Jerome McGann

THE GUTENBERG VARIATIONS

Say “TEXTUAL SCHOLARSHIP” and think “dryasdust,” and then perhaps see an obsessed St. Jerome alone in a cave or a monk’s study humped over a great tome with a pen in his hand, a skull and an hourglass somewhere nearby, and perhaps a lion and a dog sprawled at his feet. That complex figural expression was one of the central emblems of the Renaissance, when textuality—powered by the Gutenberg revolution—fairly came to define “the human condition” as a textual condition. But a certain kind of textual condition, a condition of “The Word” conceived as written or printed, a Word to be read.

Recent textual scholarship has been founded on an awareness that the uneasy balance between textual criticism, on the one hand, and interpretational studies, on the other, had reached a condition of crisis that required serious reflection and investigation. It was a crisis that had begun sometime in the 1960s; E. A. J. Honigmann’s The Stability of Shakespeare’s Text might be reasonably cited as an early defining moment.1 Honigmann’s book invited a new approach to Shakespearean criticism, editorial as well as interpretational, based on the evidence of asymmetries of various kinds in the received corpus.

Textual studies are traditionally (and perforce) grounded in minute and material particulars; they frequently concentrate on social, historical, and other contextual matters and materials; their critical reflection


© Jerome McGann, 2002. All Rights Reserved.
tends to focus on practical questions of method and procedure; and they may be practiced without making explicit connection to issues of interpretation of any kind, whether hermeneutical or ideological.

Interpretational studies, by contrast—“Theory,” so called, is a special, recent type of this general procedure—are typically speculative and conceptual. Their focus is in self-reflection, and their principal concern is with meaning, whether that is conceived in social, psychic, or formal terms. This conversation was emerging in other parts of the literary academy at precisely the same time. “New Historicism” and “Cultural Studies” both exploded on the scene in the 1980s as a consequence of these institutional tensions.

But the character of these exchanges has altered dramatically for one simple reason: the emergence of the World Wide Web in 1993. That event brought the clear realization that a new textual condition was at hand and that traditional literary and textual studies had an enormous stake in it. One could now see quite clearly that digitization was both the medium and the message concealed in the crisis that had been developing in literary and cultural studies since the mid-1960s.

Why? Because the Web exposes how the technology of archival and bibliographical exchanges can be radically expanded in both spatial and temporal terms. Scholars can interact with each other anywhere in the world, can exchange their work in various new ways, and can access materials located in remote locations. They also can execute remarkable new critical transformations of their subject matter, thus opening unexpected opportunities for investigations of many kinds, bibliocritical as well as interpretational.

We are thus entering a period when the entirety of our received cultural archive of materials, not least of all our books and manuscripts, will have to be reconceived. The initial stages of this reconception, which is well under way, have been largely confined to work with archives and
libraries whose holdings are being digitally repossessed in many new ways. And because these depositories are the ground of all traditional scholarly work, these institutional changes will have—are already having—radical effects. This is very much a material revolution, and in negotiating it, we all—not least of all traditional scholars—would do well to recall Marx’s eleventh thesis on Feuerbach, which has acquired interesting new meanings beyond those originally conceived by Marx: “The philosophers have only interpreted the world in various ways; the point, however, is to change it.”

Information scientists and systems engineers will be (already are) much involved with these changes. But it is the literary scholar, the musicologist, the art historian, etc., who have the most intimate understanding of our inherited cultural materials. Hence the importance that traditional scholars gain a theoretical grasp and, perhaps even more important, practical experience in using these new tools and languages. For “theory” in this volatile historical (and historic) situation will have little force or purchase if it isn’t grounded in practice.

We need to step back a bit from the immediate hubbub of digital enthusiasts and cultural (especially literary) doomsters and try to think coolly about what is happening and what needs to be done. For simplicity’s sake, I will focus my comments here on literary works. Nonetheless, it’s clear that the issues have the broadest kind of cultural relevance.

