (Wiberley 1983 and 1988). Bates' work and that of Dalton and Charnigo (2004) and Whitmire (2002) has also shown that those humanities scholars who use digital resources tend to be demanding of their quality, and are capable of constructing complex search strategies, given appropriate training. Lehmann and Renfro (1991) and Wiberley (2000) suggest that humanities scholars are receptive to technology as long as it demonstrates adequate savings in time or effort.

If we are to design an e-book that humanities scholars may regard as fit for their research purposes, then it follows that we must understand what they do in digital environments, what kind of resources they

need, and perhaps most importantly, what makes them decline to use resources. These questions are not well understood. Except for the work of Bates et al., on the Getty project (Siegfried et al. 1993), very few people have studied what humanities scholars do when they carry out research in an online environment. There is even less concern with what they do offline and how connections may be made between the two fields of activity. Some of these questions are being addressed by researchers involved in the HCI-Book project. However, a great deal more research remains to be done on such areas before we may be confident of designing digital resources that will be of genuine utility to the humanities researcher.

In thinking about how best to conduct a study of humanities users, we have benefited from the experience of two previous UK-based projects. Both of them list Claire Warwick of our research team as a principal or co-investigator.

The UCIS project (User Centered Information Search in context) looked at the ways in which users interact with digital libraries. UCIS studied humanities users and their interactions with information in physical and virtual environments. An important facet of this work was that users were studied in as naturalistic a context as possible, so as to gain a fuller understanding of the nature of their information work (Warwick et al. 2005). Humanities academics were interviewed and observed undertaking their usual research in both physical and digital libraries. The knowledge gained from this was used to derive requirements for features of, and for alternative interfaces to, the Greenstone digital library system.

The LAIRAH project conducted research into levels of use of digital resources in the arts and humanities in the UK. In other words, LAIRAH asked which digital resources were actually accessed, and how often, by users. The project conducted workshops in which users' reactions to different digital resources (both widely used and less used) were investigated.

Both LAIRAH and UCIS found that humanities researchers have very sophisticated information skills and complex mental models of their physical information environment, whether looking for material in a library or an office bookshelf. However, they find those skills much more difficult to apply to the digital domain (Makri et al. 2007). Humanities researchers are very much aware of the affordances, advantages, and problems of various information technologies, but are concerned about accuracy and ease of use for both physical and electronic resources. To address this concern, researchers require

as much information as possible about how such resources are constructed and what they contain (Warwick et al. 2006).

LAIRAH research in particular found that users have very high standards for the content and interfaces of the resources they use, and are easily deterred from using digital resources. Anything that makes a resource difficult to understand—such as a confusing name, a challenging interface, or data that must be downloaded—will deter them from using a given resource. Humanities researchers also find that the pleasure of interacting with the materials that they read, whether in physical or digital space, is an important element of scholarship (Blandford et al. 2006). We therefore face the challenge of designing a digital resource that is not only functional and appropriate for its scholarly purpose, but that is also pleasant to use, and above all does not deter researchers from using it.

Research has also shown that many of those who have been surveyed about their use of digital resources are enthusiasts for the medium. The research for both UCIS and LAIRAH suggests that many scholars do not make significant use of specialist digital resources (beyond the generic information resources described above). Our research must therefore aim to survey such light users and non-users to determine what factors might deter them from using an e-book. This will allow us to gain the fullest possible picture of the user community, instead of one limited to the early adopting enthusiasts who are likely to volunteer for studies on digital resources.

In addition to the above-mentioned projects, two major Canadian undertakings have influenced our thinking. Both TAPoR (Text Analysis Portal for Research) and ECO (Early Canadiana Online/ Notre Memoire En Ligne) have investigated how humanities researchers use online tools, whether text-specific or not. Two major findings emerged. First, scholars are most concerned with accomplishing their tasks; the interface must present things in the context of tasks that they might wish to accomplish, rather than of the tools that might do the job, or the technical details of the text that they might be using. Secondly, graduate students behave differently and have different needs from more senior scholars. These findings highlight the need for the HCI-Book project to identify different categories of users in order to meet their needs; these categories may be more fine-grained than we expect.

3. Features of Books and e-Books

It is difficult to come up with a list of generalized features of the electronic book, because the word "book" refers to two distinct concepts. On one hand, a book is an empirically measurable type of physical object; on the other, it is a powerful and comprehensive type of metaphor. The book, as the phrase appears in terms like "history of the book" or "culture of the book," inevitably simplifies a range of textual materials. Meanwhile the value of these materials (scrolls, manuscript and print codices, newspapers, magazines, fascicles, broadsides, bound quires of manuscript poems, and unpublished archival materials) lies for many scholars in their diversity. Members of our research group study all these materials and more, and yet humanities scholarship still assigns the coordinating metaphor of the book to a number of heterogeneous texts. As a result, in many areas of the humanities the book has of necessity tended to remain an abstract, generalized idea that emphasizes common features over historical particularities. Broad histories of writing tend to group texts into major categories based on what are perceived as as primary technologies. Typically these categories are manuscript (including the shift from papyrus to parchment, and from scroll to codex), print (including the shift from hand-press to machine-press printing), and more recently, digital writing.

The book may be too simple a metaphor to capture all the possible forms of navigable information, but nonetheless it remains a metaphor with potency even in a digital culture. We continue to make cultural investments in structures of information that offer, as features, closure and containment matched with navigability and connectivity. A book has physical boundaries—its covers set its contents off from other books—and yet its text invariably refers beyond those boundaries, such that the book becomes a metaphorical stand-in for the entire bookshelf, library, or archive. Institutional models such as libraries and publishing houses likewise depend upon books as discrete units of manageable content. As bibliographers such as D.F. McKenzie have argued in recent years, the cultural perception of the book as a discrete unit of production is thus at odds with the heterogeneity implied by a term such as "textuality;" yet "textuality" carries none of the symbolic force of the book (McKenzie 1999).

Our use of the term "book" throughout our discussions comes with an appreciation of the term's ambivalence and hidden complexity: "book" is both an ideal and an inadequate term for our objects of

study, whether manuscript, print, digital, or otherwise. The same terminological challenges arise with the term "e-book," which for our purposes means not only the book-like electronic simulations of past decades, but also electronic texts generally, from transcriptions in electronic archives to born-digital electronic texts. With this sense of complexity in mind, we propose to study the features of books and e-books, not by assigning features exclusively to print or to digital textuality, but by identifying the features most at stake in the textual economy that print books and e-books now form together. The most salient of these features include:

1. **Tangibility**, or the capacity of a book to convey information about itself through physical indicators such as size and format. E-texts are often regarded as intangible, or tangible in different ways, but still possess a physicality that should be considered in their analysis (see Kirschenbaum 2004a).

2. **Browsability**, or the book's ability to provide random access through tactile means such as flipping pages to move within the text—a feature amplified in digital text by speed, scale, and browsing tools such as rich-prospect browsing interfaces (Rucker 2003), but not necessarily different in nature from browsing in print.

3. **Searchability**, a feature available in all texts through optical scanning, is aided in print reference works like dictionaries by visual formatting, lemmatization, and alphabetization; it is dramatically enhanced in machine-readable digital text by tools for algorithmic analysis and retrieval.

4. **Referenceability**, a text's propensity for intertextual linking (explicit or implicit), as well as the degree to which parts of a text may be referenced by other texts; in print, the addition of chapter and verse numbers to the Bible made it referenceable on a more granular scale than before, while on the web, the combination of URLs (Uniform Resource Locators) and anchor tags (<a>) allow any link to target a precise point in a web page. More recently, the introduction of persistent URLs of a number of kinds, among them unique digital object identification systems, have added further potential.

5. **Hybridity**, or the composite nature of books as being composed of various discourses, genres, sources, and textual formations; literary miscellanies and anthologies, for example, might include diverse poems of various forms, prose fiction or

non-fiction, and drama by various authors; electronic forms, which are less fixed and definite, increase the possibility and likelihood of hybridity.

In drawing an analogy between printed books and e-books as artifacts, we might also consider how developments in technology enable and sometimes direct changes in the features and characteristics of the book. For example, the advent of columns and printable margins on the printing press enabled marginalia, leading to notes, which eventually became footnotes, and then (as tastes changed) endnotes. In other words, the relationship and means of connection among text, context, paratext (such as an accompanying preface or glossary), and intertext (such as a text with close thematic or allusive links to another)—verbal and non-verbal elements—has changed over time (Greetham 1997; Maclean 1991). Similarly, in the short history of e-books technological developments have enabled new ways of relating and linking texts. The result is a series of changes in the format and features of texts, and in how they are presented to the user. Theoretically, just as changes in marginalia changed reading practices in the early history of printing, so too changes in e-book technology affect reading practices (Slights 2001).

The Uses of Books

The history of the book is the history of how people have shaped the intellectual tool of writing in order to make it more efficient, more versatile, and easier to use. In Greece, before 600 BC, there was no clear decision regarding whether writing should go from left to right or from right to left (or even boustrophedon, where the direction of writing followed the path of the plough in a field and lines had to be read in alternating directions). Romans read aloud, or were read aloud to by slaves.

Gradually, the text departed from the linearity of the spoken word and became organized in a visual way, giving more control to the reader. The invention of the codex in the first century AD was a major turning point. However, though the codex was vastly more efficient than the scroll as a medium for text, it took four centuries for the codex to completely replace the scroll and to evolve from an essentially linear format to a tabular one, gradually giving the reader more control over the pace and form of reading.

Many incremental advances were necessary in order to make the book what it is today. In medieval monasteries, monks in the

scriptoria were to read silently in order to avoid distracting their colleagues. This practice helped to shape the book as a visual entity adapted to efficient reading. Reading itself was for many centuries mainly conceived of as a ruminatio (pondering at length). In the eighteenth century, "extensive reading" became pervasive with the advent of newspapers and magazines. This new practice made browsing vast amounts of textual material a legitimate form of reading.

Historically, the medium used for displaying text is important because it determines which operations are easy for the reader to perform. Each medium puts varying emphasis on the visibility of the text and on the reader's interactions with that text. In this regard, text is the nexus of a fundamental tension. On one hand, it is a product of language which is of limited duration, and which depends upon a syntactic organization of the words. On the other hand, since it is a visual entity, text may also surpass its linear bounds and play upon the resources offered by layout, typography, and colors in order to create other types of meaningful events.

