Contents

Contributors viii
Introduction 1
  Willard McCarty 1
1. Never Say Always Again: Reflections on the Numbers Game  
  John Burrows 13
2. Cybertextuality by the Numbers  
  Ian Lancashire 37
3. Textual Pathology  
  Peter Garrard 71
4. The Human Presence in Digital Artefacts  
  Alan Galey 93
   Perspective  
   Edward Vanhoutte 119
6. Electronic Editions for Everyone  
   Peter Robinson 145
7. How Literary Works Exist: Implied, Represented, and Interpreted  
   Peter Shillingsburg 165
8. Text as Algorithm and as Process  
   Paul Eggert 183
9. ‘I Read the News Today, Oh Boy!’: Newspaper Publishing 
   in the Online World  
   Marilyn Deegan and Kathryn Sutherland 203
References 219
4. The Human Presence in Digital Artefacts

Alan Galey

1. What Lies Beneath

The reader wanders at leisure over smiling fields; he plays and runs and never stumbles; and he never gives a thought to the time and tedium it has cost me to battle with the thorns and briers, while I was clearing the land for his benefit. He does not reckon [...] how great the discomforts that secured his comfort, how much tedium was the price of his finding nothing tedious.

Erasmus, letter to William Warham, Archbishop of Canterbury (1976 [1514-6]: 262)

This essay considers the tensions between the surface orderliness of scholarly resources and the stubborn irregularity of textual materials. Textual scholarship stands to contribute two key ideas to the digital humanities: first, that there is more to electronic forms than what reaches the screen; and second, that the relationship of form to content is complex and sometimes beyond exhaustive modelling. These two points may seem commonsensical enough within a book-history context, but much of the hypertext theory that dominated the previous decade gives little impression they could matter. Part of the burden of digital textual studies must be to counter the influence of those hypertext theorists who rushed to essentialize computing, the Internet, and digital textuality. More recent work by Katharine Hayles, Matthew Kirschen-

1 The work presented here was supported by the Social Sciences and Humanities Research Council of Canada. Different parts of this article were presented at the conferences of the Society for Textual Scholarship and the Society for Digital Humanities / Société pour l’étude des médias interactifs, the University of Toronto’s Faculty of Information Studies, and Texas A&M University. I am grateful to those audiences for their questions and comments and especially to Willard McCarty, Richard Cunningham, Christopher Moore, and Stan Ruecker for their comments on early drafts. Any remaining errors are mine.
baum, Lev Manovich, Alan Liu, and others associated with media-specific analysis and new software studies has shown what technically informed perspectives can bring to the study of digital textuality. Kirschenbaum describes their approach as one that ‘cultivates granular, material readings of the inevitable cultural and ideological biases encoded by particular applications and interfaces’ (2004a: 533). Such modes of reading expose continuities with past textual practices that some versions of hypertext theory were predisposed not to see — yet such ‘granular, material readings’ render new textual practices and technologies no less exciting for all that. This essay argues that it should be disquieting to see a deepening separation of material form from idealized content in our tools at the very moment when literary critics have established the materiality of texts to be indispensable to interpretation. As digital textual studies takes shape as a field, it finds itself caught between these divergent trends in computational practice and literary theory.

Textual scholarship has long been driven by an anxious desire to know what lies beneath the perceptual surface — the authorial consciousness embedded in written language; the original of multiple versions; the moments of live inscription held within inert physical artefacts. If the object of textual work is to delve beneath surfaces, the subjects who carry out that work find themselves at the threshold, negotiating between the dead past and living present. In Gary Taylor’s oft-quoted formulation, ‘Editing is a ritual we perform over the corpus of an author who has passed away’ (1988: 50), and all such rituals exist for the sake of the living. Digital textual scholars might well sympathize with Erasmus (quoted in the epigraph above) as he laments the conditions of his editorial labours on the Vulgate of St Jerome. Erasmus’s words capture the double-vision that texts demand of us with regard to their mediation of surfaces and depths. It is the same with our tools, digital and otherwise. Beneath any smoothly functioning computer interface such as a Web browser, the source code may harbour the ‘thorns and briars’ (in Erasmus’s words) of half-solved bugs, lingering after the ‘battle’ with materials and deadlines to build the digital artefact. In a similar manner, the monuments of textual, philological, and historical scholarship are often associated with what Erasmus calls ‘discomfort’ and ‘tedium’ in their production, as is the fictional Edward Casaubon’s miserable Key to All Mythologies in Middlemarch, or sometimes even spring from terrible trauma and loss, as in the case of Alfred Pollard’s ‘punishing work schedule’ (Maguire 1996: 28) while he grieved the deaths of his sons in

---

2 On media-specific analysis, see Hayles 2004.
4. The Human Presence in Digital Artefacts

World War I. Battlefield or cleared land: both metaphors for the edited text are present in Erasmus’s picture of textual labour, positing a substratum of textual remains beneath the reader’s feet. What then is ‘the price’, as Erasmus puts it, to be paid for the illusion of digital texts and environments as ‘smiling fields’ where the user never stumbles? The following discussion approaches that question by examining some theoretical assumptions that shape the digital tools of Erasmus’s inheritors.

As Erasmus implies, we cannot detach technical concerns from the aesthetic, symbolic, and hermeneutic dimensions of textual work, and his preoccupation with unseen complexity takes visual form in Figure 1. This image shows what lies beneath the main reading room in the New York Public Library’s Central Building at 5th Avenue and 42nd Street, an icon of cultural heritage. The illustration comes from the cover of an issue of *Scientific American* published five days after the new building opened to the public in 1911. The genre should be familiar — a representation of the inner workings of an interface between readers and a massive collection of texts, using a visualization technique (the cutaway view) to make the unseen mechanism intelligible to non-specialists. In the reading room at the top, patrons and librarians use a system of catalogues, request slips, and pneumatic tubes to order books housed in the stacks below. The books, once located, ascend to the reading room via mini-elevator. We can also read a foreshadowing of the digital humanities’ open-access ethos here, too, since this union of aesthetics and machinery serves a public library.

