

# IMAGINING WHAT YOU DON'T KNOW: THE THEORETICAL GOALS OF THE ROSSETTI ARCHIVE.

JEROME McGANN

---

I

In a trenchant meta-theoretical essay, Lee Patterson investigated what he called "The Kane-Donaldson *Piers Plowman*", that is to say, the 1975 Athlone Press edition of the B Text. I say "meta-theoretical" because the edition itself constitutes the primary theoretical event. Patterson's essay elucidates the theory of that extraordinary work of scholarship.

According to the editors' themselves, their edition is "a theoretical structure, a complex hypothesis designed to account for a body of phenomena in the light of knowledge about the circumstances which generated them" (212). Needless to say, this "body of phenomena" is problematic to a degree. Patterson studies the evolution of Kane and Donaldson's "complex hypothesis" about these phenomena as the hypothesis gets systematically defined in the edition itself. These are his conclusions:

As a system, this edition validates each individual reading in terms of every other reading, which means that if some of these readings are correct, then--unless the editorial principles have been in an individual instance misapplied--they must all be correct. This is not to say that the edition is invulnerable, only that criticism at the level of counterexample. . . is inconsequential. . . . Indeed, the only way [criticism] could be effective would be if [it] were part of a sustained effort to provide a contrary hypothesis by which to explain the phenomena--to provide, in other words, another edition. (69)

Patterson's startling last judgment--deliberately outrageous--is not simply a rhetorical flourish. He is aware of the intractable character of the *Piers Plowman* materials. But he admires, justifiably, the comprehensiveness and the rigor of the Kane-Donaldson work. Even more, he admires its visionary boldness. In thinking about Kane and Donaldson's project, was Patterson also thinking of Blake?

I must create my own system or be enslaved by another man's  
I will not reason and compare. My business is to create

If he wasn't, he might have been, perhaps he should have been. For Patterson's essay is acute to see what is so special about the Kane-Donaldson edition: not merely that it is based upon a clearly imagined theory of itself, but that the theory has been given full realization. "Counterexample" will not dislodge the "truth" of the Kane-Donaldson edition. Indeed--Patterson himself does not say this, though it is implicit in his argument--even a different "theory" of the *Piers Plowman* materials will necessarily lack critical force against the theoretical achievement represented in the Kane-Donaldson edition. Only another theory of the work that instantiates itself as a comprehensive edition could supplant the authoritative truth of the Kane-Donaldson text.

Why this requirement should be the case is one part of my subject in this essay. The other part, which is related, concerns procedures of theoretical undertaking as such. In this last respect my focus will be on electronic textuality.

Let me address the first issue, then: the theoretical status of what William Carlos Williams called "embodied knowledge" (which may be rendered in the Goethean proverb "In the beginning was the deed"). There is an important sense in which we should see the Kane-Donaldson project as a gage laid down, a challenge to scholars to imagine what they know or think they know. The edition begs to be differed with, but only at the highest level--only at an equivalent theoretical level, in another edition. In this respect it differs from other editions that have seen themselves as theoretical pursuits. Here I would instance Fredson Bowers' *The Dramatic*

*Works of Thomas Dekker* or almost any of the editions of American authors that were engaged under the aegis of the Greg-Bowers theory of editing. These works do not go looking for trouble, as the Kane-Donaldson project did (so successfully). They imagine themselves quite differently, as is readily apparent from the scholarly term they aspired to merit: definitive. In this line of work the scholar proceeding with rigor and comprehensiveness may imagine a *de facto* achievement of critical completeness. Not that other editions might not be executed, for different reasons and purposes. But the "theoretical structure" of the so-called critical edition, in this line of thought, implicitly (and sometimes explicitly) argues that such undertakings would be carried out within the horizon of the definitive critical edition.