The Uses of e-Books

The discussion above outlines a range of features pertaining to books and e-books. Tangibility is strongly associated with the book as a physical artifact, while hybridity is strongly associated with digital texts. Browsability, searchability, and referenceablility are also dramatically enhanced in digital media. As we examine this list of features, it becomes clear that e-books are particularly well-suited to the needs of researchers (information seekers and those engaged in text analysis). It is not surprising, then, that research and development efforts related to e-books in the humanities have tended in the direction of developing searchable digital archives and databases, establishing a method of encoding digital texts to ensure compatibility among archives, and establishing a new knowledge economy related to scholarly electronic publication.

Other research areas concern the affordances of digital media for collaborative knowledge production, and for the literary arts. The first of these, collaborative knowledge production, is increasingly facilitated by various emerging forms of social software, such as wikis and weblogs. The most obvious example is Wikipedia, a popular web-based free-content encyclopedia that maintains a relatively open policy regarding contributions. Wikipedia's model for contribution clearly interrogates notions of authorship and intellectual property rooted in print culture. This is not a new model: it is one that was

displaced by the formalized diffusion of academic writing, which saw its genesis in seventeenth-century Europe. Before our current methods of inquiry and knowledge diffusion were shaped, notes Siemens (2002), knowledge exchange and the advancement of scholarship were facilitated in large part by personal dialogue and the circulation of private manuscripts and correspondence, with the emphasis on "ensuring that valuable ideas circulated and became part of growing, documented, bodies of knowledge" (Siemens 2002, 3). Siemens suggests that the argument for proceeding ad fontes (going back to the sources) is compelling in its suggestion that we might turn again "to earlier models of scholarly exchange … and consider their possible relationship, even if only metaphorically, to what we now refer to as 'new' types of scholarly exchange that are made possible by the electronic medium" (3). The impact of such a model for future knowledge production and diffusion activities in the humanities should be examined.

The second research area referenced above, electronic literature, refers to "works with important literary aspects that take advantage of the capabilities and contexts provided by the stand-alone or networked computer" (ELO 2006, ¶2). Electronic literature includes genres such as hypertext fiction, reactive poetry, and blog novels. Older forms, such as hypertext fiction, have their roots in text adventure games (such as Will Crowther's 1975 *Adventure* and Infocom's *Zork*) and Bantam's *Choose Your Own Adventure* series of children's books. Emerging genres such as reactive poetry, on the other hand, intermingle literary arts and multimedia design. Often presented as Flash files, works in this latter class employ animated images and text accompanied by sound, in an effort to produce visually dynamic pieces. We will study further the nature of these emerging genres, and how readers interact with them.

Digital Archives, Sustainability, and Access

Basic strategies of preservation, upkeep, and ongoing access in the print medium have changed little since the arrival of the mechanically-produced book, though the details of those systems have undergone changes when technological developments enabled certain strategies. Change has also been driven by need: some examples are the emergence of classification systems at the turn of the century to handle the exponentially-increasing numbers of books, and the movement toward "universal" libraries and national collections.

Many aspects of these basic strategies are transportable to the electronic medium, but most require modification to take into account three things:

- the pace of technological change, which manifests itself, for example, in old electronic materials that already require special technological considerations to remain functional;
- the fact that electronic materials are much more malleable than those in print (so versioning becomes a more complex issue); and
- new basic functionality and interoperation possibilities that computation allows us to consider and enact on electronic materials—all, again, from the perspective of the user.

Current electronic databases are deploying a number of economic models to ensure ongoing support and access. The following list of successful examples should be investigated and emulated by any humanities project concerned with sustainability and access:

1. **Membership Model**: JSTOR, which provides database access to journal archives, was established with major Mellon Foundation funding, and is sustained on a membership basis. Memberships are available for institutions alone, with libraries and schools paying an initiation fee along with an annual fee, based on size. Access is restricted to individuals who belong to a member institution.

2. **Institutional Sponsorship Model**: Representative Poetry Online is, in good measure, the work of a local champion, Ian Lancashire, with sponsorship from University of Toronto's Office of the Provost, the Faculty of Arts and Science, the Department of English, and the Library (which hosts the database). This is seen as a public service, and the contents are open-access.

3. **Institutional Sponsor/Major Grant Model**: The principal archive in physics, arXiv.org, is sustained by funding from Cornell, which hosts it as a matter of public service, offsetting some expenses with grant funding from NSF. The contents are open-access.

4. **Subscription/Open Access Model**: Highwire Press, operated by the Stanford University Library, provides access to close to 1,000 journals principally on the basis of personal or institutional subscriptions to individual journal titles. These

journals provide around a million articles free of charge, some offering immediate open access, others using the model of delayed open access. The delay between for-pay publication and free release runs from 6 to 24 months.

5. **Endowment Model**: The *Stanford Encyclopedia of Philosophy* is seeking donations to generate sufficient funds to manage the continuing development of the project: <http://plato.stanford.edu/fundraising/>. The contents are open access.

6. **Cooperative Model**: Something of an ideal, yet to be fully realized, that combines A and E, insofar as membership among principally interested libraries comes together to support and contribute in-kind to sustaining a cooperative database that provides open access to all.

4. Aspects of the Book

Material Aspects

Understanding the material aspects of the book is a crucial pre-condition to developing e-books. The very name "e-book" alerts us to this fact. When the movable-type press created the opportunity for mass-producing texts, its operators aimed to make books look like manuscripts. It took decades for books to become something other than mechanically produced manuscripts. E-books could suffer from a similar developmental lag unless we attend closely to the processes and products that have, over the 500 years of print culture, established the printed book as a standard physical medium. Revising the printed book into electronic format while keeping to the conventions of print is not the best use of resources. Readers will be better served if we imagine a radical departure from the printed volume. To help ourselves make this paradigm shift, we might juxtapose the set of practices carried out in producing printed texts and the set of practices necessary for the production of electronic texts.

To discuss the material aspects of the book, then, is to discuss not only paper, ink, formatting, binding, and content, but also the processes that are required to produce a book. Before the paper can be folded to create a folio, quarto, or octavo, it must first be manufactured as a broadsheet or a roll. Before early modern paper could be manufactured, a slurry had to be prepared from pulped linen, lime, and water. Today, wood fiber or recycled material must be pulped in a process that is part mechanical and part chemical.

Historically, paper, vellum, or papyrus was used to enable the circulation of a written text, each with its own production processes. Before these media, there was wax; before wax, stone dash (and it should be noted that this skeletal history is decidedly Western in its slant).

Changing means of production fundamentally alter cultural notions of authorship, readership, and literary form. For example, the substantial cost of the materials required to produce the early printed book necessitated selective publication according to any number of criteria (content quality, author's status, marketability, and so on). The cost of the end product also determined the readership of printed materials: as technological developments reduced production costs, readership expanded and changed, as did the types of books that were produced. With the advent of machine-made paper, for example, periodical and newspaper publication proliferated in the nineteenth century (Rose 1995). Another consequence of reduced printing costs was the emergence of the so-called popular press. In the late sixteenth and early seventeenth centuries, small and cheap books, chapbooks, pamphlets, and broadsheets brought a new demographic to the book market. This in turn enabled new means of influence upon the reading public in sociopolitical matters such as religion, and new forms of writing to meet the interests of an increasingly diverse reading public (Chartier 1994; Spufford 1982; Watt 1991). Conversely, electronic publishing began with a cost advantage over print production. Many forms of e-books are already comparatively cheap to produce, and they do not require a large infrastructure of dedicated printers, editors, and publishers. Consequently, the practice of "self-publishing" that was once regarded with a certain amount of stigma is now increasingly legitimized. Selective publication is no longer necessitated by technology or economics. This new means of production forces a reassessment of cultural assumptions about authorship, readership, genre, accessibility, and usability (including quality control and censorship).

Symbolic Aspects

The concepts of e-book and electronic text can be defined in many ways. Most traditional definitions are based on the functional and material characteristics of e-book technology. From this perspective, e-books are defined by the following characteristics: they are electronically searchable, they are dynamic in terms of content updateability, they are adaptable in terms of cross-platform

portability, they are physically portable, they are user-interactive, and they are non-sequential—provided they use hypertext technology.

E-books are currently available in a variety of formats; most e-book formats allow for multilinear navigation. So, hypertextuality seems to be a characteristic feature of e-book technologies. However, most of the work on hypertext has considered it from a technological perspective: most of the time hypertext is explained in terms of nodes, links, markup language, and so on. Hypertextuality has a substantial impact on the symbolic aspects of the way we read texts, but very few research initiatives have tried to formalize this, especially from a semantic point of view.

Hypertextuality may be an important feature of e-book technologies, but hypertexts themselves are heterogeneous, and can be defined from many perspectives. For instance, the concept of hypertext can be defined as a hardware- or software-based computer technology, or as an abstract textual structure. In the lexical database Wordnet, for example, hypertext is defined as machine-readable text that is not sequential but is organized so that related items of information are connected. Although incomplete, this definition is a very good example of hypertext as an abstract textual structure.

Formal hypertext models can offer tools for an understanding of hypertextuality's main characteristics, but they can also neglect the importance of the user (the reader in the context of e-book technology), particularly when the objective is to understand and model the meaning of hypertext. The challenge is to model or formalize both e-books and hypertextuality, as the latter is one of the former's main characteristics, at the same time taking the reader into consideration as an interpretative agent.

Formal Aspects

Possibly the most immediate way in which a reader interacts with a book is through its material, or formal, aspects. Readers tend to enjoy the feel of the pages, the cover, and the design of the book. These formal aspects are important to consider when imagining how to transfer the reading experience to a digital environment.

One formal aspect of the book is that it is more than can be consumed in a single visual event. This distinguishes a book from a pamphlet, which could conceivably be seen in its entirety by a reader holding the pamphlet open while standing in front of a mirror. The reader could read the inside two pages of the pamphlet, and could at least see and possibly read the backwards-appearing front and back

covers of the pamphlet, without having to manipulate the pamphlet further. A book has a more complex form—one that, no matter how simple the content, can never be seen all at the same time.

Translated into the electronic environment, this distinction between book and pamphlet is comparable to the distinction between a single web page and a website (which is a collection of web pages). However, a book could conceivably be transformed into a single web page. There is nothing inherent in the form of the book that demands it be translated into a multi-page website.

This may suggest a genuine problem. If the e-book is held to be inherently more complex than its print counterpart, then the above distinction between book and pamphlet would seem to provide support for that position. Further, if one is inclined to view the print book as less complex, then one might point to the inadequacy of the analogy (that is, the relationship between web page to website is not comparable to the relationship between pamphlet and book).