But *Scientific American*’s tribute to fin de siècle design and engineering also illuminates anxieties about what lies beneath it. The image embodies the kind of metaphor for archiving that Thomas Richards uses to describe the symbolic importance of the British Museum at the data-collecting height of the British Empire, and especially its basement as a chaotic space that was symptomatic of the material pressures of data overload (Richards 1993: 4-16). The stacks beneath the New York Public Library reading room emphasize a volume of information that, as a totality, goes unseen by its users. As in Richards’ metaphor, there is even a basement at the very bottom of the New York Public Library’s stacks, where boxes appear in disordered contrast to the stacks and retrieval system above. Such are the pressures felt by structures that must be at once monument and infrastructure.

---

The image also displays a fixation with human presence in the form of the tiny figures that populate the stacks (forty-seven of them). Where Erasmus claimed to feel alone in his task of textual management, no such isolation seems possible in this system. One could imagine this image without the figures, as a strictly technical blueprint, but the purposeful distribution of humans throughout makes this a representation not just of mechanical automation but also of human labour. The book is a text on a human scale,
and this image goes out of its way at least forty-seven times to reassert a human presence in a system that holds books, readers, and machines within its compass. Yet none of the humans or books are individuated anywhere in the image; we are shown neither individual readers nor recognizable books that matter in their specificity. As the accompanying article states approvingly, the system ‘distribute[s] the reader rather than the volumes which he reads’, and ‘automatically […] divides the thousands of readers who wish to consult the books into the intellectual classes in which they belong’ (Anon. 1911: 527). Even with the focus on the scale of ‘thousands of readers’ and their intellectual subclasses, this representation of the library seems to forget roughly half of the library’s potential users: only male readers are represented. The living moments of encounter between individual readers and texts, in all their diversity and idiosyncrasy, remain deferred in this representation, implicit in the image but inscrutable to our eyes. The logic of the image, then, is as much temporal as spatial: celebrate resource-building now; understand the particularities of material later.

This large-scale mode of representing reading stands in contrast to humanist depictions of individual reading and writing in Erasmus’s time. That instinctive humanist desire to draw closer, like Tantalus, to some idealized but elusive textual encounter finds powerful expression in Vittore Carpaccio’s 1502-3 painting The Vision of Saint Augustine (Figure 2, below).

Carpaccio’s painting depicts Augustine penning a letter to Jerome at the precise moment of the latter’s death, as the ghostly presence of the letter’s addressee fills the room. In Carpaccio’s probable source, a 1485 Venetian edition of the life of Jerome, Augustine in his cell at Hippo is attempting to quantify the joy of souls in the presence of God, and is just putting pen to paper to ask Jerome in Bethlehem for his thoughts. The information Augustine seeks comes to him in a moment of miraculously instant communication that accords more with our present than with the epistolary of the ancient world. Jerome, the archetype of textual scholars, rebukes Augustine’s reduction of knowledge to human numbers: ‘Augustine, Augustine, what are you seeking? Do you think that you can put the whole sea in a little vase? […] Will your eye see what the eye of no man can see? […] By what measure will you measure the immense?’ (quoted in Roberts 1959: 292).}

4 On the painting’s probable source and the misidentification of the painting as a depiction of Jerome himself, see Roberts 1959.
5 The translation is Roberts’s from Hieronymus. Vita et transitus, Venice 1485 (Huntington Library, transcribed by Eugene Brunelle): ‘Augustine, Augustine, quid queris: putasse brevi inimittere vasculo mare totum […] Que oculus nullus hom-
Human text gives way to divine voice, and the letter to Jerome becomes the perfect interface, instantly receding before an unmediated presence. Where the *Scientific American* image shows a perfectly synchronic system, outside of time and history, Carpaccio shows us the instant of collaborative intellectual work deeply embedded in history. In Alexander Nagel and Christopher S. Wood’s reading, ‘The fluttering pages of the open codices, the fall of shadows, the alerted dog, the poised pen all suggest the momentariness of that moment, the evening hour of compline, as Augustine tells us. This is secular time, the time of lived experience, whose each moment repeats but differs from the previous moment’ (2005: 403). Most importantly, *The Vision of Saint Augustine* is not our vision, and the miraculous text Augustine receives from Jerome lies beyond the limits of human representation.

Both of these scenes depict technologies for managing multiple texts — Carpaccio places numerous writing implements, books, and a horizontal reading wheel in Augustine’s study — but his painting meditates on the partialness of human knowledge, while the *Scientific American* image celebrates the abstraction of a mechanical system. Both use encounters with inum videre potuit tuus videbit? [...] Immensa, qua mensura metieris?” (Roberts 1959: 297).

6 On Carpaccio’s depiction of temporality and its relation to the history and nature of reading, see Bringhurst 2006.
documents to reflect on orders of experience that exceed human capacities. Reflecting on a copy of the Carpaccio painting given to him as a student, K. Anthony Appiah comments that ‘the shelf of books behind the saint — his library — contained most of the works he would have thought worth reading’; ‘he would almost certainly have read all of them’ (2005: 45). Today, Appiah notes, more is printed in a single city in a week than Augustine or Jerome could have read in their lifetimes: ‘we are, in short, drowning in the particulars we humanists study’. In essence, the *Scientific American* image is 1911’s answer to Appiah’s fears, and to the question posed in the title of a colloquium and article by Gregory Crane: ‘What Do You Do with a Million Books?’ (2006). One could read the Carpaccio and *Scientific American* images’ differences as emblematic of the digital humanities in its present state, which emphasizes abstract, large-scale approaches such as linguistic corpora and data mining, the social-science version of literary history practiced by Franco Moretti (so-called distant reading), and text analysis techniques that derive patterns from multitudinous low-level observations rather than situated acts of subjective interpretation. These approaches represent a movement away from the humanities’ traditionally idiographic tendency (to seek local knowledge about specific cases) and toward the natural and social sciences’ nomothetic tendency (to seek abstract patterns and general laws). These approaches also share something in common with the *Scientific American* image in that they place the viewer or critic in a superhuman position, showing systems of words and texts from a perspective that no single human could occupy in real space and time. Moretti stresses that distance can be ‘a condition of knowledge’ (2000: 57; emphasis added).

---

7 Appiah actually refers to ‘one of Carpaccio’s great murals of Saint Jerome’ but it is more likely that he has the Augustine image in mind. Carpaccio’s other paintings of Jerome do not feature libraries, and this image was often mistaken as a depiction of Jerome (see note 4 above).