In the past 15 years or so scholars have all but abandoned the theory of the "definitive edition", although the term still appears from time to time. The Kane-Donaldson theoretical view, that a critical edition is an hypothesis "designed to account for a body of phenomena in the light of" our given historical knowledge, must be judged to have gained considerable authority during this period. As Patterson's essay suggests, theirs is fundamentally a dialectical and dynamic theory of critical editing. Not of course that a Greg-Bowers approach need fail to appreciate the indeterminacy of particular editing tasks and problems. On the contrary. But the general theoretical approach is different. Bowers, for example, inclines to technical rather than rational solutions to problematic issues, as his famous insistence on collating multiple copies of a printed work clearly demonstrates. This is a procedure that flows from a disciplined theoretical position. But it differs from the theoretical posture adopted by Kane and Donaldson, who take a much more skeptical view of the authority of positive data.

Over against these two theoretical approaches to editing stands that great tradition of what Randy McLeod would call (I think) "un-editing": that is, the scholarly reproduction of text in documentary forms that reproduce more or less adequate replicas of the originary materials. Until recently this approach has scarcely been seen as "theoretical" at all. But McLeod and others have been able to show the great advantages to be gained by theoretically sophisticated forms of documentary procedures. Many doors of perception have been cleansed by R. W. Franklin's *The Manuscript Books of Emily Dickinson* (1981), by Michael Warren's *The Parallel King Lear* (1989), and by the astonishing genetic texts that have come to us from Europe, like D. E. Sattler's *Friedrich Holderlin. Samtliche Werke* (1984).

Let us remind ourselves about what is at stake in these kinds of work. In another day--say, in the late 19th century--an edition like Warren's would have emerged from the influence of institutions such as the Early English Text Society. To that extent it would be seen as an archival work meant primarily to preserve and make accessible certain rare documents. But of course Warren's edition is very different, it is an investigation into the character and status of documents and their relationships (intra- as well as extra-textual). Like Sattler's great edition, it instantiates a self-conscious and theoretical argument. Moreover, Warren's immediate subject, *King Lear*, is implicitly offered as a strong argument for rethinking the textuality of the Shakespeare corpus as a whole. The play isn't seen precisely as representative because the case--which is to say, the documentary material--is too idiosyncratic. This unusual documentary survival, however, is used to encourage and license new acts of attention toward the whole of the Shakespeare canon, as well as to analogous texts beyond.

More speculative theoretical undertakings operate very differently from works like Warren's and Sattler's. Having emerged from the genre of the scholarly essay and monograph, speculative theory tends to move an argument through processes of (as it were) natural selection. Paul De Man was a careful builder of the absences he presented, seiving his materials with great discrimination. In textual and editorial works, by contrast, the whole of each phyla as they have ever been known--every individual instance of all the known lines--lays claim to preservation and display. Everyone comes to judgment: strong and weak, hale and halt, the ideal and the monstrous.

They come, moreover, in *propria persona*, and to that extent they come on their own terms. Franklin's edition of Dickinson points up the theoretical advantage that flows from this method of proceeding. His fidelity to the original manuscripts was so resolute that the documents would eventually be called to witness against him--or rather, against certain of Franklin's less significant ideas about Dickinson's texts. Franklin's work exploded our understanding of Dickinson's use of the physical page as an expressive vehicle. We now see very clearly that she often designed her textual works in the manner of a visual or graphic artist.

These unmistakable cases have come to function something like the case of *King Lear*--strange survivals helping to elucidate surfaces that might otherwise seem commonplace and unremarkable. Franklin himself has resisted and even deplored many of the critical moves that his own work made possible. His edition did not set out to demonstrate some of its most important ideas: that Dickinson used her manuscript venue as a device for rethinking the status of the poetic line in relation to conventions of print display, for example; or that the execution of the (private) fascicles and the (public) letters together comprise a "theory" of verse

freedom every bit as innovative as Whitman's; or that fragmentary scripts might possess an integrity that develops through a dynamic engagement between a text and its vehicular (material) form. These ways of thinking about texts are real if unintended consequences of Franklin's work. The edition itself, however, was clearly undertaken through a different set of ideas. Most apparent, it was a kind of preliminary move toward producing a new print edition of the poems, this time organized by fascicle rather than by hypothetical chronology or topical areas (the two previously dominant ordering systems). Franklin is now completing that print edition. And while it may have considerable success--Dickinson is one of our central American myths--it is unlikely to match the theoretical achievement of *The Manuscript Books*. That achievement is being pursued elsewhere, in the textual works being organized through the so-called *Dickinson Archive*, a scholarly venue sponsoring a variety of approaches to editing Dickinson materials.