In addition to length, there are other formal properties typically associated with books. There is, for instance, the cover—the upper and lower boards, and the spine—which plays a role in communicating the contents of the book. We say that a book cannot be judged by its cover precisely because the natural human response to covers is to use them to judge books. Designers use this tendency to communicate various kinds of information, such as genre, author and title, publisher, and often something about the narrative content. Immediately beneath the cover are the pages that constitute the front matter. These reiterate in more standardized formats some of the suggestions made by the cover, and provide additional information such as the metadata required for cataloguing, as well as an outline of the contents.

The body of the text contains significant layers of design, such as white space and page numbers, that exist in addition to the text itself, but make the act of reading more efficient. White space in the margins for instance, allow the book to be held without obscuring any text, and provide room for binding. Sequential page numbers allow the pages to be accessed non-sequentially with the aid of a table of contents or an index. Page numbers also serve to allow the reader to track progress through the book.

Experimentally developed over centuries, these conventions are now subject to remediation in the digital environment. We have an opportunity to reconsider them, and to decide which are important enough to keep and which have become irrelevant. We also have the chance to extend and elaborate on the features that are most valuable,

reconceiving them in ways impossible in print. Electronic searching is one such enhancement; hyperlinking is another.

5. Prototyping

Interface

It is in some ways unusual to discuss the book as though it were an interface, but the term begins to seem more appropriate when we shift our focus into the discussion of the e-book. An interface is the intermediary technology between a set of services or functions on one side, and a person on the other. For the conventional book, the service consists of the text and its apparatus, and the interface—of paper, ink, binding, typography, and so on. The set of services provided by the e-book has the potential to expand in a wide range of directions, from the inclusion of more sophisticated hyperlinks, to audio and video clips, to metadata of various kinds, to data mining processes and their results, to visualizations of information. Ruecker (2005) points out, for instance, that the e-book's table of contents has the potential to develop into a sophisticated research tool, where the reader can add or subtract elements of interest in order to obtain a dynamic overview of the contents.

Our research into e-book interfaces will include developing prototypes[3] in several directions, from content visualizations, to overview displays coupled with tools for manipulating the displays, to experimental browsers that allow exploration of both collections and content. Our research questions primarily involve remediation and new affordances. That is, what elements of print books are valuable or even indispensable in the new digital context, and what new functions are significant enough to readers and researchers that they should be

---

[3]   A prototype in this context is an interface or visualization that can be seen to actualize theoretical suppositions, so the theory informing the creation of the prototype can itself be tested by having people use it (as per Rockwell and Sinclair [2007]; see also McCarty's discussion of modeling in this context [2008]). Computational prototypes such as those we propose to be part of our work are distinct from production systems in that the prototype focuses chiefly on providing research-pertinent functionality within a larger framework of assumed operation (perhaps within the visual depiction of a fully-operational system, or perhaps in understood isolation from it); production systems require full functionality, and they are often derived from multiple prototyping processes.

developed? Furthermore, how should the conventions of reading and scholarship influence the design of the e-book form and function?

We have already been making progress in a number of these areas, with projects for dynamic text playback, for nuanced browsing of XML collections and files—the Mandala browser (Cheypesh et al. 2006)—and for making data mining tools accessible to humanities scholars (Ruecker et al., 2006). In addition, members of our team have been working on online tools to support editors and readers of scholarly editions (e.g., Best, Galey, Werstine), and we are working with a wide range of researchers who are actively editing such editions (e.g., Flanders, Schreibman, Siemens).

Interface work often involves technical software and hardware development that are not research contributions, but that are needed to answer new research questions. Our methods will help us to obtain a useful academic result without engaging in extensive development activities. We recognize that the research life cycle moves from theory and sketches to interactive prototypes to development projects. By coupling the design research strategy of iterative constructive diagnostic research (Sless 1997) with approaches based on the affordance strength vector model (Ruecker 2006), we hope to maximize exploration of novel research questions at the early stages of the life cycle, while avoiding getting drawn into expensive, time-consuming development tasks.

Preservation

Digital texts offer opportunities for storage, distribution, access, analysis, and forms of interaction and use that are not possible with print texts. However, the same features of malleability and replicability that make digital text so versatile also increase its vulnerability. Unlike printed books, which are self-sufficient textual objects, digital texts are machine- and software-dependent, which means that our ability to use them is subject to transient and unstable technologies (Beagrie and Jones 2002). Digital texts are more easily altered than print texts, so it is more difficult to ensure their integrity and authenticity (Lavoie and Gartner 2005). The vulnerability of digital texts, and the implications for long-term use of textual artifacts, are widely recognized in the library and archive communities. International initiatives in North America and elsewhere exist to study and develop guidelines for digital preservation (Giaretta 2006; Waters and Garret 1996). Examples of such initiatives are PLANETS, CASPAR and the Digital Curation Centre. Their goal is not only to preserve the texts themselves, but

"to preserve the information integrity; that is, to define and preserve those features of an information object that distinguish it as a whole and singular work" (Waters and Garret 1996).

The main consideration in trying to preserve the texts themselves, and to ensure that they are available to future generations, is technological obsolescence. A complex layer of technologies mediates between digital texts and users. These technologies—storage devices, data formats, interaction and viewing applications—have a limited life expectancy (Lavoie and Gartner 2005). Digital storage media break down surprisingly quickly; versions and types of applications and systems change rapidly due to technological advances (Beagrie and Jones 2002; Lazinger 2001). Some approaches to this problem have been developed, and they need to be considered in any preservation strategy.

To combat instability, stored data need to be regularly backed up and may be periodically "refreshed," that is, copied onto fresh storage media. Large data archives should be stored on multiple sites and in more than one format (Linden, Martin, Masters, and Parker 2005). To combat obsolescence, both emulation and migration can be used (Hedstrom, Lee, Olson, and Lampe 2006). Emulation is the use of current technology to mimic older systems needed to access legacy data (Granger 2000). Migration is the practice of periodically moving data onto new systems to keep up with current technology. Both practices can be effective; however, it is essential that they be well-documented and use clear standards and methods, because both have the potential to change the digital artifacts, either through data corruption or changes in the structures or capabilities of the systems.

This susceptibility to change is a serious challenge for digital preservation initiatives, as it is difficult to ensure that texts are actually what they claim to be, and that they have not been altered in any way. Even if interactivity and changes in texts are desirable in current and future use scenarios, it is important to be able to preserve earlier versions of texts, and understand how changes are made over time. The principal approach to this problem has been to develop models and standards for the archiving and description of digital texts. The Open Archival Information System (OAIS) reference model is an ISO standard (2002) that describes the environment, functional components, and information objects of systems designed to preserve digital materials. The key concept in this model is that objects are information packages, which include both the content and the associated representation information needed to make the object

understandable (Caplan 2006). Although not explicit in the model, this type of framework can be used to design "trusted repositories" that preserve the integrity of digital texts. Implicit in OAIS and other related standards (such as RM standard ISO 15489) is the need for preservation metadata, "information about the technical environment in which records are created and exist" (Duff 2003). Preservation metadata serve as documentation for digital texts and may cover a number of areas:

- Provenance—custodial history and changes in ownership over time;
- Authenticity—information to show that a text is what it claims to be and has not been altered (fixity);
- Preservation—any action taken to preserve the text, and possible impacts on its functionality or appearance;
- Technical—any hardware or software requirements needed to render and/or use the text; and
- Rights—recorded intellectual property rights. (Lavoie and Gartner 2005)

Nested within the broad OAIS model are more specific frameworks for preservation metadata. The Metadata Encoding and Transmission Standard (METS) provides a coding and structure standard for metadata elements, useful in shaping OAIS information packages. METS is an empty container that can include structural metadata for complex objects. It is also a wrapper for administrative, descriptive, and structural metadata, and it defines a metadata exchange syntax (Day 2005; Duff 2003). A set of core metadata elements that can be used in the METS framework is defined by the PREMIS DATA Dictionary (2005), which also provides guidelines for the development of preservation metadata schemas.

However, no preservation metadata schema can be developed in the abstract. One of the major conceptual challenges associated with digital preservation is to determine what needs to be preserved, given that digital texts can be defined at different levels—as data bits, text, structure, format, intellectual meaning, interpretations and so on (Caplan 2006). Furthermore, the decision as to what features determine the integrity of a text is not based on universal standards, but is dependent upon current and future user and task contexts (Ross 2002). For example, some digital texts are fluid by nature, and so a fixed ordering of the parts may not be necessary to preserve their integrity (Steemson 2002). Thus it is important that digital

preservation issues be considered at the outset of any project involving digital texts, so that the creators of the texts and the infrastructure may participate in determining the key elements for preservation. These decisions should be developed in the framework of a data model capable of accommodating diverse types of primary sources and preservation metadata. The data model should be based on an understanding of user roles, tasks, and use contexts. The technology infrastructure should be capable of protecting the integrity of digital texts and supporting the concept of self-documenting objects by providing a mechanism for both automatic and manual creation of metadata.

Digital preservation issues should be considered from the outset of the HCI-Book project in order to develop conceptual and physical models of the e-book and e-book repositories that are consistent with developing standards in archival information systems and preservation metadata. In particular, we need to take a broad perspective on identifying the features of an e-book that are essential to its integrity, and begin to think about an infrastructure capable of documenting and preserving those features. This should be done in coordination with current initiatives, such as the PLANETS project, which are focusing on these questions.

Encoding/Algorithm Issues

A history of writing is also a history of textual encoding. Every writing technology—from shapes traced into clay tablets to HTML tags used on web pages—conforms to established rules (that may be more or less explicit), without which recorded language would be incomprehensible. Humanists, as students of the history of human communication, have always been at the forefront of developing systems for describing textual artifacts and establishing common practices for their study. As such, it is not surprising that humanists have played such a prominent role in the development and standardization of digital encoding practices, from early uses of punch cards for encoding (shortly after World War II) to the more recent development of markup languages such as SGML and XML. One of the most sustained efforts has been the Text Encoding Initiative (TEI), which has been striving since 1987 to establish flexible standards for digitally encoding information about an infinitely large range of textual objects. The diversity of extant texts explains why the TEI's efforts are ongoing and as needed and relevant today as ever.

Digital humanists working on text encoding are confronted with two specific sets of challenges:

1. how to digitize texts—that is, how to transcribe information about print, manuscript, and other texts into a digital format (with all of the risks inherent to any form of transcription); and
2. how to encode information in ways that are useful to both the human reader and the computer.