First of all, textual studies, bibliography, and editorial theory and method have to be replaced at the center of scholarly work. They are “foundational” because they are the means by which we preserve an accessible and accurate cultural memory. Always threatened by the erosions of time and its patient marauders, our material archives are now further...

threatened by the power and the very promise of these new technologies. Digitizing the archive will be a major undertaking for many decades to come, but it is a process that involves many difficult and problematic choices. How and by whom will these choices be made?

That is an ominous question given this fact: Only a handful of disciplinary scholars are positioned, practically and intellectually, to make or assess these decisions. This diminished number is the direct result of some seriously misguided academic/institutional decisions of the past 30 or 40 years, when the requirement of textual studies and bibliography was gradually dropped from most Ph.D. programs in English language and literature. It must be said, of course, that these decisions arrived because of the situation of the previous thirty or forty years, when an agreement was reached between textual scholars and literary interpreters to segregate their two worlds. Like other segregation programs, this one brought dire results.

Only the force of historical circumstance—the emergence of digital technology and its invasion of the cultural arena—has driven home the true scale of the problem. The digitizing of the archive needs direction by scholars who have a high-level understanding of the theory and method of textuality. But just at the moment when these persons are most needed, we discover how few we have trained. We have permitted too wide a gap to emerge between the disciplines of “the two cultures.” We have neglected a programmatic approach not only to instruction in bibliocritical skills and methods, but also of the theoretical and practical relation of these skills to the broad problems and questions of cultural interpretation.

Here is a true story of the level to which the general ignorance about texts and textuality has descended. I have in front of me a draft report of the U.S.-based Task Force on the Artifact in Library Collections. This group of “scholars, librarians, and archivists” is making recommenda-
tions on how to “preserve the valuable evidence in research institutions and libraries.” Remarking on the “consequences of the adoption of print technology,” the authors observe that the primary one has been “the creation of a fixed and stable text.” They go on to gloss this deeply problematic idea in the following way: “Copies of a single edition of a book are made to be indistinguishable one from another; the text does not vary among copies. People have grown used to the idea of a canonical version of any given text; there is, for example, one fixed text of Leo Tolstoy’s *War and Peace.*”

This statement is so depthlessly ignorant that it takes one’s breath away. It also is frightening. I don’t mean frightening merely because a group of (presumably distinguished) “scholars, librarians, and archivists” hold such views, nor even because, holding them, they might actually try to implement them in practical ways in our libraries. The statement signals a widespread scholarly and intellectual condition.

At my own university, the University of Virginia, we have a distinguished tradition of textual scholarship. Our current English department was shaped, after all, by Fredson Bowers. (It’s true that Bowers’s work often fostered “the idea of a fixed and stable text,” but he did so in the sure knowledge that such an idea was quite unrealizable—an enabling function for controlling an analytic method.) We also have a number of related institutional resources devoted to the study and preservation of books and the translation of these things into digital form. With certain exceptions, however, even here most of our graduate students have no clear understanding of theory of textuality and its socio-institutional formations, let alone a practical grasp of bibliography or scholarly editing.

3. The final report is available at http://www.clir.org/pubs/reports/pub103/contents.html. This is the final report where the cited quotation from the draft report (section 3.1) was expunged when various scholars raised an outcry.
Nicholson Baker and others register the problem at the practical level of the library and the decisions that have to be made about what to preserve and what to give up.\(^4\) The jeremiad of *Double Fold* has garnered a lot of press, and anyone who cares for paper artifacts has to be grateful for the wake-up call it makes. But precisely because so much is at stake, precisely because cool judgments are now so much needed, we want to avoid being either put off or swept away by Baker’s often self-indulgent rhetoric. One key thought in his book should be held fast: “Very few people knew any of this was going on.”\(^5\) Disastrous acts of archival culling have occurred, and Baker has exposed some recent and flagrant mistakes. These happen, alas, all the time, which is why scholars in particular must be vigilant and regularly involved in these kinds of institutional decisions. Preserving our cultural memory is a complex, difficult, and expensive task. Because not everything can be preserved in original forms (something Baker perhaps tends to forget), careful judgment has to be exercised and many different interests need to be consulted.