Projects of this kind have a strong documentary orientation. They are also electronic. Why? Because digital technology has created a field of significant new possibilities for facsimile and documentary editing projects. In this respect, remarkable genetic editions like Sattler's, although they come to us in codex form, prophecy an electronic existence. They are our age's incunabula, books in winding sheets rather than swaddling clothes. At once very beautiful and very ugly, fascinating and tedious, these books drive the resources of the codex form to its limits and beyond. Think of the Cornell *Wordsworth*, a splendid example of a postmodern incunable. Grotesque systems of diexis and abbreviation are developed in order to facilitate negotiation through labyrinthine textual scenes. To say that such editions are difficult to use is to speak in vast understatement. But their intellectual intensity is so apparent and so great that they bring new levels of attention to their scholarly objects. Deliberate randomness attends every feature of these works, which are as well read as postmodern imaginative constructions as scholarly tools. This result comes about because their engineering of scholarship is often as obdurate and non-transparent as the material being analyzed.

The works I have been talking about suggest how editions may constitute a theoretical presentation. Their totalized factive commitments gives them a privilege unavailable to the speculative or interpretive essay or monograph. Because the three documentary editions just discussed--Franklin's, Warren's, and Sattler's--call special attention to the theoretical status of textual materials, including their own mediating qualities and procedures, these works represent a vanguard for new levels of critical reflexivity.

Their greatest significance, however, only appeared after they were drawn into the orbit of another more encompassing textual innovation: electronic hypertextuality. It gradually became clear that had each of these editions been conceived and executed in digital forms, their documentary and critical imperatives would have discovered a more adequate vehicle.

At that point I began to imagine something that scholars have not, I think, known before. Lee Patterson might have called it a "true" theory of documentary editing, Randall McCloud an (un)serious effort at Unediting. Or call it a hypermedia archive with a relational and object-oriented database. Its truth as theory is two-fold: as a fully searchable set of hyperrelated archival materials; as a reflexive system capable of self-study at various scales of attention. In 1993, The Rossetti Archive was begun as an effort to realize this double theoretical goal. In describing it then I said that its aim was to integrate for the first time the procedures of documentary and critical editing.

But this initial purpose was governed by received understandings of these two approaches. Formed through a long history of scholarship grounded and organized in codex forms, these understandings would have their imaginative limits searched and exposed in the practical work of designing and executing The Rossetti Archive. This result was inevitable. Although The Rossetti Archive was not conceived as a tool for studying the theoretical structure of paper-based textual forms, it has proven very useful in that respect. Translating paper-based texts into electronic forms entirely alters one's view of the original materials. So in the first two years of the Archive's development I was forced to study a fundamental limit of the scholarly edition in codex form that I had not been aware of. Using books to study books constrains the analysis to the same conceptual level as the materials to be studied. Electronic tools raise the level of critical abstraction in the same way that a mathematical approach to the study of natural phenomena shifts the theoretical view to a higher (or at any rate to a different) level.

## II

That last idea, which still seems an important one, led me to write "The Rationale of HyperText" about four years ago. The further development of The Rossetti Archive since that time has brought new alterations to the work's original conception and purposes. Or perhaps they are not so much alterations as supplements. For the project "to integrate the procedures of documentary and critical editing" keeps turning to worlds unrealized. The Rossetti Archive seemed to me, and still seems to me, a tool for imagining what we don't know. The event of its construction, for example, gradually exposed the consequences of a crucial fact I did not at first

adequately understand: that the tool had included itself in its own first imagining. We began our work of building the Archive under an illusion or misconception--in any case, a transparent contradiction: that we could know what was involved in trying to imagine what we didn't know. Four years of work brought a series of chastening interdictions, stops, revisions, compromises.