The first set of challenges, regarding strategies of digitization, requires researchers to make (often difficult) decisions about what types of information to preserve when passing from print (or manuscript, stone, etc.) to a digital format. Manuscripts and printed texts have a conceptually limitless amount of information associated with them—the recorded linguistic signs, the physical traits of the object (size, materials used, preservation level), and metatextual information (date of creation, publisher, sales). Digitization imposes compromise and sacrifice: it is not a question of whether information is lost, but of how much. Few attempts have been made to produce electronic editions that encode a wide spectrum of information about print objects. The HCI-Book project will be able to contribute substantially to efforts to standardize encoding practices of textual objects as multidimensional and multimedia objects, and not just as sequences of linguistic signs. Moreover, we anticipate that the ability to encode a wider array of features found in textual objects will create new opportunities for representing and studying them.

The question of designing encoding schemes that may useful to both humans and computers has been raised throughout the development of digital encoding formats and markup languages. Since most markup is performed manually (graphical editors are more useful for markup languages with relatively small vocabularies, such as HTML), there has no doubt been a historical tendency to favor human readability of markup over computability. In recent years, however, that trend has somewhat reversed. For instance, eXtensible Markup Language (XML) is in many ways an attempt to create an encoding scheme even more constrained than SGML (Standard Generalized Markup Language). XML saves processing, analysis, and rendering tools from having to anticipate and deal with a wide array of variations and exceptions. Moreover, a new generation of analysis tools designed and developed by the digital humanities community is prompting a rethinking of how text encoding can be better adapted

to the demands of algorithmic processing. (Examples of such next-generation tools may be found on the TAPoR website.)

The HCI-Book project will be well-positioned to contribute to the collaborative process of adapting existing text encoding practices for a range of purposes that relate to creating rich electronic reading environments. These environments must be able to represent information relevant to traditional textual critics (interested in the material characteristics of textual artifacts), and to provide a variety of integrated analytic tools that can exploit the underlying encoding and facilitate interpretation.

Textual Analysis Tools

Several studies have indicated that people generally prefer reading materials in print rather than on a screen (for reasons of comfort, portability, and cost, among others). Advances in display technologies have narrowed the gap since earlier days of flickering CRT monitors. Still, it remains that computer screens generally provide content at a much lower resolution than print. While ongoing efforts by several companies to produce electronic paper are promising, one does not have to wait to experience several of the benefits of working with electronic texts rather than print texts. Electronic texts can be readily indexed, searched, browsed dynamically, reconfigured and analyzed. Those are some of the benefits that the HCI-Book project will strive to fully exploit in the design and development of electronic text interfaces.

Despite some preliminary research on user perceptions of text analysis tools, much work remains to be done in studying the practices of readers of electronic texts, particularly of humanities scholars. The design and conception of analytic tools for electronic editions can be approached from two directions: 1) the needs and expectations of users (based in large part on current models), and 2) the innovation of new forms of analytic tools (which would be difficult for most users to imagine, because they would represent an entirely new paradigm of interacting with texts).

There are several ways in which analytic tools can be associated with texts:

- as visual cues embedded in the text itself (typographic indications; for example, color-coding of documents in HyperPo);
- as links or buttons in a text's margins that spawn results (text pushing; for example, "Reading Tools" in Open Journal System); and

- through a tool's ability to access a given text (text pulling; for example, TaporWare Tools).

It should be emphasized that the HCI-Book project is chiefly concerned with conducting research on how to integrate analytic tools and texts, and with determining which tools in particular are most appropriate for users of electronic texts. The focus will be on developing prototypes that could subsequently be expanded and/or commercialized in a different context.

We begin in this work with the assumption that text analysis tools are, in effect, an interface to the book—an unusual type of interface. Some of the questions we want to answer in this vein are:

1. How do people use text analysis tools in reading or studying a book? What informal ways of analyzing a text have emerged, whether in popular culture or the scholarly community?
2. What are the expected and broadly understood text analysis functions? What types of tools do most scholars understand, and how do these functions appear in reading environments? For example, we can hypothesize that most scholars understand searching and KWIC (keyword in context) displays; what other analytic functions do scholars understand?
3. What could the analytical interface to the book be? What new types of analytical interfaces are emerging? In particular we want to look at visual interfaces to the text and emerging paradigms.

To pursue these questions we imagine the following activities:

1. A visual survey of text analysis tools and their interfaces. The survey's result would be a visual thesaurus and a topology of interactive analytical interface elements.
2. A survey to find out what sorts of text analysis needs currently exist in the scholarly community. This would be a online survey, possibly followed up by phone interviews. It will try to tease out emerging needs and functions, which should also draw out emerging reading methods independent from tools that can be computer supported.
3. Prototyping of new analytical interfaces. Working within an interdisciplinary context that includes computational linguists, interface designers, and text analysis tool developers, the project could develop imagined interfaces that respond to activities 1 and 2.

Reading Tools and Contexts

Our aim is to build and test reading tools that are designed to support and extend readers' use of databases and texts. These tools would do so by facilitating specialized analytical approaches to the resources, and by connecting those resources to related materials both within and outside of academic circles. The reading tools would be used to facilitate the link between the text or document at hand and its related materials, with some examples being primary sources, historical contexts, related scholarly literature, glossaries, and educational materials. The tools would use the documents' metadata to automatically bring precision and focus to the search of related materials, some of which may be restricted to members of research libraries, others open for public use.

The addition of tools to a new reading environment for electronic texts, with the option for readers to use or remove these tools, is an aspect of such an environment's design. New reading environments will need to incorporate educative functions for less experienced readers (e.g., students, interested public) in the specific domains of the resources developed. These environments must offer opportunities for readers to engage with texts through the tools, and to connect the individual texts to larger realms of learning. They must also provide a way for readers to contribute back to the texts and to their fellow readers through a process of participatory critical engagement, with readers controlling what they encounter and configuring the nature of their reading environment, always with the possibility of going farther than expected.

Socially Prominent Electronic Media and Platforms

The web browser continues to be one of the most significant socially prominent platforms for electronic text, reflecting and guiding assumptions about texts and reading. With the rise of the World Wide Web, text-based browsers such as Lynx, Mosaic, and early versions of Netscape became the most common venues for everyday users to experience hypertext (contrary to some predictions that hypertext fiction or digital library interfaces would gain primacy). As browsers began to support a range of media types beyond linked text alone, authors of web pages could integrate images, sounds, movies, and animations into electronic texts without the need for advanced coding. As Netscape Communicator and Microsoft Internet Explorer became the dominant browsers of the late 1990s, their support of

scripting languages meant that designers had access to an electronic text publishing system, combined with a multimedia interface, combined with a reasonably advanced programming environment, all contained within software distributed over the web for free. Most browsers supported Javacript and Java; in addition, Internet Explorer supported scripts in the proprietary language Visual Basic. Although the prototype-based language Javascript is sometimes seen as a limited language compared to the class-based language Java, both are now used to deliver application-level software via web browsers. For example, the Ajax architecture uses Javascript to integrate several web technologies in web applications that use the browser as their main venue—quite different from the simple text delivery for which browsers were originally conceived (see Garrett).

Of particular consequence to computing humanists is the browser's handling of interface and functionality, normally viewed as separate components elsewhere. The interface and the underlying functionality that manipulates the data may be closely integrated in a browser-based application, written in the same language (Javascript) and located in the same files. Browser-based tools for software building such as Ajax are not the first to offer this integration, but Kirschenbaum suggests why this kind of integration is important: with much existing technology, "from a developer's perspective, the interface is often not only conceptually distinct, but also *computationally* distinct. … [I]t wasn't until the comparatively recent advent of languages like Visual Basic that it even became practical to program both a user interface and an application's underlying functionality with the same code" (2004b, 524-25, emphasis in original). For this latter point Kirschenbaum draws upon Carroll, who notes that in past research on user interface software and tools "an early objective was the separation of the user interface and application functionality into distinct layers. This approach modularized the user interface in user interface management systems […]. However, layering entrained limitations on the granularity of user interface interactions" (Carroll 2002, xxxii). We can place Ajax within the recent general developments that Carroll describes: "Current approaches favor developing user interfaces and functionality in the same language, either in new languages invented for this purpose, like Visual Basic, or through extensions to standard languages for implementing functionality, such as libraries and toolkits for C++ or Java" (Carroll 2002, xxxii). Perhaps the most important term here is "granularity," which implies a level of complexity in text-reader interactions that

approaches literary scholarship's treatment of form and content. Such developments merit the attention of computing humanists as well as web developers.

Browsers are also significant to humanities computing because of their implementation of the OHCO model of textual structure (that is, text as an ordered hierarchy of content objects). The OHCO thesis and the debate surrounding it are usually associated with text encoding (see DeRose et al. 1990; Renear 1997; Schreibman 2002; and Buzzetti and McGann), but the standard Document Object Model (DOM) of web browsers also assumes an underlying tree structure in all texts, where logical nodes cannot overlap. The DOM, as implemented in Explorer, Firefox, Safari, and most other recent browsers, is a programming interface to documents assumed to be pure logical structures, and thus provides many web technologies (such as Javascript and XML) with a shared system for manipulating parts of a document. Without the DOM, browser-based web applications would be impossible to implement.

The DOM is a noteworthy aspect of browser development because of the trade-off it represents to computing humanists: its manipulable logical structure for text allows designers to further integrate functionality with interface in a manner that Kirschenbaum suggests is consonant with humanistic ideas of integrated form and content. Yet the DOM can only provide this structure by embodying a model of text that is far from epistemologically neutral, as the OHCO debate has shown. While most of the discussion of the OHCO thesis has focused on text markup, a surprisingly small amount of attention has been given to the ways in which the OHCO model stands behind every single web page—including some of the very web pages that refute the OHCO thesis (see Buzzetti and McGann, and the online version of McGann's "Rethinking Textuality": <http://jefferson.village.virginia.edu/%7Ejjm2f/old/jj2000aweb.html>). Web browsers neither validate nor disprove the OHCO thesis and the positivistic view of textuality it represents; browsers do, however, suggest that questions of fundamental importance to computing humanists are woven into even the most ubiquitous of software applications.