8 The colloquium was co-sponsored by the University of Chicago and the Illinois Institute of Technology; see Crane 2006 and the Million Books Project, a collaboration between Carnegie Mellon University Libraries and the Internet Archive: http://www.archive.org/details/millionbooks [accessed 19/10/2008].


10 The terms *idiographic* and *nomothetic* originate with the neo-Kantian philosopher Wilhelm Windelband (1998/1894: 13), and today see more frequent use in anthropology and psychology than in literary studies. On the tradition of thought about the distinction they name, including Max Weber, see Manicas 1998.
removed), but The Vision of Saint Augustine complicates the metaphors we use to represent distance, proximity, and the quantifiable. The viewer of that painting remains all too human. Carpaccio invites his viewers just to the threshold of human experience, inviting us to cast our eyes, like Augustine in the painting, beyond the frame of human perception even as we accept its limits. The scale of the human is the preoccupation of both images.

Such is the power of tools and representations alike: to shape thinking, both through the conclusions they enable and the metaphors they deploy. The concerns this essay advances have tended to remain tacit in the digital humanities, a field whose sustaining progress narratives and investments in fundable projects foster a sense of itself as an onward march into the future — an avant garde that was the first to embrace computing as a tool for humanities scholarship. Yet the tool-building enterprise risks falling into a binary in which digital tools represent innovation, dynamism, and provocative instability, while the materials they operate upon — very often literary texts — represent availability, continuity, and unproblematic stability. This binary makes it easy to forget textual work always has an interpretive dimension that depends upon the complexity of humanities materials, especially after bibliographically aware literary scholarship in the wake of D.F. McKenzie and Jerome McGann has established the value of joining interpretation with the materiality of texts. Our understanding of that relationship has become intertwined with another, less obvious one: the tension between tools and materials in the digital humanities.

2. Digital Textual Scholarship

Every tool is a weapon,
If you hold it right.
Ani DiFranco, ‘My IQ’ (2002)

Whenever we ask what new technology can do for textual scholars, we must not lose sight of a deeper question: what is at stake in the work textual scholarship does, digitally and otherwise? What makes this work worth doing? Progress narratives almost always leave something important behind, and information culture itself has been accused of systematically forgetting its own history (Day 2001: 3), and of succumbing to a ‘rhetoric of newness’ and ‘rhetoric of amnesia’ (Rabinovitz and Geil 2004: 2). Indeed, we have been here before. The digital humanities now occupy
much the same position that W.W. Greg and the other New Bibliographers did in the era when the New York Public Library’s Central Building first opened its doors. At that time, the cultural pressures that went with social and technological change required the assimilation of vast amounts of technical knowledge about the transmission of texts — and by extension, the seemingly transmissible parts of culture itself — into a coherent progress narrative. This narrative had to account not only for the literary documents that had survived, but also for the practical means by which culture could be preserved and disseminated into the future through editing (and through the related activities of historical bibliography and bibliographic appraisal, enumeration, and preservation). Like digital humanists today, the New Bibliographers lived in a time of new media and information technology; they had to articulate their work to a changing academy that often did not understand it; they were obliged by their material to command a detailed knowledge of how texts, humans, and machines interact; and they had to respond to the often-contradictory imperatives of explaining and making.

Today, textual studies stands not only as a beneficiary of new tools to solve old problems — and, let us hope, to find new problems — but also as a well-developed perspective on new kinds of cultural artefacts. Throughout this essay, the term artefact encompasses products of human artifice that can be studied for interpretative purposes, like books, but also what McKenzie somewhat awkwardly called ‘non-book texts’.11 (Digital artefact can also mean an instance of visual noise in a digital image, but that usage does not enter this discussion.) The terminological challenge is to find a noun that includes books as easily as cultural productions like video games, films, and paintings, but avoids the scientific and programming connotations of the term object. Anthropologists and archaeologists have also thought about this problem. For example, Anders Andrén makes a distinction between artefacts and texts, though not a rigid one (1998: 146-53), and Karin Barber takes text to include artefacts and verbal performances (2007: 1-29). Book historian Matthew Brown helps to focus the concept by describing artefact as ‘a term which suggests an authentic, extant source, not a copied, transcribed, and edited version’ (2004: 702-3). However, Brown’s description becomes complicated when we consider whether a copy of a video game can be an artefact, since there are no instances of, say, the

game Myst in the world that are not copies — or can such cultural works never attain artefactual status? What then of a copy of the Shakespeare First Folio, itself an edited text created in part from scribal transcriptions, but which advertises itself as ‘Published according to the True Original Copies’ (Blayney 1996: 3)? A thorough definition of artefact is beyond the scope of this essay; instead, it may be more useful to consider the intellectual contexts in which we define a term like digital artefact.

On its present course, digital textual scholarship may well turn out to be a continuation of the project of D.F. McKenzie. By the time of his death in 1999, his work and influence had gone a long way toward disentangling the field from the orthodoxies of the New Bibliography, and had reintroduced historical and interpretive perspectives into editorial theory, which predecessors such as W.W. Greg had tended to regard as a closed system of transmissible texts, human agents, and mechanical constraints. This essay takes its title from a phrase of McKenzie’s, for whom bibliography’s great virtue was that it could ‘show the human presence in any recorded text’ (1999/1985: 29). These are words to conjure with: the phrase ‘any recorded text’ opens the scope of textual scholarship’s materials to all manner of what McKenzie called ‘non-book texts’, including ‘films, recorded sound, static images, computer-generated files, and even oral texts’ (1999/1985: 4), to which we could add software, born-digital fiction and poetry, and now blogs, wikis, and social networking websites — the kinds of intensely socialized digital texts whose existence in a Web 2.0 world would likely have fascinated McKenzie had he lived to see it. It is worth noting that he tended not to describe computers simply as new tools for the textual scholar’s toolbox, but rather as a welcome challenge in a continuing professional obligation to account for new forms of communication. As McKenzie suggested in his centenary lecture for the Bibliographical Society in 1992,

---

12 These currents run throughout most of McKenzie’s work, but see in particular his chapter ‘The Book as an Expressive Form’ (1999/1985: 9-30). For an overview of responses to McKenzie’s position see van der Weel 2005; the most pointed criticisms may be found in Tanselle 1991.