In the end, despite these events, I see more clearly how one can indeed imagine what you don't know. You can build The Rossetti Archive, which is just such an imagining, and it can be fashioned to reveal its various (and reciprocal) processes of knowing and unknowing. These can be either intramural or extramural. The Rossetti Archive's self-exposures, for example, emerge from two of the project's basic understandings (they are reciprocals): first, the Archive can build the history of its own construction into itself (the ongoing histories of its productions and its receptions); second, it can expose those historicalities to each other at various scales and levels.

These self-exposures might be of many kinds. To give you a good general sense of what we have done let me focus on one aspect of the Archive's work: the decision to make digital images the center of the project's computational and hypertextual goals. This decision followed our aspiration to marry critical and facsimile editing, and our belief that electronic textuality had arrived at a point where the desire could be realized. So in 1992 we began trying to bring about the convergence of these ancient twin.

Let me recapitulate a bit of the recent history of humanities computing initiatives. Current work in electronic text and data management fall into two broad categories that correspond to a pair of imaginative protocols. On one hand we have hypertext and hypermedia projects--information databases organized for browsing via a network of complex linkages. These characteristically deploy a mix of textual and image materials that can be accessed and traversed by means of a presentational markup language like HTML. On the other hand are databases of textual materials organized not so much for browsing and linking/navigational moves as for in-depth and structured search and analysis of the data. These projects, by contrast, require more rigorous markup in full SGML. If they deploy digital images, the images are not incorporated into the analytic structure. They will be simple illustrations, to be accessed--perhaps even browsed in a hypertext--for reference purposes.

One kind of project is presentational, designed for the mind's eye (or the eyes' mind); the other is analytic, a logical structure that can free the conceptual imagination of its inevitable codex-based limits. The former tend to be image-oriented, the latter incline to be text-based.

To date, almost all of the most impressive electronic text projects have been text-based: Robinson's Chaucer project, McCarty's *Ovid*, Duggan's *Piers Plowman*, the University of Bergen's Wittgenstein project. Set beside these kinds of works, even the most complex hypertext--*The Perseus Project*, for instance, or the literary "webs" developed by George Landow at Brown University--will seem relatively primitive scholarly instruments. (The great exception to this generalization would be those non-proprietary and wholly decentered projects that grow and proliferate randomly, like the network of gopher servers or--the spectacular instance--the World-Wide Web's [WWW] network of hypermedia servers.)

So far as localized hypermedia projects are concerned, works like those developed in Storyspace have a distinct attraction, as the amazing success of WWW demonstrates. Crucially, they operate at the macro-level of our human interface. Elaborated front-end arrangements reinforce an important message: that however strange or vast the materials may sometimes appear to be, the user can maintain reasonable control of what happens and what moves are made. The buzzwords that iconify such a message are "user-friendly" and, most widespread of all, "interactive".

Not without reason do hypertext theorists regularly imagine their world in terms of spatial and mapping metaphors. Not without reason did the greatest current hypertext project (WWW) decide to code its data in HTML (it could have supported a more rigorous DTD for its materials), or make the accessing of images (rather than the analysis of their information) a key feature of its work. WWW's success derives from its humane--indeed, its humanistic--interface. Of course WWW, like all hypermedia engines, is grotesquely pinned down by the limits of the color-monitor. Still, though limited by the monitor (whether in two or three dimensions), hypertexts like WWW can simulate fairly well the eye-organized environment we are so used to.