Various plug-ins have made the browser more reader-friendly, notably Adobe's Acrobat plug-ins for reading PDF (Portable Document Format) files that offer the look and feel of the printed page in addition to various modes of navigation, powerful search possibilities, highlighting and commenting capabilities. Acrobat has been integrated with a plug-in from Macromedia that gives a

sophisticated rendition on screen of the turning of the page, making the reading of the digital text closer to the actual experience of reading a book. For an example, see <http://www.lemonde.fr/> and click on the link under "Journal électronique." A similar development is the extension system of tools for Mozilla-based browsers such as Firefox. (See the Mozilla Developer Center portal for extensions: <http://developer.mozilla.org/en/docs/Extensions>). Pertinent development proceeds apace in these areas at terrific speed.

However, we would like to implement new metaphors in order to give users more control on their reading activity. One conspicuous issue that has not yet been addressed in scholarly work is "continuous reading," briefly described above, which is exemplified by the novel whose reading may continue over many days or even months. Some other metaphor-based tools that would be useful to develop in the HCI-Book project are:

- a library metaphor, as a complement to the desktop metaphor for organizing reading activities and maintaining visible the documents that are being read or that the user plans to read soon. At least two types of bookmarks should be available; various intuitive marks should make it easy to spot the most frequently used;
- a visual tool, which should reveal the importance of a document, how much has been read and how much is left to be read; and,
- a series of icons indicating possible ways in which the text might be reformatted, depending on the type of text and the goals of the reader. In the printed world, there is a variety of formats, each of which accommodates the nature of its contents; a newspaper is not formatted like a book, and a scholarly book is not formatted like a novel.
- Similar metaphors, and others, have been adopted successfully in projects such as the British Library's *Turning the Pages* project and beyond.

Reading Culture

Early discussions of new media from the perspective of literary theory devote much time to a consideration of how reader and writer roles are modified in the fragmentary, multilinear writing spaces of the digital medium. Such discussions frequently propose that engaging with networked texts requires more active participation by the reader. Some of these same discussions conclude that technology is destined

to improve the experience of literary reading. For example, noting the roots of electronic literature in computerized text adventure games (also called interactive fiction or IF), Bolter highlights the game-like quality of textual spaces where reader–players fashion their own paths through documents (2001, 126 ff). Bolter also argues that electronic texts are like scripts or musical scores that readers must enact, or perform (173). Douglas, following a similar line of thought, calls on the metaphor of dance to describe electronic fiction: "a work of hypertext fiction can act as a blueprint for a series of potential interactions … a dance choreographed by an absent author" (2000, 23).

Of course, according to certain schools of thought, readers have always been active "players" because they necessarily reconstruct texts within their own worldviews. They do this by reading themselves into narratives, drawing inferences in order to fill in temporal, spatial or causal gaps, and so on (Rosenblatt 1938/1968; Iser 1978). And yet, clearly hypermedia extends such processes in important ways. Readers of IF, for example, must do more than reconstruct a narrative that exists in a predetermined order; they must engage in an activity of construction and reconstruction, both determining sequence and filling perceived gaps in meaning. The result, according to Landow, is "an active, even intrusive reader" who feels a sense of agency because hypertext has infringed "upon the power of the writer, removing some of it and granting that portion to the reader" (1997, 90). This scenario leads Landow (1994, 14) and others (e.g., Rosenberg 1994) to coin a new term for the act of engaging with electronic texts: "wreading."

This neologism of course recalls earlier metaphors promoting the idea of reading as an act of writing. Barthes' notion of lisible and scriptable texts tends in this direction (Barthes 1974), as does the Derridean doctrine of écriture (Derrida 1976). But although Derrida sees the notion of reading as a form of writing and rewriting (as in, for example, "Plato's Pharmacy"), he also cautions that such assertions are wanting in and of themselves, for it is important to take the discussion a step further and to consider exactly what kind of "writing" reading is (Derrida 1981).

In spite of such calls for more careful attendance to the nature of reading–writing relations, the subject has been largely overlooked by hypermedia theorists, who often promote metaphors suggesting the conflation of the two processes, but who fail to examine whether or how those metaphors play out in practice. Questions regarding the interdependencies of reading and writing, what sort of cognitive

processes each demands, and how the two might effectively be combined to promote learning have long been a subject of study among literacy researchers. Fitzgerald and Shanahan (2000) provide a useful overview of such research, detailing the traditional separation of reading and writing in North American school curricula, movements toward a more integrated literacy curriculum, and approaches to the reading–writing relationship from various perspectives. They focus particularly on research examining shared knowledge and cognitive processes between reading and writing. The results of this research are anything but definitive. Studies dating back 75 years in which students' reading and writing abilities are compared have demonstrated only moderate correlations between the two processes; contrary to popular wisdom, good readers are not always good writers, and vice versa (Stotsky 1983). Fitzgerald and Shanahan (2000) also point to studies of various groups, including individuals with brain injuries, which reinforce the separability of the two processes. Particularly interesting in this regard are documented instances of individuals who suffer partial aphasia following strokes or other brain injuries. Such individuals may be able to be able to write, but not to read their own writing, or to speak, but not to understand speech, and so on. Neuroscience delineates a range of highly specific language disorders (aphasias), each of which is tied to injuries to or lesions in distinct areas of the brain (Ganong 2003).

What is clear in all of this is that reading and writing, while related, are cognitively and experientially separate. Metaphors conflating the two fail to consider the complexities of both processes, the phenomenological and cognitive differences between the two, whether one or the other process might be better for facilitating acquisition of particular forms of knowledge, how the two may support each other in knowledge acquisition, and so on. Computer-based environments for reading and writing offer an interesting venue in which to consider anew the question of reading–writing relations.

According to Thérien (1985), there are at least five distinct processes active every time we read a text:

- neuro-physiological (eye movements, the brain's functions, etc.);
- cognitive (the basic cognitive functions as studied by cognitive science);
- argumentative/narrative (the act of following a complex sign such as a discourse, a narrative, etc);
- affective (emotional response); and

- symbolic (interpretation of the text within the context of our own body of knowledge and establishment of relations between the text being read and other texts).

These five processes can be said to define three tasks: manipulation (the material dimension of the reading process), comprehension, and interpretation. In other words, to read a text is to be able to progress through it, which implies both manual and neuro-physiological aspects. In a post-typographic era, we must consider how each of these tasks changes and how the interrelations between them change as well. What does it mean to manipulate an e-book? What new strategies must be developed when the basic element of reading, the page, is not present? Can manipulation be transposed from one context (book culture) to another (screen culture or e-book culture) without any problems?

The second task, comprehension, implies the semiotic dimension of the reading process. To read a text is to understand what is written, which implies linguistic, cognitive, and affective aspects. E-books and hypermedia help produce new forms of texts, requiring new strategies of comprehension.

The third task, interpretation, refers to the symbolic dimension of the reading process. To read is to establish a relation between the text being read and other texts that explain, illustrate, complete, or expand what is being read. If interpretation is the minimal relation established between two texts by a reader, the second text facilitating understanding of the first, then a networked reading environment would presumably help bring about interpretation. And yet, this is not the case, as several early studies of readers working in hypertext environments demonstrate (e.g., Kim and Hirtle 1995; Foss 1989; Rouet and Levonen 1996). The problem here may be one of over-interpretation. When a reader makes connections that are not based on a complete or complex knowledge of the text being read, that connection may confound rather than facilitate interpretation. If the text is not "read"—if it has not been the object of an act of appropriation—then its interpretation may quite easily be "uote non fondé." That is, superficial, divergent instead of convergent (Eco 1992).

The three tasks involved in reading (interpretation, comprehension, and manipulation) are logically implied; interpretation logically implies comprehension, which logically implies manipulation. This recalls C. S. Peirce's notion of precision: to have a 3, you need a 2,

and to have a 2, you need a 1. 1 can stand by itself, but 2 needs a 1 in order to exist, and so on. This is to say that that we cannot have complex forms of interpretation if we do not have adequate forms of comprehension, which themselves require satisfactory forms of manipulation.

What is at stake now, with e-books, is the mastery of basic forms of manipulation. We still do not know how to read, manipulate, and work with e-books. To establish a reading culture of e-books, we must find ways to help establish robust forms of manipulation, which is the first step in facilitating strategies of comprehension and enabling interpretations to be put in motion.

6. Afterword: Consultation, Consensus, and the Needs of Interdisciplinary Research

Interdisciplinary and multidisciplinary approaches are key to this research—textual studies and bibliography, design and interface design, user and reader studies, and information management and computer science, at the very least—because of the complex and sophisticated nature of the questions being asked, which are themselves beyond the capability of any one individual or even one discipline to answer. This interdisciplinarity stands in contrast to traditional research contributions within scholarly fields where academics tend to make solo efforts with little direct collaboration with others—a model often reinforced through the conventional structure of doctoral studies and of hiring and promotion decisions (Hara et al. 2003; Newell and Swan 2000). At the same time, however, funding agencies are providing incentive for collaboration between researchers in universities and other research institutions as well as members of the community through their funding programs (SSHRC 2004; Newell and Swan 2000). As a discipline, digital humanists are capitalizing on these trends, undertaking collaborative research projects that require a variety of skills, both technical in nature and content/discipline specific. These collaborations draw upon individuals who are based at single institutions as well as with others who are located at institutions across the country and internationally. Interdisciplinary collaborative research projects of this kind offer clear benefits; however, the very nature of collaboration presents challenges which must be managed carefully if the success of such a project is to be ensured. A successful research team must develop mechanisms to

maximize those factors that contribute to success while minimizing the potentially negative impact of the associated challenges (Amabile et al. 2001; Cuneo 2003). Examination of the experience of the HCI-Book team—now, in 2010, the INKE team—provides insight into the nature of interdisciplinary research and suggests potential "best practices" for other teams to consider.

Consultation Process

This particular research team has been in development for approximately six years. The first meeting brought together several individuals who had previously worked together and had similar research interests, though from a variety of academic disciplines, to explore potential areas of collaboration. From this initial discussion, others with similar research interests were drawn in. In the second year, we met again to discuss potential research questions and potential team members. Over the next year, we continued online discussions and met for another weekend to further the conversation of proposed research directions. These efforts were supported by a week-long symposium held the following summer, which resulted in a more refined research question and set of methodologies, and more fixed team membership. Later that fall, the research question, methodologies, and team membership were finalized along with rules for interaction and collaboration, which were articulated in an application for the first stage of the granting program. In the fourth year of the collaboration, we were then invited to write a complete grant application. We were ultimately successful on our second attempt to the granting program, nearly six years after the first conversations had occurred. The final research team has 35 active researchers across four countries, several universities and institutions, and almost 90 academic disciplines and sub-disciplines, with a budget of approximately $13 million in monetary and in-kind contributions. The represented fields range from philosophy and cultural studies to visual communication design and robotics.