My immediate concern is not with materials preservation, however, but with the related one that deals with the digital transformation of the archive in a general way. Partly this transformation involves the electronic organization of library resources, and partly it involves the digitizing of those resources. In each case, informed judgment rests in an informed textual scholarship.

The sad truth is that most people—even many scholars—have only a casual interest in the material complexity of texts and books or in the need to distinguish their various particular differences. Don McKenzie once told me that he liked to begin his bibliography seminar by passing around a printer’s dummy book and then asking his class the question:

\(^5\) Ibid., p. XX.
“When this book was printed, what was in it?” The initial blank response, a parody of the book, yielded to an interesting discussion as soon as McKenzie encouraged his class to think about the material object and the implications of its formal properties.6

Here is another tale, one I tell to my graduate students.

Byron published the first installment of his masterpiece *Don Juan* in 1819 in a handsome quarto on good rag paper. Because the book did not have either his name on the title page or the name of his publisher, it appeared with no copyright safeguards. Consequently, the quarto text was immediately reissued in cheap duodecimos by a number of piratical publishers. Observing this publishing scene, the *Quarterly Review*’s conservative and anonymous reviewer of the poem understood that although he had basically a single text on his hands, he was faced with at least two different literary works carrying very different meanings. On the one hand was a poem of remarkable elegance and wit, an ornament for those who could afford to buy the guinea quarto edition; on the other was a poisonous work pampering to depraved imaginations. And what made this Jekyll and Hyde? In the reviewer’s own words: “*Don Juan*’ in quarto and on hot-pressed paper would have been almost innocent—in a whity-brown duodecimo it was one of the worst of the mischievous publications that have made the press a snare.”7

This story, first told in some detail in 1965 by Hugh J. Luke, has long been available to scholars of romanticism. Many similar scholarly tales might be told, for the truth is that every literary work proliferates in multiple material forms with multiple, often contradictory, meanings that are a function of those material forms. Nonetheless, the import of

such facts—perhaps even their material existence—remains widely unrealized as that task force of “scholars, librarians, and archivists” sadly shows. They are not alone in their misery: Many people, many scholars, have to be reminded that Don Juan (like War and Peace, etc.) is neither fixed nor stable. Its meaning—its very identity—varies with its documentary variations, and those variations are a function of the work’s material, historical, and dialectical existence. Every document, every copy of every edition, is unique; and this uniqueness grows more clear, more articulate, more significant as the document undergoes its historical passage. The task of describing and explaining such things—it is an art and a science—is a primary obligation of the literary scholar, perhaps the primary obligation.

As human beings we have many other obligations, but as scholars our chief vocation is memorial—to preserve and transmit an accurate historical record. Such a record will be, perforce, multiple and complex. It is a record not so much of “primary texts” as of primary transmissions. And we cannot know what this record means, or imagine what it means in interesting ways, unless we know how to locate and assess its multiple material forms.

The removal of documents from an archival depository is thus a hazardous event, as Baker’s work should remind us all. It can be undertaken only with the greatest possible care and thought. Equally hazardous is the process, which gains ground every day, of translating these primary materials into digital forms. The importance as well as the difficulty of doing this may be measured precisely by the power and flexibility of these new transformational processes. So much more can be done, for better and for worse, through digitization than could be done, say, with microfilming.