By contrast, SGML-type projects need take little notice of the eye's authority. They are splendid conceptual machines, as we see when we reflect on the relative unimportance of sophisticated monitor equipment to text-based SGML projects. The appearance of text and data is less crucial than their logical organization and functional flexibility.

The computerized imagination is riven by this elementary split, as everyone knows. It replicates the gulf separating a Unix from a

Mac world. It also represents the division upon which The Rossetti Archive was consciously built. That is to say, from the outset I held the project responsible to the demands of hypermedia networks, on one hand, and to text-oriented logical structures on the other. This double allegiance is fraught with difficulties and even with contradictions, as would be regularly shown during the first period of the Archive's development (1992-1995). Nevertheless, I determined to preserve both commitments because each addressed a textual ideal that seemed basic and impossible to forego. We knew that we did not have the practical means for reconciling the two demands--perhaps they can never be reconciled--but even products like Dynatext, imperfect as they were (and are), held out a promise of greater adequacy that spurred us forward. Besides, the tension fostered and exacerbated by this double allegiance might prove a kind of *felix culpa* for the project, a helpful necessity to mother greater invention. This was my initial belief, and events have only strengthened that faith.

So our idea was to build the Archive along a kind of double helix. On one hand we would develop a markup of the text data in SGML for a structured search and analysis of the Archive's materials. On the other we would design a hypertext environment for the presentation of the primary documents--Rossetti's books, manuscripts, proofs, paintings, drawings, and other designs--in their facsimile (i.e., digital) forms. A key problem from the outset, then, was how to integrate these different organizational forms. We arrived at two schemes for achieving what we wanted. One involved a piece of original software we would develop, now called [Inote](#). Although I won't be discussing that part of the project today in great detail, I shall come back to it briefly later in connection with another matter. The other plan was to develop an SGML markup design that would extend well beyond the conceptual framework of TEI, the widely-accepted text markup scheme that had spun off from SGML. TEI has become the standard protocol for organizing the markup of electronic textual projects in humanities.

This is not the place to enter a critique of the limitations of a TEI approach to text-encoding. Suffice it to say that the linguistic orientation of TEI did not suit our documentary demands. Bibliographical codes and the graphic design of texts are not easily addressed by TEI markup. But those features of texts are among our primary concerns; after all, we had chosen Rossetti as our model exactly because his work forced us to design an approach to text markup that took into account the visibilities of his expressive media. What we wanted was a text-markup scheme that could deal with the whole of the textual field, not simply its linguistic elements. So in 1992 we began the effort to design an SGML-based documentary markup for structured search and analysis of all the work of Dante Gabriel Rossetti.

Those initial design sessions immediately exposed an unanticipated problem: the presence of textual concurrencies. SGML markup of text organizes its fields as a series of discrete textual units. Each unit can comprise imbedded subseries of the same logical form, and further subseries can be developed indefinitely. But SGML processors have no aptitude for is markup of textual features that are concurrent but logically distinct. A classic instance would be trying to permit a simultaneous markup of a book of poems by page unit and by poem. In SGML you are led to choose one or the other as the logical basis of the markup design.

At that point we had two options: to abandon SGML and look for a markup language that could process concurrent structures; or to try to modify SGML to accommodate the needs of The Rossetti Archive. In choosing the latter option, as we did, we were consciously committing ourselves to an inevitable set of unforeseeable problems. For the truth is that all textualizations--but pre-eminently imaginative textualities--are organized through concurrent structures. Texts have bibliographical and linguistic structures, and those are riven by other concurrencies: rhetorical structures, grammatical, metrical, sonic, referential. The more complex the structure the more concurrencies are set in play.