13 See also McKenzie 1999/1985: 13 and 39. This broad scope has proven easier to embrace in theory than in critical practice. For example, David Greetham’s review of Burnard, O’Brien O’Keeffe, and Unsworth (2006) highlights several problems with the collection’s overall conception, particularly that ‘the absence of painting, dance, film, television, video games, music (about all of which there has been some very challenging discussion of late) makes the collection almost relentlessly text- (or linguistics-) based’ (2007: 135).
4. The Human Presence in Digital Artefacts

That obligation has acquired a new urgency with the arrival of computer-generated texts. The demands made of bibliography and textual criticism by the evolution of texts in such forms, the speed with which versions are displaced one by another, and the question of their authority, are no less compelling than those we accept for printed books. By the logic of our discipline, we’re equally committed to acknowledge that these textual artefacts also embody the conditions of their construction. (McKenzie 2002/1992: 272-3)

This is a remarkable statement for being both progressive and conservative at once. In the progressive sense, McKenzie naturalizes the expansion of textual scholarship’s circle of knowledge to encompass the digital, such that the modifier digital becomes redundant in digital textual studies. By his logic, to reject inquiry into digital artefacts is to reject the very essence of textual scholarship. But this vision of textual studies also conservatively extends the traditional concerns of print and manuscript bibliography to digital artefacts, with McKenzie’s first thoughts tending toward the enumeration of versions and the establishment of authority among them.

Does digital textual scholarship then consist of applying existing descriptive and analytic methods to digital artefacts? To an extent, this conservative approach works, and the single most edifying example so far may be Kirschenbaum’s article ‘Editing the Interface: Textual Studies and First Generation Electronic Objects’.14 Taking as his subject canonical electronic literature such as Michael Joyce’s afternoon Kirschenbaum deftly applies a McGannian awareness of bibliographic codes in reading the material nuances of born-digital objects. This mode of reading raises a question the field is still working to answer: ‘what if a textual scholar, well-versed in theories of textual editing, were […] to be given the task of preserving the original text of afternoon in some stable and standardized electronic format for the sake of the scholarly record? How would our scholar go about it?’ (Kirschenbaum 2002: 33). This is the kind of question that should keep textual scholars awake at night, not to mention librarians, archivists, and literary scholars.

Two ways of approaching the answer emerge: first through interpretation, by showing how interface elements such as icons and windows in different operating systems and versions may affect how we understand the work; and second through description, adding to our vocabulary terms such

---

14 When it appeared in 2002, this article had a catalyzing effect on many textual scholars, especially those of the generation that had grown up with personal computers in the home. Its importance was recognized with the Society for Textual Scholarship’s prestigious Fredson Bowers Memorial Prize in 2003. The ideas presented in the article were developed in Kirschenbaum 2008.
as *layer, version, release, object, state, instance,* and *copy.* These terms bridge the formalized languages of programming and descriptive bibliography, two worlds that make remarkably similar investments in precise language and meaningful distinctions. The bibliographical edge to Kirschenbaum’s approach allows him to delve beneath the surfaces of digital artefacts, illuminating the facets of material construction and software design that many literary hypertext enthusiasts and cyberculturalists have tended to pass over or ‘mystify’ (Kirschenbaum’s word) with weak, off-the-shelf interpretations of poststructuralist theory (2002: 25). Searching for an exemplar of digital textual scholarship, Kirschenbaum’s article hearkens back to two recognizable strengths of the past century’s bibliographical tradition, one being McGann’s materialist hermeneutics, and the other, the New Bibliography’s rigour in accounting for the physical features of books.¹⁵

Yet for all its innovation, this early example of digital textual scholarship also relies upon a conservative view of scholarly editing as fundamentally preservational — an updated version of Greg’s 1932 dictum that ‘Books are of value in proportion as they preserve the past’ (1998/1932: 136). As textual scholarship extends its scope to include digital artefacts, it must do so while itself changing from within. In seeking to avoid the weak version of poststructural criticism, with its ill-informed descriptions of digital texts as inherently unstable and non-physical, Kirschenbaum’s analysis risks jettisoning what we might call strong poststructuralism, whose influence on textual studies has prompted resistance to the idea of stable origins, interest in texts as mediators of power and not just as bearers of aesthetic worth, questioning of the construction and uses of canons, and valuing of multiple authority as richness.¹⁶ If hypertext theory in the nineties failed to understand how digital texts work beneath the surface, the computing humanists who did understand tended to underestimate poststructuralism’s abiding influence. Susan Hickey, for example, mischaracterizes the relationship between textual studies and electronic editing: ‘the major difference between a printed and an electronic edition is that a fairly standard and well-documented model has developed for a printed edition, but no such thing exists for an electronic edition’ (2000: 133). Even eight years later this remains an insightful statement about electronic editions, but it overlooks the profound changes the print ‘model’ underwent in the wake of the

¹⁵ The appendix to Kirschenbaum’s article, titled ‘Towards Some Principles of Computational Description,’ is a deliberate echo of Fredson Bowers’s landmark 1949 book, *Principles of Bibliographical Description.*

¹⁶ The phrase ‘multiple authority is richness’ comes from McLeod 1982: 421.
New Bibliography’s dethroning through the eighties and nineties, which
drew force from the influx of poststructuralist theory in literary studies.
Although textual scholarship often presents itself in a conservative light
as a conduit of tradition and guardian of cultural heritage, its own future
depends upon recognizing, pace Greg, that all recorded texts are also of
value in proportion as they provoke thought and change in the present.

3. Interface and the Stakes of Design

Long-term preservation of digital heritage begins with the design of reliable
systems and procedures which will produce authentic and stable digital
objects.