Challenges

As we developed our collaboration, we had to negotiate several key challenges stemming from the disciplinary differences within the team.

Disciplines play a very important organizing role within the academy. First, disciplines provide guidance to scholars within a

particular field on the nature of appropriate research questions, methodologies, research output, quality standards, funding levels, and career progression. These factors tend to be unique to each discipline (Bruhn 2000; Cech and Rubin 2004; Russell 1985). In many respects, this disciplinary insularity creates a common understanding of the "world" which allows members within the discipline to communicate easily and quickly in short-hand. However, as discussed above, some research questions require collaboration among a variety of disciplines, and this creates a potential for conflict and miscommunication. As the number of disciplines within a research project increases, the greater this potential becomes. Key issues that create conflict include, but are not limited to, the definition of research problems, the level of interdependence among team members, the determination of appropriate methodologies, the authorship attribution, the selection of appropriate publication and dissemination venues, and the development of effective work and communication patterns among the collaborators.

Ultimately, to collaborate effectively, researchers from different disciplines must find common ground in areas of theory, language, value system, methodology, and research style (Northcraft and Neale 1993; Bagshaw, Lepp and Zorn 2007; Bruhn 1995). In some cases, the research team may even need to develop a new working vocabulary specific to the interdisciplinary/multi-disciplinary project (Bracken and Oughton 2006; Lutz and Neis 2008). Even a common term such as "model" may be understood very differently in different disciplines (Derry, DuRussel and O'Donnell 1998). As a result, team members must exhibit flexibility, an eagerness to communicate, and a preparation for compromise (Bracken and Oughton 2006; Bruhn 1995). In early papers on their collaboration, Liu and Smith (2007) suggest that the analogy of ambassador may be appropriate in this context. An effective ambassador must be willing to bridge cultures and understand differences while finding the similarities upon which relationships can be built.

The attribution of academic credit and authorship is a particularly important issue facing multidisciplinary teams, particularly given the different conventions among disciplines (Kraut, Galegher and Egido 1987-1988; Choi and Pak 2007). For example, publications in the sciences tend to list all contributors to a project—not merely the authors—while in the humanities, credit is generally granted to the first author, who is the individual deemed to have done the most work (Kraut, Galegher and Egido 1987-1988; Fanderclai 2004). If

teams do not discuss and resolve this issue of attribution in advance, they often face conflict (Bagshaw, Lepp and Zorn 2007).

Our research team dealt with the challenges associated with disciplinary differences in the following ways before they became potentially damaging conflicts (Bagshaw, Lepp and Zorn 2007; McGinn et al. 2005). First, through the almost six years of discussion that took place before our successful grant application, we met face-to-face and through online mechanisms such as e-mail, project planning spaces, and Skype. Through these conversations, we found that while team members used the same words, they often defined the terminology in very different ways. The terms "book," "text," "reading," "authority," and "prototype" presented particular difficulties, requiring us to create a working vocabulary which allowed effective communication (Siemens 2009). Through this process, we were exposed to different perspectives and forced to work through our differences to find common ground and commitment to the collaboration as a whole (Siemens and INKE Research Group 2009). It is important to note that many of these face-to-face meetings combined formal discussions with informal ones over meals, which allowed us to get to know each other on a personal as well as professional level. The trust that is developed through these interactions can sustain a team through the usual stresses and strains of collaboration (Siemens 2008; Kraut and Galegher 1990). For any research project, but particularly for one of this magnitude in terms of budget, membership, and research scale, the development of the collaboration cannot be rushed if it can hope to be successful (Bagshaw, Lepp and Zorn 2007; Massey et al. 2001).

Second, given the variety of disciplines represented within our group and the resulting potential for conflict, our research team articulated an authorship strategy fairly early in the collaboration. Our negotiated convention of authorship comprises an acknowledgement of two or three key individuals as named authors along with the attribution "INKE Research Group," which clearly signals the nature of this particular working relationship. Any published work and data represents the collaboration of the whole team, past and present, not the work of any sole researcher.

Finally, beyond our discussions of research questions, objectives, and approaches during and after our grant application development, we also formally articulated roles, expectations, and task interdependence through a project charter (for the grant application stage) and administrative governance documents (for the grant working stage).

The charter was supplemented by a Gantt chart which outlined key tasks, timing, and research area responsibility. Further, the collaboration was portrayed as a figure eight diagram, which showed the integrative flow of research across our four research areas and the involvement of partners and stakeholders in the project. These documents reduced any potential for future conflict, and also provided an opportunity for us to "get to know each other and to build trust" (McGinn et al. 2005, 564). Other digital humanities and digital libraries teams have also found formal documentation to be necessary for effective team research. As one participant in an earlier study on collaboration in these communities stated, "formal documents sound cheesy, but in a multi-ethnic, multi-lingual, multi-generational, multi-talented work group (as every work group is) they are essential for setting a baseline understanding of what the project is and who is supposed to do what" (Siemens et al. 2009).

While our research team is still in the early phases of its 7-year research project, it has been a team in fact for almost 6 years. Through this time, we have been building the necessary relationships and processes that we will need to navigate the challenges associated with multidisciplinary research collaborations.

7. Bibliography

Amabile, Teresa M., et al., "Academic-Practitioner Collaboration in Management Research: A Case of Cross-Profession Collaboration," *Academy of Management Journal*, 44/2 (2001): 418-31.

Astleitner, H., and Leutner, D., "Learning Strategies for Unstructured Hypermedia—A Framework for Theory, Research, and Practice," *Journal of Educational Computing Research*, 13/4 (1995): 387-400.

Bagshaw, Dale, Margret Lepp, and Cecelia R. Zorn, "International Research Collaboration: Building Teams and Managing Conflicts," *Conflict Resolution Quarterly*, 24/4 (2007): 433-46.

Barrett, A., "The Information-Seeking Habits of Graduate Student Researchers in the Humanities." *Journal of Academic Librarianship*, 31/4 (2005): 324-331.

Barthes, R., *S/Z*, R. Millar, Trans. (New York: Hill and Wang, 1974).

Bates, M. J., "The Cascade of Interactions in the Digital Library Interface," *Information Processing & Management*, 38/3 (2002): 381-400.

Beagrie, N., and Jones, M., *Preservation Management of Digital Materials: A Handbook* (Digital Preservation Coalition, 2002).

Best, Michael, "Introduction: A Booth at the Fair," *Monitoring Electronic Shakespeares*, Spec. issue of *Early Modern Literary Studies*, 9/3 (2004): <http://extra.shu.ac.uk/emls/09-3/bestintr.htm>.

Best, Michael, ed., *The Internet Shakespeare: Opportunities in a New Medium*, Spec. issue of *Early Modern Literary Studies*, 3/3 (1998): <http://extra.shu.ac.uk/emls/si-02/si-02toc.html>.

Blandford, A., Rimmer, J. & Warwick, C., "Experiences of the Library in the Digital Age." *Proceedings of the Third International Conference on "Cultural Covergence and Digital Technology"* (Athens, 2006).

Bolter, Jay David, *Writing Space: Computers, Hypertext, and the Remediation of Print*, 2nd ed. (Mahwah, NJ: Lawrence Erlbaum, 2001).

Bracken, L. J., and E. A. Oughton, "'What Do You Mean?' the Importance of Language in Developing Interdisciplinary Research," *Transactions of the Institute of British Geographers*, 31/3 (2006): 371-82.

British Academy, *E-resources for Research in the Humanities and Social Sciences - A British Academy Policy Review*, (2005): <http://www.britac.ac.uk/reports/eresources/report/sect3.html#part5>.

Bruhn, John G., "Beyond Discipline: Creating a Culture for Interdisciplinary Research." *Integrative Physiological & Behavioral Science*, 30/4 (1995): 331-42.

Bruhn, John G., "Interdisciplinary Research: A Philosophy, Art Form, Artifact or Antidote?" *Integrative Physiological & Behavioral Science*, 35/1 (2000): 58-66.

Buzzetti, Dino, and Jerome McGann, "Electronic Textual Editing: Critical Editing in a Digital Horizon." *Electronic Textual Editing*. (Text Encoding Initiative Consortium, 2005): <http://www.tei-c.org/Activities/ETE/Preview/mcgann.xml>.

Caplan, P., "Preservation metadata." *Digital Curation Manual*: <http://www.dcc.ac.uk/resource/curation-manual/chapters/preservation-metadata/preservation-metadata.pdf>.

Carroll, John M., "Introduction: Human-Computer Interaction, the Past and the Present." In John Carroll (ed.), *Human-Computer Interaction in the New Millennium* (New York: ACM Press, 2002).

Cech, Thomas R., and Gerald M. Rubin, "Nurturing Interdisciplinary Research," *Nature Structural & Molecular Biology*, 11/12 (2004): 1166-69.

Chartier, Roger, *The Order of Books*, trans. Lydia G. Cochrane (Stanford: Polity Press, 1994).

Chatman, S., *Story and Discourse: Narrative Structure in Fiction and Film* (New York: Cornell University Press, 1978).

Cheypesh, Oksana, et al., "Centering the Mind and Calming the Heart: Mandalas as Interfaces," *Society for Digital Humanities (SDH/SEMI)* (Toronto: York University, 2006).

Choi, Bernard C. K., and Anita W. P. Pak, "Multidisciplinarity, Interdisciplinarity, and Transdisciplinarity in Health Research, Services, Education and Policy: 2. Promotors, Barriers, and Strategies of Enhancement," *Clinical & Investigative Medicine*, 30 (2007): E224-E32.

Chu, C. M., "Literary Critics at Work and their Information Needs: A Research-Phases Model," *Library & Information Science Research*, 21/2 (1999): 247-73.

Cuneo, Carl, "Interdisciplinary Teams - Let's Make Them Work," *University Affairs* (November 2003): 18-21.

Dalcher, D., and Genus, A., "Introduction: Avoiding IS/IT Implementation Failure," *Technology Analysis & Strategic Management*, 15/4 (2003): 403-407.

Dalton, M. S, and Charnigo, L., "Historians and their information sources," *College & Research Libraries*, 65/5 (2004): 400-425.

Day, M., "Metadata," *Digital Curation Manual*, (2005): <http://www.dcc.ac.uk/resource/curation-manual/chapters/metadata/metadata.pdf>.

DeRose, Steven J., David G. Durand, Elli Mylonas, and Allen H. Renear, "What Is Text, Really?" *Journal of Computing in Higher Education*, 1/2 (1990): 3-26.