We made our choice for SGML largely because we could find no system for dealing with concurrencies that possesses the analytic depth or rigor of SGML, and because the project was not to design a new markup language for imaginative discourse. True, building a general model for computerized scholarly editing depends on an adequate logical conception of the primary materials, and it does not bode well to begin with a logic one knows to be inadequate. On the other hand, what were the choices? If natural languages defeat the prospect of complete logical description, an artistic deployment of language is even more intractable. In such cases adequacy is out of the question. Besides, SGML is a standard system. We are aware of its limitations because the system is broadly used and discussed. As Hamlet suggested, we seemed better off bearing the ills we had than flying to others we knew nothing of. And there was one other important consideration: the basic concurrency of physical unit v. conceptual unit might be addressed and perhaps even accommodated through other parts of the design structure of the Archive--through the markup of images, through software for analyzing image information, and through the hypermedia design.

So in 1992 we began building The Rossetti Archive with what we knew were less than perfect tools and under clearly volatile conditions. Our plan was to use the construction process as a mechanism for imagining what we didn't know about the project. In

one respect we were engaged in a classic form of model-building whereby a theoretical structure is designed, built, and tested, then scaled up in size and tested at each succeeding juncture. The testing exposes the design flaws that lead to modifications of the original design. That process of development can be illustrated by looking at one of our SGML markup protocols--[the DTD for marking up Rossetti Archive Documents \(or RAD\)](#). This DTD is used for all textual (as opposed to pictorial) documents of Rossetti's work, as well as for important related primary materials (like the Pre-Raphaelite periodical *The Germ*). It defines the terms within which structured searches and analyses of the documents will be carried out. My immediate interest is not in the SGML design as such but in the record of modifications to the design. That record appears as the list of dated entries at the top of the document.

Before discussing some of these entries let me point out two matters of importance. First, note that the date of the first entry is "6 Oct 94". That date is just about one year after we completed the first design iterations for the Rossetti Archive DTDs. A great many modifications to the initial design were made during that year, but we did not at first think to keep a systematic record of the changes. So there is a pre-history of changes held now only in volatile memory: i.e., the personal recollections of the individuals involved, and in paper files that contain incomplete records of what happened in that period.

Second, the record does not indicate certain decisive moments when the Archive was discovering features of itself it was unaware of. In these cases no actual changes were made to the DTDs. For example, we regularly discovered that different persons implementing the markup schemes were liable to interpret the intent of the system in different ways. We tried to obviate this by supplying clear definitions for all the terms in use, as well as a handbook and guide for markup procedures. But it turned out--surprise, surprise--that these tools were themselves sometimes ambiguous. The Archive is regularly reshaped, usually in minor ways, when we discover such indeterminacies.

External factors have also had a significant impact on the the form and content of the Archive, and we found ourselves driven into unimagined directions. One of the most interesting shifts came about because of our problems with permissions and copyrights. The cost of these exploded as the Archive was being developed, and in certain cases we were simply refused access to materials. This problem grew so acute--the date is 1994--that I decided on a completely new approach to the issue of facsimile reproduction of pictures and paintings. Rather than construct the first installment of the Archive around digital facsimiles made from fresh full-colour images (slides, transparencies, photographs), I determined to exploit a vast contemporary resource: the photographs made of Rossetti's works during and shortly after his lifetime, many done by friends and other early pioneers in photography. Rossetti is one of the first modern artists to take a serious interest in photography--the photographs he made of Jane Morris and Fanny Cornforth with J. R. Parsons are themselves masterpieces of the art.

This shift to early photographic resources--the materials date from the mid-1860s to about 1920--has two great advantages, one both scholarly and practical, the other scholarly. The move allows us to temporize on the extremely vexed issue of copyright. We use whatever fresh full-colour digital images we can afford and work toward developing standards for the scholarly use of all such materials. These procedural advantages bring a number of significant scholarly gains as well. On one hand we now comprehensively represent Rossetti's visual work in the medium that was probably its major early disseminating vehicle. On another, we create a digital archive of great general significance for studying both the history of photography and the history of painting.