UNESCO Charter on the Preservation of Digital Heritage, article 5

Tensions between tools and materials in the digital humanities manifest
themselves in both the design and analysis of digital artefacts. In particu-
lar, the preservation imperative described above brings cultural pressures
to bear upon all textual scholarship, digital and otherwise, such that Greg
uses loaded words when he speaks of books as a ‘precious inheritance’
(1988/1932: 136). Digital texts lack the same symbolic status as documents
like the Magna Carta, Shakespeare First Folio, or United States Declaration
of Independence, each of which confers a sense of material origin upon
master narratives. We can see tensions at work in some of these documents’
digital counterparts on the Web, specifically by reading their URLs for con-
notations of stability and authenticity. Here are two examples:

Universal Declaration of Human Rights
www.un.org/Overview/rights.html

Canadian Charter of Rights and Freedoms
laws.justice.gc.ca/en/charter/

In both cases, human-readability coincides with machine-readability in the
form of the Web address, which in turn confirms the stability of the content
of these foundational documents. Future stability of such digital artefacts
is the concern of UNESCO’s Charter on Preservation of the Digital Heri-
tage, which states that ‘The purpose of preserving the digital heritage is to
ensure that it remains accessible to the public’ (article 2). But contrast that
document’s own URL with the ones above:

UNESCO Charter on Preservation of the Digital Heritage

Although it is possible to find a slightly simpler URL that brings us to a PDF
version of the document (UNESCO 2003), this unwieldy chunk of code is
the closest thing we have to a stable address for the Charter in the native
format of the Web. As a digital document, the Charter says one thing but
does another, creating a contradiction between its content and form: the
aspirations of cultural heritage pull in one direction while the design of the
code pulls in another.

These tensions become visible in digital objects through the double-
vision that characterizes textual scholarship: to see at once both the signifi-
cyng surface and what lies beneath. By nature textual scholarship resists
the fallacy of screen essentialism, the tendency to essentialize digital text
as ‘easily erasable pixels of light flickering on the screen’, as Marie-Laure
Ryan does in one of the canonical articles of hypertext theory (1999: 95).
In Kirschenbaum’s definition, screen essentialism depends upon ‘the bias
towards monitors and display devices in new media studies, where the
vast preponderance of critical attention has been focused on what happens
on the windowed panes of the looking glass’ (2004b: 95). The term comes
from Nick Montfort’s critique of certain biases in new media studies:

When scholars consider electronic literature, the screen is often portrayed
as an essential aspect of all creative and communicative computing — a fix-
ture, perhaps even a basis, for new media. The screen is relatively new on
the scene, however. Early interaction with computers happened largely on
paper: on paper tape, on punchcards, and on print terminals and teletype-
writers, with their scroll-like supplies of continuous paper for printing out-
put and input both. (Montfort 2004: [n.p.])

Under such conditions there was a more consequential distinction between
input and output processes than we generally experience with PCs, some-
times involving a gap of days between the submission of input and the
receipt of output from a large, shared mainframe. This is not to suggest
that screens are unimportant, but rather that critics need to balance their
attention to computers as objects with an understanding of computing as
process, in which the screen is but one layer of interface. To see the algo-
4. The Human Presence in Digital Artefacts

rithm within the UNESCO document’s cumbersome URL is to understand the contextualizing system, just as the Scientific American illustration makes a point of revealing the system that humans normally cannot see (at least without a wrecking ball). When reading digital artefacts, textual scholars might question the conventional wisdom that the only good interface is a transparent one.

If textual scholars tend to position themselves at the threshold between the surfaces of texts and their mysterious depths — between Erasmus’s ‘smiling fields’ and the New York Public Library’s buried stacks — then digital materials may lead them to new kinds of thresholds. As in bibliography, questions about preserving and reading digital artefacts lead inevitably to the topic of their design. Reading the human presence in a digital artefact requires knowledge of markup, encoding, and even programming, though such complexity can be productive.

The greatest conceptual difference between book history and interface design lies in their temporal orientations. If textual scholarship remains focused on the past, interface design is naturally oriented toward the future. Interface design is all about how things should be, how to improve the deliverable yet to be delivered. This temporal orientation manifests itself rhetorically. Design gurus like Jakob Neilsen and Bruce Tognazzini tend to intone their advice in the imperative, often synthesizing vast amounts of data into PowerPoint bullets. For example, Neilsen offers the sensible dictum that ‘Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution’.17 In this sentence Neilson implicitly looks forward to a time when error messages make sense, when cryptic ‘404: not found’ errors in web browsers no longer lead novice users to wonder if there were really at least 403 other mistakes they could have made. This orientation spans the continuum from professional to academic writing on design, including work that is not overtly part of the ‘how-to’ genre.18

---

18 For example, see Dillon 2004 and any of Tufte’s books, such as Visual Explana-
By contrast, book history is just that: history. It looks back to how things were, even in the very recent past, and how they came to be as they are. To those ends, its chief products are narratives in the form of scholarly books and articles. Increasingly though, the term history of the book is expanding into history and future of the book, generally a positive development but by no means a straightforward one, since the future is not available for study in the same way as the past. The digital humanities’ most productive response to this difficulty has been to ask ‘why speculate when we can prototype?’; that is, to regard the future of the book as something we create, not just observe and comment upon.

This difference in temporal dispositions does, however, lead to trade-offs. If traditional textual scholarship can seem too historical, with its preservation imperative acting as a brake on any experimental tradition from within, then so too can interface design be accused of not being historical enough, sometimes uncritically assuming a synchronic view of readers and texts that ignores cultural, historical, and political contexts. This is not to suggest that design lacks an historical discourse of its own; quite the contrary. For example, the history of design is a core component of most degree programs in the field, and Edward Tufte’s books on information design exemplify a breadth of historical and cultural materials that most textual scholars would admire. However, a substantial part of computer interface design has nonetheless developed in a way that disregards historical understanding as central to the knowledge it produces. Similarly, Tufte himself draws on the work of book historians to offer his insightful readings of ‘visual confections’ in seventeenth-century English books (1997: 122-5, 134-6), and Andrew Dillon begins a chapter on reader studies with an epigraph from Bacon’s essays (2004: 3), but neither really contextualizes his argument within the specific concerns of the seventeenth century, the way a literary critic or cultural historian would. History here is not even background, let alone context; it is only a source of materials. Writing like a cultural historian is not Tufte’s nor Dillon’s purpose, since the epistemological context for their approaches is not history but cognitive science, just as it is with the interface design coming from computer science. That is by no means a weakness — both of their works cited here are excellent introductions to their topics — but it is a difference that must be acknowledged.