The opportunities and prospects, which are beyond doubt fascinating, are also beyond doubt dangerous and replete with illusions. (One of the
The worst current illusions is that book space is stable while digital space is volatile—an illusion that Baker, with his laudable devotion to books, sometimes promotes.) The materials and tools, the new as well as the traditional, are complex and difficult. Furthermore, the process of translation is expensive in terms of human and financial resources. Finally, the work of transformation is being undertaken by people—ourselves—who have much less knowledge than we need to do the work well. We simply don’t understand the technology of computerization as well as we should. But the problem lies not simply in the novelty of computers. When we handle these new instruments, we expose as well the novelty, so to speak, of our books. Translating paper-based work into digital forms involves regular and humbling revelations for traditional scholars, who discover in this process how much we don’t know about the materials we’ve worked with all our lives. As literary scholars we think we know what books and their novels and poems are. But, in fact, these amazing works—they are our “humanist” machines—have long defeated our best efforts to understand them. What we do is make them. How little we understand what we are capable of making becomes painfully, ludicrously clear when we try to explain to a computer how they work.

So in negotiating this epochal moment, we come to realize the daunting scale of what we need to know. And yet, the work has begun and will proceed, as it must, with increasing speed. We cannot wait for better knowledge to arrive; we have to create the knowledge we require. Singing Gutenberg elegies has been understandable, but now, as Our Lady of the Flowers said to his judge, we are already beyond that.8 The notable increase in collaborative work by scholars in very different disciplinary fields testifies to what we are being called to do.

Traditional disciplinary scholars will contribute most to this effort not by becoming programmers or by switching to computer science, but, first, by trying to gain a clearer grasp of our own disciplines and their materials, and, second, by working with scholars in new kinds of interdisciplinary arrangements. These collaborations (for example, between engineers and hermeneuts) have rarely been pursued before. Now they are imperative. The digitization of the archive can only be done responsibly if it can be carefully and accurately described to computers, which have—press reports to the contrary—very simple (and therefore very difficult) minds. The human mind, on the other hand, is very complex (and so also very difficult). It is, after all, human beings who build computers capable of defeating chess masters.

The technical problem of translating our received cultural heritage into digital forms thus opens a remarkable humanistic opportunity. Resolving the technical issues hangs upon the need to rediscover and re-imagine the functional forms of paper-based works. Think about what books do. Like computerized information tools, the book performs two basic functions: It is a medium of data storage and transmission; and it is an engine for constructing simulations. The first is primarily informational, the second primarily aesthetic and rhetorical.

As our bibliographical world develops what Steven Johnson calls an “Interface Culture”; these two semiotic realms—textual and digital—find themselves caught in a series of tense and complex struggles that make progress difficult.9 But we should not dare to go forward without, like Aeneas, bringing our household gods along with us. Traditional semiotic tools instruct us about digital technology as much by their differences as their similarities, and vice versa. The study of the book in this emerging digital age is a critical imperative for exactly those reasons. The pragmatic problem of digitizing the received archive

involves us in foundational questions of textuality and general semiotics.

I can best illustrate what is involved here with another story, the last and longest I shall tell. Until the emergence of the Web and its amazing graphical presence(s), scholars paid little attention to the imaging capabilities of digital processing. The discipline of humanities computing developed first in the field of linguistics, where computers could be designed to carry out useful automated pattern searches across large bodies of purely linguistic data. As a consequence, computerized approaches to textual corpora between the mid-1950s and the early 1990s implicitly reinforced text-theoretical approaches that minimized, even “invisibilized,” the graphical features of textual works.