Whether extra-mural or intra-mural, however, these changes to the Rossetti Archive are, first, the realized imaginings of what we didn't know; and second, clear instances of a theoretical power beyond the range of strictly speculative activities. Let's look again for a moment at some intramural examples coded in the historical log of the RAD DTD. The recorded alterations in that DTD design were made as we scaled up the project from its initial development model (which involved only a small subset of Rossetti documents). This is a record of a process of imagining what we didn't know. The imagining comes through a series of performative moves that create a double imaginative result: the discovery of a design inadequacy, and a clarification of what we had wanted but were at first unable to conceive.

Some of the modifications are relatively trivial--for example, this one:

```
<!-- div1 ornLb added to titlePage 11-20-96 A.S.-->
```

The change permits the markup of an ornamental line break on title pages. Small as it is, the change reflects one of the most important general demands laid down by our initial conceptions: to treat all the physical aspects of the documents as expressive features.

A more obviously significant change is the following:

<!-- revised: 9 Mar 95 to add r attr to l, lg and lv (seg)-->

This calls for the introduction of the attribute "r" (standing for "reference line") to all line, line group, and variant line values in the Archive. The small change defines the moment when we were able to work out a line referencing system for the Archive that permits automatic identification of equivalent units of text in different documents. We of course knew we wanted such a system from the outset, but we were unable to feel confident about how the system should be organized until we had three years of experience with many different types of textual material.

Working out this scheme for collating Rossetti's texts revealed an interesting general fact about electronic collating tools: that we do not yet have any good program for collating units of prose texts. The poetic line is a useful reference unit. In prose, the textual situation is far more fluid and does not lend itself to convenient division into discrete units. The problem is especially apparent when you try to mark up working manuscripts for collation with printed texts. The person who discovers a reasonably simple solution to this problem will have made a signal contribution not just to electronic scholarship, but to the theoretical understanding of prose textuality in general.

But let us return to the history of RAD revisions. Look at the notation for 14 June 1995:

<!-- revised: 14 Jun 95 to add group option to rad for serials -->

A large-scale change in our conception of the Archive's documentary structure is concealed in this small entry. The line calls for the use of the "group" tag in the markup structure for the serials to be included in the Archive (like *The Germ*). Behind that call, however, lies a difficult process that extended over several years. The problem involved documents with multiple kinds of materials (like periodicals). The most problematic of these were not the periodicals, however, but a series of primary Rossetian documents--most importantly, composite manuscripts and composite sets of proofs. In these materials the problems of concurrency became so extreme that we began to consider the possibility of abandoning SGML altogether--which would have meant beginning the whole project from scratch.

As it turned out, we found a way to manipulate the SGML structure so as to permit a reasonably full presentation of the structure of these complex documents. That practical result, however, was not nearly so interesting as the insights we gained into general problems of concurrency and into the limitations of SGML software.

Consider the following situation. Rossetti typically wrote his verse and prose in notebooks of a distinctive kind. Two of these survive intact to this day, but the fragments of many others are scattered everywhere. Many are loose sheets or groups of sheets, many others come down to us as part of second-order confederations of material that Rossetti put together, or that were put together by others (during his lifetime or after his death) as other second or even third-order arrangements. Problem: devise a markup scheme that will reconstruct on-the-fly the initial, but later deconstructed, orderings. Or--since in many cases we can't identify for certain which pages go with which notebook phylum--devise a markup scheme that constructs on-the-fly the various possibilities. Or: devise a system that lays out an analytic history of the re-orderings, including a description of the possible or likely lines by which the distributed documents arrived at their current archival states.

An instrument that could perform any or all of these operations would have wide applicability for textual scholars of all kinds and periods. I am sure it could be developed, perhaps even within SGML. It is an instrument that was imagined into thought by building the Rossetti Archive. We saw it as we were trying to devise markup systems that would accommodate the composite proofs and manuscripts that are so characteristic of Rossetti's extant textual materials. It is an instrument that we would like to develop ourselves--except we're far too busy with so many other basic problems and demands.