Derrida, Jacques, *Of Grammatology* (Baltimore: Johns Hopkins University Press, 1976).

Derrida, Jacques, *Dissemination*, trans. Barbara Johnson (Chicago: University Press, 1981).

Derry, Sharon J., Lori Adams DuRussel, and Angela M. O'Donnell, "Individual and Distributed Cognitions in Interdisciplinary Teamworks," *Educational Psychology Review*, 10/1 (1998): 25.

Dillon, Andrew, *Designing Usable Electronic Text*, 2nd ed, (Boca Raton: CRC Press, 2004).

Dobson, T.M., "For the Love of a Good Narrative: Digitality and Textuality," *English Teaching: Practice and Critique*, 5/2 (2006): 56-68.

Dobson, T.M., and Luce-Kapler, R., "Stitching Texts: Gender and Geography in Frankenstein and Patchwork Girl," *Changing English*, 12/2 (2005): 265-277.

Dobson, T.M., and Miall, D.S., "Orienting the Reader?: A Study of Literary Hypertexts," *SPIEL*, 17/2 (1998): 249-62.

Douglas, J. Y., *The End of Books—Or Books Without End?: Reading Interactive Narratives* (Ann Arbor: University of Michigan Press, 2000).

Duff, W., "Metadata in Digital Preservation: Foundations, Functions and Issues," in F. M. Bischoff, H. Hofman and S. Ross (eds), *Metadata in Preservation: Selected Papers from an ERPANET Seminar at the Archives School Marburg* (Marburg: Archivschule Marburg, 2003).

Eco, Umberto, *Interpretation and Overinterpretation*, S. Collini (ed.) (Cambridge: Cambridge University Press, 1992).

Ellis, D., and Oldman H., "The English Literature Researcher in the Age of the Internet," *Journal of Information Science* 31/1 (2005): 29-36.

Fanderclai, Tari, "Collaborative Research, Collaborative Thinking: Lessons from the Linux Community," in James A. Inman, Cheryl Reed and Peter Sands (eds) *Electronic Collaboration in the Humanities: Issues and Options* (Mahwah, NJ: Lawrence Erlbaum Associates, 2004).

Fitzgerald, J. and Shanahan, T., "Reading and Writing Relations and their Development," *Educational Psychologist*, 35/1 (2000): 39-50.

Flowers, S., *Software Failure: Management Failure: Amazing Stories and Cautionary Tales* (Chichester, NY: John Wiley, 1996).

Foss, C., "Tools for Reading and Browsing Hypertext," *Information Processing and Management*, 25/4 (1989): 407-18.

Frascara, J., *User-Centred Graphic Design* (London: Taylor and Francis, 1997).

Galey, Alan, "Dizzying the Arithmetic of Memory: Shakespearean Source Documents as Text, Image, and Code," *Early Modern Literary Studies* 9/3 / Special Issue 12 (2004): 4.1-28, <http://purl.oclc.org/emls/09-3/galedizz.htm>.

Ganong, W.F., *Review of Medical Physiology*, 21st ed. (New York: Lange Medical Books, 2003).

Garrett, Jesse James, "Ajax: A New Approach to Web Applications," *Adaptive Path*, <http://www.adaptivepath.com/publications/essays/archives/>.

Gibson, J. J., *The Ecological Approach to Visual Perception*. (Boston: Houghton-Mifflin, 1979).

Granger, S., "Emulation as a Digital Preservation Strategy," *D-Lib Magazine*, 6/10 (2000).

Green, R., "Locating Sources in Humanities Scholarship: The Efficacy of Following Bibliographic References," *Library Quarterly*, 70/2 (2000): 201-229.

Greetham, D.C., *Textual Scholarship: An Introduction* (New York: Garland, 1994).

Greetham, D.C., *The Margins of the Text* (Ann Arbor: University of Michigan Press, 1997).

Hansen, W. J., "User Engineering Principles for Interactive Systems," *Proceedings of AFIPS Fall Joint Computer Conference* 39 (1971): 523-32.

Hara, Noriko, et al., "An Emerging View of Scientific Collaboration: Scientists' Perspectives on Collaboration and Factors That Impact Collaboration," *Journal of the American Society for Information Science and Technology*, 54/10 (2003): 952-65.

Hedstrom, M. L., et al., "'The Old Version Flickers More': Digital Preservation from the User's Perspective," *American Archivist*, 69 (2006): 159-187.

Ingwersen, P., and Jarvelin, K., *The Turn: Integration of Information Seeking and Retrieval in Context*, vol. 18 (Berlin: Springer, 2005).

Iser, W., *The Act of Reading: A Theory of Aesthetic Response* (Baltimore: Johns Hopkins University Press, 1978).

Johns, Adrian, *The Nature of the Book: Print and Knowledge in the Making* (Chicago: University of Chicago Press, 1998).

Kim, H., and Hirtle, S. C., "Spatial Metaphors and Disorientation in Hypertext Browsing," *Behaviour & Information Technology*, 14/4 (1995): 239-50.

Kintsch, W., "The Role of Knowledge in Discourse Comprehension: A Construction Integration Model." *Psychological Review*, 95/2 (1988): 163-82.

Kirschenbaum, Matthew G., "Extreme Inscription: Towards a Grammatology of the Hard Drive." *TEXT Technology*, 13/2 (2004): 91-125.

Kirschenbaum, Matthew G., "'So the Colors Cover the Wires': Interface, Aesthetics, and Usability," in Susan Schreibman, Ray Siemens, and John Unsworth (eds.), *A Companion to Digital Humanities* (Malden, MA: Blackwell, 2004): 523-42.

Kraut, Robert E., and Jolene Galegher, "Patterns of Contact and Communication in Scientific Research Collaboration," in Jolene Galegher, Robert Kraut, and Carmen Egido (eds.), *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work* (Hillsdale, NJ: Erlbaum, 1990).

Kraut, Robert E., Jolene Galegher, and Carmen Egido, "Relationships and Tasks in Scientific Research Collaboration," *Human-Computer Interaction*, 3/1 (1987-1988): 31-58.

Landow, George, *Hyper/Text/Theory*. (Baltimore and London: Johns Hopkins University Press, 1994).

Landow, George, *Hypertext 2.0: The Convergence of Contemporary Critical Theory and Technology*. Rev. ed. (Baltimore: Johns Hopkins University Press, 1997).

Lavoie, B. F., and R. Gartner, *Preservation Metadata* (No. 05-01). *Digital Preservation Coalition*, 2005.

Lazinger, S., *Digital Preservation and Metadata* (Englewood: Greenwood, 2001).

Lehmann, S., and P. Renfro, "Humanists and Electronic Information Services: Acceptance and Resistance," *College and Research Libraries*, 52 (1991): 403-413.

Leventhal, L. M., et al., "Sleuthing in HyperHolmes: An Evaluation of Using Hypertext vs. a Book to Answer Questions," *Behaviour and Information Technology*, 12 (1993): 149-64.

Linden, J., et al., *The Large Scale Archival Storage of Digital Objects*. (Digital Preservation Coalition, 2005).

Liu, Yin, and Jeff Smith, "Aligning the Agendas of Humanities and Computer Science Research: A Risk/Reward Analysis," *SDH-SEMI 2007*.

Luce-Kapler, R., and T. M. Dobson, "In Search of a Story: Reading and Writing e-Literature," *Reading Online*, 8/6 (2005): np.

Lutz, John Sutton, and Barbara Neis, "Introduction," in John Sutton Lutz and Barbara Neis (eds.), *Making and Moving Knowledge: Interdisciplinary and Community-Based Research in a World on the Edge* (Montreal, Quebec: McGill-Queen's University Press, 2008).

Maclean, Marie, "Pretexts and Paratexts: The Art of the Peripheral," *New Literary History* 22/2 (1991): 273-79.

Makri, S., et al., "A Library or Just another Information Resource? A Case Study of Users' Mental Models of Traditional and Digital Libraries," *Journal of the American Society of Information Science and Technology*, 58/3 (2007): 433-45.

Massey, Anne P., et al., "When Culture and Style Aren't About Clothes: Perceptions of Task-Technology 'Fit' In Global Virtual Teams," *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work.*

McCarty, Willard. "Knowing True Things by What Their Mockeries Be: Modelling in the Humanities," in Ray G. Siemens and Susan Schreibman (eds.), *A Companion to Digital Literary Studie*s (Oxford: Blackwell, 2008): 391-401.

McGinn, Michelle K., et al., "Living Ethics: A Narrative of Collaboration and Belonging in a Research Team," *Reflective Practice* 6.4 (2005): 551-67.

McKenzie, D.F., "The Broken Phial: Non-Book Texts," *Bibliography and the Sociology of Texts* (Cambridge: Cambridge University Press, 1999): 31-53.

Miall, D. S. and T. M. Dobson, "Reading Hypertext and the Experience of Literature," *Journal of Digital Information*, 2/1: (2001). <http://journals.tdl.org/jodi/issue/view/9>

Miall, D. S., and D. Kuiken, "Beyond Text Theory: Understanding Literary Response," *Discourse Processes*, 17 (1994): 337-52.

Newell, Sue, and Jacky Swan, "Trust and Inter-Organizational Networking," *Human Relations*, 53/10 (2000): 1287-328.

Northcraft, Gregory B., and Margaret A. Neale, "Negotiating Successful Research Collaboration," in J. Keith Murnighan (ed.), *Social Psychology in Organizations: Advances in Theory and Research* (Englewood Cliffs, NJ: Prentice Hall, 1993).

Rabinowitz, P. J., *Before Reading: Narrative Conventions and the Politics of Interpretation* (Ithaca, NY: Cornell University Press, 1987).

Renear, Allen, "Out of Praxis: Three (Meta)Theories of Textuality," in Kathryn Sutherland (ed.), *Electronic Text: Investigations in Method and Theory* (Oxford: Clarendon, 1987): 107-26.

Rose, Jonathan, "Worker's Journals," in J. Don Vann and Rosemary T. VanArsdel (eds.), *Victorian Periodicals and Victorian Society* (Toronto: University of Toronto Press, 1995): 301-310.

Rosenberg, M.E., "Physics and Hypertext: Liberation and Complicity in Art and Pedagogy," in George Landow (ed.), *Hyper/Text/Theory* (Baltimore: John Hopkins University Press, 1994): 268-98.

Rosenblatt, L., *Literature as Exploration.* (New York: Noble and Noble, 1938/1968).