The coming of the Web proved a revolutionary textual event because of the difference it made not only for our conception of digital environments, but also for our understanding of traditional texts. The lack of concern with graphical (or auditional) meaning and information in the first 40 years of humanities computing replicated the influential distinction made by traditional textual scholars between the so-called substantives and the accidentals of a text (in gross terms, the punctuation and paratexts versus the words). These distinctions ultimately reinstalled the idea that one can break units of text into a “content” unit and a “form” unit. Humanities scholars working with computers thus began, and in the early 1990s completed, a major effort to develop a text-encoding scheme for preparing humanities texts for digital environments. This movement involved the extrusion from the logical markup language SGML (Standard Generalized Markup Language) a subset specially designed for typical humanities texts. This subset, developed by the Text Encoding Initiative (TEI), began to be widely adopted in humanities disciplines in 1993, precisely the year the Web made its dramatic appearance.
A contradiction of great significance thus arose into view in 1993. The Web began unfolding its Interface Culture at the very moment that a nongraphical, purely linguistic approach to textuality was striving for hegemony in humanities computing. The contradiction proved especially illuminating because it clearly overlapped with a controversy that had been raging for 10 years among traditional textual scholars. The latter involved a conflict between the dominating line of text and text-editing theory, which focused on the linguistic object, and a so-called social text approach promoted, for example, by my *Critique of Modern Textual Criticism*¹⁰ and by Don McKenzie’s great Panizzi Lectures of 1985, *Bibliography and the Sociology of Texts*.¹¹ In this social approach to textuality, the whole of the bibliographical field—minimally, linguistic codes as well as bibliographical codes (which would include Genette’s “paratexts”)—would have to be theorized for inclusion in any actual editing project that aspired to comprehensiveness.

Although this social approach to textuality gained considerable ground in interpretation theory and literary pragmatics generally, it was hotly debated among editors and editorial theorists. Early in the debate, Fredson Bowers argued that a social approach to editing was not an editorial method at all but, rather, a type of hermeneutics;¹² and Trevor Howard-Hill delivered what he took to be the most damning critique of all: However attractive such a method might appear in theory, it could not be implemented.¹³ Editors edited texts, not books.

But the truth was that the emergence of digital culture had made the theory of a social text a practical editorial possibility, even a demand.

---

We began to design and build The Rossetti Archive at University of Virginia in 1993 in order to show one way of implementing the theoretical proposals offered in the 1980s by McKenzie and myself. At this date, a number of scholars have embarked on electronic editions that follow the design principles of a “social theory of texts.” In such an approach, editors are called to bring their critical apparatus to bear on the entirety of the bibliographical field and not on the (linguistic) text alone.

The relevance of this for electronic scholarship and “editing” becomes very clear as soon as one examines the theory of textuality being implemented by TEI out of its SGML base. The theory is called the OHCO Thesis by its proponents. It holds that for computational purposes a literary work must be conceived as an Ordered Hierarchy of Content Objects (OHCO). Three implications of the thesis must be stressed. First, the thesis follows a traditional text-theoretical approach in regarding the object of attention to be a linguistic object. Second, the linguistic object is conceived to be ordered according to a hierarchical structure of parts. Third, these parts are “content objects,” that is to say, linguistic forms ideally conceived. In Saussurean terms, signifiers of various kinds are marked so that an ordered hierarchy of signifieds may be prepared for computerized analysis. Briefly, then, the OHCO Thesis treats its texts—humanities texts, we want to remember—as hierarchically arranged structures of information.

Anyone interested in literary works, however, or in textual objects that develop their aesthetic and/or their material expressive resources, is immediately struck by the serious limitations of such a theory of textuality. Poetical works, insofar as they are poetical, are not expository or informational. Furthermore, because works of imagination are built as complex nets of repetition and variation, they are rich in what informational models of textuality label redundancy and even “noise.”
Such phenomena are generically intrinsic, however, and no imaginative work can exist without systems of “overlapping structure.”

The most important achievement of the implementation of SGML/TEI may well turn out to be the exposure of its limitations in face of the complex laws that operate in so-called natural language. (In fact, there is no such animal; in text-encoding terminology, all language is marked language.) The developers of TEI, to their great credit, were the first to call attention to some of the limitations of their own model. The system continues to be modified and adapted, but its basic structure inhibits what can be gained through modification. So it appears certain that other approaches to computable text markup will emerge exactly because they are needed by the persons who are most invested in the study and use of texts, both traditional and digital.