For digital textual scholars the problem is not that the humanities and (social) sciences are different, but that a psychological or historical
4. The Human Presence in Digital Artefacts

perspective may present itself as the only valid one. This essay sides with the historical perspective mainly because it has been neglected in interface design, which looks instead to cognitive science for its epistemological framework. Ronald Day suggests that the origin of this trade-off between disciplines was the constitution of information studies in response to corporate and military needs in postwar America, resulting in

willful ignorance of Marxist, non-quantitative, non-‘practical’, and, largely, non-American analyses of information — analyses of information and society and culture have almost totally been given over to so-called information specialists and public policy planners, mainly from computer science, business and business schools, the government, and the quantitative social sciences. (Day 2001: 5)

The consequence for related fields such as human-computer interaction has been what Day calls a problematic ‘focus on quantitative methods of analysis, a neglect of critical modes and vocabularies for analysis, a dependence on naive historiographical forms […]’, and a neglect of art and culture outside of conceptions of historical transmission (that is, “cultural heritage”’) (2001: 5; emphasis added). Tufte’s work deftly bridges the cognitive and the qualitative, but the overall disciplinary trade-off Day describes may well account for blind spots such as the design of the UNESCO Charter as a digital artefact, as well as problems in other digital projects that lock their materials in a conceptual box labelled cultural heritage. Many of the literary texts which bear that label — Hamlet, Ulysses, The Canterbury Tales, The Prelude — have complex histories of transmission intertwined with interpretive concerns, and textual scholars may receive new kinds of illumination from the material histories of films, audio recordings, graphic novels, and video games. The humanities’ investment in the inner complexity of materials productively complicates the task, as McGann describes it, ‘of re-editing — of representing — in digital form the entirety of our received textual and documentary archive’ (McGann 2001: 194). Digital textual scholars have found themselves charged with building a new humanities archive using someone else’s tools.¹⁹

4. Religious Issues: Form and Content

The sonnets of Shakespeare remain the sonnets of Shakespeare even in the most abominable edition. Nor can the finest printing improve their quality.

Aldous Huxley, Introduction to *Printing of To-Day* (1928:1)

With a few exceptions, accounts of the intellectual and institutional transformations of the digital humanities often overlook the tension that results from using other disciplines’ tools: pulling in one direction, computational practices mandate the abstraction of content from the details of its presentation; pulling in the other direction, literary studies now values those very presentational details as integral to the interpretation of texts. Many of the software tools computing humanists use today embody the design principle of treating form and content as not only distinguishable (as literary critics do in order to talk about them), but also as divisible into components like XML (eXtensible Markup Language) files and stylesheets. Such matters reach beyond pragmatics; as Alan Liu argues, the ‘cardinal needs of transformability, autonomous mobility, and automation resolve at a more general level into what may be identified as [a] governing ideology […]: the separation of content from material instantiation or formal presentation’ (2004: 58; emphasis changed from original). Liu names this ‘governing ideology’ *transcendental data*, in which the separability of data from their presentation via technologies like XML means that ‘our interfaces today are ever more transparently just […] skins or, put technically, templates, schemas, style sheets, and so on, designed to be extricable [from content]’ (2004: 62; emphasis in original). This ideological formation, when manifested in pragmatic terms, confronts digital textual scholars with the kind of dilemma known as a ‘religious issue’ in programming jargon: is it desirable, let alone possible, to divide the content of a text from its material form for the purposes of machine-readability and large-scale computation?20

It may be helpful to consider first a related question: how did this problem overtake digital textual scholarship? The answer lies in how the theoretical and practical conversations about digital texts have unfolded in the humanities. The topic of interface arose late in the critical discourse, arriving only after others like hypertext and multi-linear narrative had asserted their centrality. A case in point is the final paragraph of McGann’s influential ‘Rationale of Hypertext’ (2001, itself a gesture back to Greg’s ‘Rationale

---

20 On religious issue and related terms, see the *Jargon Lexicon*, http://www.jargon.net/jargonfile/r/religiousissues.html [accessed 19/10/2008].
This frank admission is belated in more ways than one, since the paragraph exists only in the latest published version of McGann’s ‘Rationale’ in Radiant Textuality (2001), not in the earlier versions published in the journal TEXT (1996a) and the book collection Electronic Text: Investigations in Method and Theory (1997). The textual history of McGann’s own ‘Rationale’, one of the most influential critical examinations of hypertext, thus stands as a synecdoche for the critical discourse as a whole.

No less significant, however, is that McGann ends on the word ‘tools’, a term that marks the borderline between the older hypertext theory and newer digital humanities. Hypertext theorists observe the effects of digital technologies; digital humanists actively develop them, embodying the design ethos of thinking through making. (Some, like McGann, do both.) Yet the no-nonsense pragmatism of digital humanists has not insulated them from repeating the mistakes of their predecessors. The late arrival of interface to the theoretical conversation has replicated itself in the lateness of interface design tools native to the humanities.

Although the belatedness of interface has not received due attention, neither has it gone unremarked. Kirschenbaum notes two dangers of deferred interface design in a digital humanities project: first, that a hasty, under-resourced design phase is disproportionate to the influence of that design in the reader’s experience; and second, that deferring the interface assumes content is distinct from, and precedes, form (2004a: 524-5). Presently, the first of these two dangers is diminishing as textual projects like digital scholarly editions incorporate the lessons of usability studies from fields like human-computer interaction (HCI). The more insidious danger

Footnotes:
21 For example, McGann brings up interface as a way of grounding a recent exchange about databases and archives (2007: 1588). From within archival studies, Hedstrom (2002) offers one of the best explorations of the links between interfaces, archives, and digital resources.
22 An example is the Electronic New Variorum Shakespeare (described in Werstine 2008), whose prototype interface was the subject of a Killam Trust-funded usability study in 2007-8 (Moore, Galey, and Ruecker 2008). See also the Orlando
is the methodological creep toward the separation of form and content, which runs counter to the humanist tendency to see the two as distinguishable but fundamentally indivisible. For example, McGann points to Steven DeRose’s proposition that a book is the same regardless of variables such as format (quarto versus octavo) and font (Garamond 24-point versus Times 12-point): ‘So far as I can see, nearly all the leading design models for the scholarly treatment of imaginative works operate from a naïve distinction between a text’s “form” and “content”’ (McGann 2001: 185). Kirschenbaum elaborates on why McGann’s word ‘naïve’ might be warranted:

the weight of established wisdom in a field like interface design rests on a fundamental disconnect with the prevailing intellectual assumptions of most humanists — that an ‘interface’, whether the windows and icons of a website or the placement of a poem on a page, can somehow be ontologically decoupled from whatever ‘content’ it happens to embody. (Kirschenbaum 2004a: 524)

However, what McGann and Kirschenbaum describe here is not merely loose thinking on the part of designers, nor a matter of critical inattention to discipline-specific theoretical discourses, but a basic conflict of values between the text-oriented humanities and other, data-oriented disciplines. Methodology reflects epistemology, and tools can invisibly import assumptions from other fields into the humanities.