Ross, S., "How to Preserve and Prove the Integrity of Digital Objects?" *DigiCULT Thematic Issue: Integrity and Authenticity of Digital Cultural Heritage Objects* (2002).

Rouet, J. F., and J. J. Levonen, "Studying and Learning with Hypertext: Empirical Studies and their Implications," in J.F. Rouet et al. (eds.), *Hypertext and Cognition* (Mahwah, NJ: Lawrence Erlbaum, 1996): 9-23.

Ruecker, Stan, *Affordances of Prospect for Academic Users of Interpretively-tagged Text Collections*. Unpublished PhD Dissertation (Edmonton: University of Alberta, 2003).

Ruecker, Stan, "The Electronic Book Table of Contents as a Research Tool," *Congress of the Humanities and Social Sciences: Consortium for Computers in the Humanities/Consortium pour Ordinateurs en Sciences Humaines (COCH/COSH) Annual Conference* (London, Ontario, 2005).

Ruecker, Stan, "Proposing an Affordance Strength Model to Study New Interface Tools," *Digital Humanities 2006* (Paris, 2006).

Ruecker, Stan, et al., "The Clear Browser: Visually Positioning an Interface for Data Mining by Humanities Scholars," *Digital Humanities 2006* (Paris, 2006).

Russell, Martha Garrett, "Administering Interdisciplinary Collaboration," in B.W. Mar, William T. Newell, and Borje O. Saxberg (eds.), *Managing High Technology: An Interdisciplinary Perspective* (Amsterdam: Elsevier, 1985).

Schreibman, Susan, "Computer-Mediated Texts and Textuality: Theory and Practice," *Computers and the Humanities*, 36 (2002): 283-93.

Shneiderman, B., and Plaisant, C., *Designing the User Interface: Strategies for Effective Human Computer Interaction.*, 4th ed. (Reading, Mass.: Pearson—Addison Wesley, 2005).

Siegfried, S., M. J. Bates, and D. N. Wilde, "A Profile of End-User Searching Behavior by Humanities Scholars: The Getty Online Searching Project Report no. 2," *Journal of the American Society for Information Science*, 44/5 (1993): 273-91.

Siemens, Lynne, and the INKE Research Group, "From Writing the Grant to Working the Grant: An Exploration of Processes and Procedures in Transition." *INKE 2009*.

Siemens, Lynne, et al., "'It Challenges Members to Think of Their Work through Another Kind of Specialist's Eyes': Exploration of the Benefits and Challenges of Diversity in Digital Project Teams," *Thriving on Diversity - Information Opportunities in a Pluralistic World*, (ASIS&T 2009 Annual Meeting).

Siemens, Lynne, "The Balance between on-Line and in-Person Interactions: Methods for the Development of Digital Humanities Collaboration," *SDH-SEMI 2008*.

Siemens, R.G., "Cultures Separated by a Common Language? Homography/Homophony and Collaborative Understanding, in Discipline and Context," *Enhancing International Research Collaborations, Strengthening Quality, Connections and Impacts: A SSHRC Congress 2009 Session.*

Siemens, R.G., Claire Warwick, Richard Cunningham, Teresa Dobson, Alan Galey, Stan Ruecker, Susan Schreibman, and the INKE Research Group. "Codex Ultor: Vers une base conceptuelle et theorique pour la nouvelle recherche sur les livres et les environnements documentaries," *Memoires du livre/Studies in Book Culture* 1.1 ([2010 for] 2009). <http://www.erudit.org/revue/ memoires/ 2009/v1/n1/038636ar.html>.

Siemens, R.G., Claire Warwick, Richard Cunningham, Teresa Dobson, Alan Galey, Stan Ruecker, Susan Schreibman, and the INKE Research Group. "Codex Ultor: Toward a Conceptual and Theoretical Foundation for New Research on Books and Knowledge Environments." *Digital Studies/Le Champ Numerique* 1.2 ([2010 for] 2009). <http://www.digitalstudies.org/ojs/index. php/ digital_studies/article/view/177/220>.

Siemens, R.G., Richard Cunningham, Alan Galey, Stan Ruecker, Lynne Siemens, Claire Warwick, and the INKE Research Group, "Implementing New Knowledge Environments: Laying Research Foundations for Understanding Books and Reading in the Digital Age," Proceedings of *BooksOnline 2009* at the 13th European Conference on Digital Libraries, Corfu, October 2009. <http://research.microsoft.com/en-us/um/cambridge/events/ booksonline09/papers/p6.pdf>.

Siemens, R.G. et al., "The Credibility of Electronic Publishing: A Report to the Humanities and Social Sciences Federation of Canada," *TEXT Technology*, 11/1 (2002): 1-128.

Sinclair, Stéfan and Geoffrey Rockwell. "Reading Tools, or Text Analysis Tools as Objects of Interpretation," *Digital Humanities 2007*. 199-200. Web <http://www.digitalhumanities.org/ dh2007/ abstracts/xhtml.xq?id=226>.

Sless, D., and R. Wiseman, *Writing About Medicines for People*, 2nd ed. (Melbourne: Communication Research Institute of Australia, 1997).

Slights, William W. E., *Managing Readers: Printed Marginalia in English Renaissance Books* (Ann Arbor: University of Michigan Press, 2001).

Spufford, Margaret, *Small Books and Pleasant Histories: Popular Fiction and its Readership in Seventeenth Century England* (Athens, GA: University of Georgia Press, 1982).

Social Sciences and Humanities Research Council of Canada (SSHRC), *From Granting Council to Knowledge Council: Renewing the Social Sciences and Humanities in Canada* (Ottawa, Ontario: Social Sciences and Humanities Research Council of Canada, 2004).

Steemson, M., "DigiCult Experts Search for e-Archive Permanence," *DigiCULT Thematic Issue: Integrity and Authenticity of Digital Cultural Heritage Objects*, (2002).

Stone, S., "Humanities Scholars-Information Needs and Uses," *Journal of Documentation* 38/4 (1982): 292-313.

Stotsky, S., "Research on Reading/Writing Relationships: A Synthesis and Suggested Directions," *Language Arts*, 60 (1983): 627-43.

Talja, S. and J. Maula, "Reasons for the Use and Non-Use of Electronic Journals and Databases—A Domain Analytic Study in Four Scholarly Disciplines," *Journal of Documentation*, 59/6 (2003): 273-91.

Thérien, G., *Sémiologies* (Montréal: Université du Québec à Montréal, 1985).

Van Peer, W., and Chatman, S. (eds.), *New Perspectives on Narrative Perspective* (New York: SUNY, 2001).

Warwick, C., A. Blandford, and G. Buchanan, "User Centred Interactive Search: A Study of Humanities Users in a Digital Library Environment," *Association for Computers and the Humanities- Association for Literary and Linguistic Computing 2005* (Victoria, BC: University of Victoria, 2005).

Warwick, C., et a.l, "'If You Build It Will They Come?' The LAIRAH Study: Quantifying the Use of Online Resources in the Arts and Humanities Through Statistical Analysis of UserLog Data," *Proceedings of Digital Humanities 2006* (Paris: Universite de Paris-Sorbonne, 2006): 225-28.

Warwick, C., et al., *The LAIRAH Survey, A Report to the Arts and Humanities Research Council: AHRC ICT Strategy Scheme*, 2007.

Warwick, Claire, "Print Scholarship and Digital Resources," in Susan Schreibman, Ray Siemens, and John Unsworth (eds.), *A Companion to Digital Humanities* (London: Blackwell, 2004): 366-82.

Waters, D., and J. Garret, *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information. Commission on Preservation and Access and the Research Libraries Group*, 1996.

Watson-Boone, R., "The Information Needs and Habits of Humanities Scholars," *Reference Quarterly*, 34 (1994): 203-16.

Watt, Tessa, *Cheap Print and Popular Piety, 1550-1640* (Cambridge: Cambridge University Press, 1991).

Whitmire, E., "Disciplinary Differences and Undergraduates' Information-Seeking Behaviour," *Journal of the American Society for Information Science and Technology*, 53/8 (2002): 631-38.

Wiberley, S., "Time and Technology," *College and Research Libraries*, 61/5 (2000): 421-31.

Wiberley, S. E., "Subject Access in the Humanities and the Precision of the Humanist's Vocabulary," *Library Quarterly* 53 (1983): 420-433.

Wilberley, S. E. "Names in Space and Time: The Indexing vocabulary of the Humanities," *Library Quarterly* 58 (1988): 1-28.

Zwaan, R. A., J. Magliano, and A. C. Gaesser, "Dimensions of Situation Model Construction in Narrative Comprehension," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21/1 (1995): 386-97.

SOMMAIRE

Cet article fait la synthèse d'un exercice de consultation auquel a procédé entre 2006 et 2008 une équipe de chercheurs de HCI-Book (Human-Computer Interaction/Interface with the Book *Interaction homme-machine/Interface avec le livre*), mise sur pied grâce à un octroi monétaire accordé par le Conseil de recherches en sciences humaines (CRSH) en vue de souligner le rôle et l'importance des livres numériques et analogiques dans l'attribution des subventions de recherche dans les Humanités. Ce groupe a cherché à connaître (1) ce que nous savons réellement au sujet de ces nouveaux textes qui remplacent l'imprimé et qui représentent à nos yeux les connaissances et l'expérience du passé de même que ceux qui nous sont livrés sous forme de documents numériques directs; (2) comment nous interagissons avec ces objets du savoir et l'information qu'ils contiennent; et (3) de quelle manière la rencontre du caractère médiatique dans ces supports numériques influence l'usage que nous en faisons permettant ainsi d'avoir une meilleure interaction avec ces nouveaux supports numériques. Par le biais de cette consultation, le groupe de travail a commencé à identifier les principaux points touchant la numérisation du document à contenu humaniste et d'en

prendre acte dans le but (1) de comprendre et décrire les principes de base de l'interaction humaniste avec les objets du savoir (tant numériques qu'analogiques), (2) d'élaborer des stratégies de base afin de créer un modèle d'objet du savoir humaniste, notamment les livres électroniques et ce, en fonction de cette vision commune et (3) d'énoncer des principes de base pertinents en vue d'évaluer et d'implanter les technologies en cours non sans explorer aussi celles à venir. Cette recherche a été rendue possible en 2009 grâce à un montant octroyé par le programme des Grands travaux de recherche concertée du CRSH dans le cadre d'un projet d'implantation des nouvelles connaissances qui a avalisé vigoureusement ces échanges de vue constructives suivant le plan établi.