In our present computational view of natural language forms—clearly a primitive view—textualities are overrun with ambiguities and redundancies. But, in fact, these rich forms are the basic units of meaning and expression, as the poets have always known and shown. That we have yet to theorize the laws governing the way language works comes into sharp focus when digital tools demand an explanation of the rules of language—that is to say, an explanation of the markup protocols that natural languages incarnate as texts. At that point, our conceptual deficiencies leap to attention. When Susan Hockey observes (in her recent book Electronic Texts in the Humanities) that “There is no obvious unit of language” she has, implicitly, defined the real magnitude of the issues at stake.14 These are not, needless to say, programming issues.

Hockey’s remark is no cause for despair, however, nor a debunking of the efforts of those who pursue text encoding for digital tools. The failure of these pursuits spurs new efforts, and the efforts are fueled by

The quest for better and clearer ideas about the complex power of our traditional texts. Though the Gutenberg elegists do not see this, one of the most important consequences of computerization has been new and more precisely realized understandings of the structure of textuality and the behavior of its forms. In a notable essay, John Unsworth has called these results “The Importance of Failure.” Failure is the mother of invention, as we see in the improvements in humanities computing that have emerged over the past decade. Great changes loom even at the horizon of interpretation theory and method. Computational procedures, for example, allow us to expose textual activity at probabilistic levels and to express those quantum readings in analogue and graphical forms. These capacities clearly offer the prospect of new kinds of critical dialogues with our textual inheritance, thus lifting our intuitions to greater self-consciousness.

In similar ways, the coming of the Web is lifting those veils of familiarity, as Shelley called them, that imprison our thinking in old habits. The dramatic character of thousands of interface forms has surely deepened the significance of the famous modern proverb “A picture is worth a thousand words.” We want to know why and how that statement is true in its many different functional conditions: in cathedrals, in films, in traditional textual works, in media. Because all of these forms and languages can be rendered digitally, we are led to rethink them in foundational ways, beyond their powerful and familiar transparencies. The exercise of building digital simulations of phenomena forces a reflexive exposure of one’s views about these things and thence a critical analysis of the limits and flaws in those views. Indeed, such a focus on images and graphical forms has begun to dissipate the common conception, powerfully reinforced by the advent of scientific thinking, that the clearest and most powerful ideas are abstract ideas. The

history of mathematics itself is littered with breakthrough tales that hinged upon the appearance of revelatory images. The more abstract computations followed those visionary events.

Years before the emergence of our full-blown Interface Culture, Gregory Bateson made the following shrewd remark: “Speculation suggests that image formation is perhaps a convenient or economical method for passing information across some sort of interface.” The thought underlies all the interesting parts of Edward Tufte’s studies in *Envisioning Information*. Why should these graphical forms appear so crucial? Because data and information function dialectically. Even the effort to reduce data to self-identity is a dialectical move, an action designed to facilitate data transformations of a richer order at the level of an interface.

It is possibly (probably?) true that traditional text space comprises such a powerful markup code because of its complex spatial features. Readers who concentrate on a work’s linguistic codes tend to overlook the interface in which they are incarnated. But we gain access to our linguistic information only through the economies of bibliographical and typographical design, whose ordering laws we have sadly neglected. No one should be surprised that Douglas Hofstadter’s remarkable semiotic enquires have regularly orbited, as he himself has told us, around the mysteries of typefaces and their design. The deepest meanings—isn’t this what Oscar Wilde said?—are ornamental. Hidden at the surface of things, like Poe’s purloined letter.

Marshall McLuhan was an important prophet of the events I have been describing, but not even he foresaw the changes that digital technology would bring. In many literary quarters, and not without

reason, this upheaval is perceived in dismal terms, as we know. But perhaps
what we are hearing now are not Gutenberg Elegies but, rather, Gutenberg
Variations.

Information scientists and systems engineers will be (already are) much
involved with the changes I’ve been discussing. But traditional disciplinary
scholars have the most intimate understanding of the cultural materials
that are at the center of these transformations. Hence the importance that
they—we—gain a theoretical grasp of these new tools and languages and,
perhaps even more important, a practical facility in their use.