What makes the ‘religious issue’ of form and content so fraught is that non-textual scholars like DeRose have only been repeating what had become conventional wisdom in their knowledge domains. That con-

24 Specifically, DeRose invokes the idea that all texts have an essential structure in the form of an Ordered Hierarchy of Content Objects (OHCO), a tree structure of non-overlapping nodes that conveniently matches the structure of all XML documents. The debate over the OHCO theory of text divided critics along the question of the materiality of texts — though some participants might characterize the debate differently — with DeRose, Allen Renear, and their co-authors on the pro-OHCO side, and opposing them McGann, Hayles, and others with links to textual scholarship. From a textual studies perspective, the OHCO thesis lost in theory but won in practice. The materialist hermeneutics and media-specific analysis of McGann and Hayles, respectively, have lost no ground in literary and textual studies, but the OHCO model is everywhere in our digital tools, from the structure of XML documents, to the historical core of the TEI guidelines (see http://www.tei-c.org/
votional wisdom has made its way into the humanities, for example in James Cummings’s explanation of XML’s usefulness to the Records of Early English Drama project: ‘The increasing separation of content from presentational aspects has been fundamental for the interoperability and flexibility that makes XML so valuable’ (2006: 181). But how do we determine value? Practically, such separation is a welcome convenience; intellectually, its value is suspect since it impairs textual scholars’ ability to use text encoding to model the complexity latent in their materials. Although Cummings accurately describes XML’s advantages over HTML for the encoding of born-digital documents like academic articles, his description becomes deeply problematic when applied to the encoding of non-digital materials like manuscript poems (he takes many of his examples from Chaucer). Cummings’s explanation overlooks the crucial distinction between prescribing presentational details and recording them — a non-trivial distinction in textual scholarship, which searches for the human traces left upon material artefacts.25

The seductiveness of form-content abstraction derives from its ability to simplify the task of encoding, even for those trained to appreciate complexity in texts. We can see the different temporal orientations of design and book history mirrored here, as the conventional wisdom governing the design of new digital documents extends, anachronistically, to the reading of textual materials from the past. Cummings and DeRose represent a tradition which, in the spirit of the Huxley epigraph above, regards content as literally that — meaning contained in the language of a text. (Tom Davis, in a tongue-in-cheek paraphrase of Karl Popper, terms this attitude toward information as the ‘bucket theory’ of communication [1998: 106].)26 On the other side of the doctrinal divide, McKenzie’s simple dictum stands in contrast: ‘forms effect meaning’ (1999/1985: 13).27

Guidelines/ [accessed 11/2/2010]), to the Document Object Model that underpins browsers and other Web technologies. See Schreibman 2002 for a balanced overview of the positions; for key entries in the debate, see DeRose et al. 1990; Renear 1997; Renear, McGann, and Hickey 1999; McGann 2001, as cited elsewhere here; Hayles’s chapter on ‘Translating Media’ (2005); and Robinson 2009a.
25 Peter Robinson, for example, argues for the value of encoding the minutest details of presentational information in Chaucer’s manuscripts (1996a). The problem is not that Cummings is unaware of editorial theory; evidence to the contrary may be found in Cummings 2007.
26 The history of the term information is also relevant here; see Nunberg 1996 and Capurro and Hjørland 2003. For a discussion of the concept of information within the context of theorizing tools, see McCarty 2002: 382-3 and 2005: 110.
27 As if to prove McKenzie’s point by accident, a recent book history anthology
Paradoxically, the digital humanities have produced valuable practical venues for thinking through such religious issues (for example, the Text Encoding Initiative), but also communities of practice whose theoretical assumptions are now difficult to dislodge (for example, again, the Text Encoding Initiative).\textsuperscript{28} McGann’s closing argument about interface in the ‘Rationale’ (2001), however belated, was all the more valuable for pointing a way out of hypertext theory’s closed circle of self-confirmation. In retrospect, hypertext theory now seems like a part of the conversation about digital textuality that mistook itself for the whole conversation, but the digital humanities’ practical orientation should not mean a dismissal of theory. ‘The implementation of the theoretical design’, as McGann (2001: 74) calls it, has become no less a moment of theoria for the digital humanities, in the sense of Hans-Georg Gadamer’s distinction of theoria as more than passive observation: ‘It does not mean a mere “seeing” that establishes what is present or stores up information. […] Theoria is not so much the individual momentary act as a way of comporting oneself, a position and condition. It is “being present” in the lovely double sense that means the person is not only present but completely present’, as when one is present in an audience that is fully ‘engrossed in their participation as such’ and fully aware of each other (1998: 31). In this sense of theoria and presence as synonyms, the modelling of human presence in interface design becomes not just the implementation of a theory, but an act of theoria that enables one to think through these complex relationships. Textual scholarship has long exemplified this philosophy, as in Bernard Cerquiglini’s pithy formulation, ‘every edition is a theory’ (1999: 79). Unfortunately, humanities computing practice has bypassed such moments of theoria in its tendency to think of interface design and text encoding as separate activities, each happening at opposite ends of the research plan. As Kirschenbaum points out, interface design too often comes as an afterthought, late in the project schedule as time and resources run out.\textsuperscript{29} The deferral of interface thus represents not so much work left undone as a missed opportunity to articulate what’s at stake in how the humanities understand texts.

\textsuperscript{28} For discussions of the structural impositions of the TEI tagset, with particular reference to the sometimes-vexing <tw> (forme-work) tag, see Lancashire 1996: 123-4, and Bjelland 2000: 24-6.

\textsuperscript{29} See Kirschenbaum 2004a: 524-5.
5. Serving the Particular

[T]he universal in the humanities is in the service of the particular.


The search for human presence in digital artefacts is also the search for the humanities’ place in digital scholarship. Both searches are underway even as the design of digital tools is moving in one direction while the theories of textual scholarship are moving in another. Digital humanists caught between the tensions described in this essay should be wary of two species of the same essentialist fallacy: one, perpetuated by theorists with only a screen-deep understanding of technical matters, asserts that digital texts are inherently unstable (productively or otherwise); the other, perpetuated by computing practitioners who neglect theory, asserts that that the only aspects of texts worth knowing are those which may be modelled digitally. Screen essentialism thus has its counterpart in computational essentialism, and digital textual scholars must somehow navigate between the two. The challenge is to design interface tools that do not force business or (social) science models upon humanists. This step involves developing traditions of programming native to the humanities, and recognizing that programming and computer science are not the same thing. A further challenge is to determine how designing and prototyping, as keystones of the experimental tradition within the digital humanities, should relate to the archival and historical strengths of textual scholarship.

To put all this another way, what does the conjunction and signify in a term like history and future of the book? Is it merely a hasty splice between disciplines, or an expansion of an established field into new territory? In the most optimistic light, the and represents the imperative to find a synthesis between reading historical artefacts and designing for the future. Although interface is by no means the only important aspect of digital textuality, it is the area where such a synthesis is most needed. The response of textual studies should therefore be more strategic than simply pointing out when computing practitioners do not understand humanities materials. For a text encoder working under a computer science model to treat data as extricable from their presentation is consistent with best practice. For a literary scholar to treat texts as inextricable from their presentation

---

30 On the importance of decoupling programming from computer science in the digital humanities, see Crane et al. 2007: 54; for a pedagogical perspective on programming in the humanities, see Rockwell 2003.
is also consistent with best practice. This is the methodological crux facing digital textual scholars of the present and future.

The solution lies partly in rearticulating a simple but fundamental idea to all sides: even as we design new digital artefacts, we are still learning how books work, as well as manuscripts and other textual materials. By looking for something more than the authorial genius loci in texts, McKenzie introduced new ways to answer the question of what is at stake in textual scholarship’s search for ‘the human presence in any recorded text’ (1999/1985: 29). The phrase human presence implies no less complexity than any recorded text, and digital artefacts, like any other kind, bear traces of the social worlds they occupy, not just of their creators. Just as Kirschenbaum rightly questions the easy binary that produces digital texts as unstable, so must we also be vigilant against the myth that, with the rise of digital technology, the human record is now moving from a period of fundamental stability to one of instability. As print- and manuscript-oriented textual scholars have long argued, past textual forms were never so immutable to begin with.\(^\text{31}\) Textual scholarship may be the contrarian voice within the digital humanities, resisting those progress narratives which, in order to justify investments in tool-building, make texts computationally tractable by sacrificing their complexity on the altar of expediency. The serious study of digital artefacts does not replace that of pre-digital materials; rather, the two must progress together or not at all.

Readers of the May 27th, 1911 issue of Scientific American were shown a figure of the New York Public Library that revealed to them the textual depths beneath their feet. In the same gesture, the figure asserted a sense of mastery over the space of the archive, rationalizing it by means both of the retrieval mechanism and of the image’s power to depict it. Digital humanists today are presented with similar representational systems promising a similar mastery. ‘What do you do with a million books?’ remains a worthwhile question, but only if we remember that the humanities’ great advantage is the power to produce new knowledge using only a few books — sometimes even one text. That uniquely powerful economy of scale defines what Appiah calls the humanities’ ‘deeply idiographic character’: discovery comes from ‘a particular poem, a particular painting, a particular sonata’ (2005: 42).\(^\text{32}\) For digital tools and methods to share in that distinctly qualitative power, they must be able to serve the particulars of texts with-

\(^{31}\) For example, see the recent debate between Elizabeth Eisenstein and Adrian Johns in American Historical Review, mediated by Anthony Grafton (Grafton 2002).

\(^{32}\) On the term idiographic as opposed to nomothetic, see note 9 above.
4. The Human Presence in Digital Artefacts

out sacrificing them to the exigencies of the universal. The digital humanities now face a paradoxical challenge: being digital comes naturally; it's the humanities we have to earn.
References


Corpus of Middle English Prose and Verse, quod.lib.umich.edu/c/cme/ [accessed 8 February 2010].


Works Cited

—, David Bamman and Alison Jones. 2007. ‘ePhilology: When the Books Talk to Their Readers’. In Siemens and Schreibman 2007: 29-64.


Works Cited


Fish, Stanley. 1980. *Is There a Text in This Class? The Authority of Interpretive Communities*. Cambridge, MA: Harvard University Press.


Works Cited


Gobet, Fernard and Gary Clarkson. 2004. ‘Chunks in Expert Memory: Evidence for the Magical Number Four … or is it Two?’ Memory 12.6: 732-47.


Grafton, Anthony. 2002. ‘How Revolutionary was the Print Revolution?’ American Historical Review 107.1: 84-87.


Works Cited


— 2004a. ‘“So the Colors Cover the Wires”: Interface, Aesthetics and Usability’. In Schreibmann, Siemens and Unsworth 2004: 523-42.


— 2004b. ‘Cybertextuality’. TEXT Technology 2: 1-18


Text and Genre in Reconstruction


— 2006. ‘Presence and Genre in Digital Scholarship’. Paper delivered at Master Classes in Textual Studies, De Montfort University, 1 June.


Miller, George A. 1956. ‘The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information’. Psychological Review 63: 89-97.
Works Cited


Works Cited


Robinson, Peter, ed.


238 Text and Genre in Reconstruction


— 2003a. ‘Where We Are with Electronic Scholarly Editions, and Where We Want to Be’. In Jahrbuch für Computerphilologie 5: 125-46. computerphilologie.uni-muenchen.de/jg03/robinson.html [accessed 19 January 2009].


— 2004. “Where We Are with Electronic Scholarly Editions, and Where We Want to Be”. In Jahrbuch für Computerphilologie 5. 123-143. computerphilologie.uni-muenchen.de/ejournal.html [accessed 8 February 2010].


Text and Genre in Reconstruction


Works Cited


The oldest human artifact we can find, is in the paleolithic age—the period in human prehistory distinguished by the original development of stone tools that covers approximately 95% of human technological prehistory. The period begin ca. 500,000 years ago. Some of the oldest artifacts on this list predate Homo sapiens and were most likely created by early human ancestors such as Homo erectus. Venus of Hohle Fels. Löwenmensch Figurine (Lion-man of the Hohlenstein-Stade)