### III

These examples illustrate what I would call the pragmatics of theory, and the sharp difference between theory, on one hand, and hypothesis or speculation on the other. In humanities discourse this distinction is rarely maintained, and the term "theory" is characteristically applied to speculative projects--conceptual undertakings (*gnosis*) rather than specific constructions (*poesis*).

Scientists work within the former distinction, and in this sense The Rossetti Archive seems a "scientific" project. Patterson's discussion of the Kane-Donaldson edition of *Piers Plowman* implicitly affirms the same kind of distinction, where "theory" operates through concrete acts of imagining.

The virtue of this last kind of theorizing is that it makes possible the imagination of what you don't know. Theory in the other sense--for instance, Heideggerian dialectic--is a procedure for revealing what you do know, but are unaware of. Both are intellectual imperatives, but in humanities disciplines the appreciation for theory-as-poeisis has grown attenuated. The need to accommodate electronic textualities to humanities disciplines, which are fundamentally document- and text-based, is bringing a radical change in perspective on these matters.

The force of these circumstances has registered on nearly every aspect of The Rossetti Archive, as I've already indicated. To this point, however, I've used examples that illustrate the praxis of theory as it is an example of a methodological process we are very familiar with, though perhaps not so much in a humanities context: the process of imagining what you know, testing it, scaling it up, modifying it, and then re-imagining it; and then the process of repeating that process in an indefinite series of iterations.

At this point I want to give one more example of that process. It is the history of the development of a piece of software I mentioned earlier, Inote (formerly called the Image Tool). More than an exemplum of theory-as-poeisis, the story indicates, I believe, the "strange days" that lie ahead for humanities scholars as we register the authority of these new electronic textualities.

Inote was originally an idea for computing via images rather than with text or the data represented in text. Because information in bit-mapped images cannot be coded for analysis, our technical people were asked if it would be possible to lay an electronic transparency (as it were) over the digital image and then use that overlay as the vehicle for carrying computable marked-up data and hypertext links. The idea was to treat the overlay as a kind of see-through page on which one would write text that elucidated or annotated the imaged material "seen through" the overlay. (The idea originates in scholarly editions that utilize onionskin or other transparent pages to create an editorial palimpsest for complex textual situations.)

As with virtually all work undertaken at IATH, this tool's design was influenced by many people who came to have an interest in it. Consequently, because I was initially most preoccupied with designing The Rossetti Archive's markup structure, my interest in the development of Inote hung fire. My own early thought had been that such a tool might enable The Rossetti Archive to incorporate images into its analytic text structure and thus establish a basis for direct searches across the whole of the Archive at the image level. As I worked more and more closely with SGML markup, however, I began to suspect that the same result might be achieved through the design of a DTD for images. That idea, plus the technical difficulties in building Inote, drew my attention away from the tool's development.

Inote thus began to evolve in ways I (at any rate) had not anticipated. As others looked for features that would answer their interests, it emerged as a device for editing images with multiple-style overlays that if clicked would generate a text file carrying various annotations to the image. These annotations would be saved as part of the total Archive structure and hence could be imbedded with hypertext links to other images or archival documents.

At that point--the date is early 1995--my practical interest in the tool was revived. This happened because my work on the DTDs for the Archive, nearing completion, began to expose certain limitations in the overall design structure. It was growing very clear that the Archive's two parallel universes continued discontinuous in fundamental and (in this case, I thought) unhelpful ways. Inote had become a device with two primary functions:

1. it allowed one to build a random set of image points or areas to which one could attach text materials of varying kinds;
2. it allowed one to imbed hypertext links to those materials.

So while the tool created navigational paths from text to image and vice versa, thus connecting the two basic (and different) kinds of objects in the Archive; and while it drew these image-related texts (and hence the images as well) into the full computational structure, it did not organize these materials within a logical structure readable in the Archive. Any searches of the materials would have to be in effect string searches. (Inote in its first iteration, for example, could not function in close cooperation with the indexable fields of information as established through the Archive's DTDs).

This limitation in the tool recalled my attention to the Archive's basic contradiction and double allegiance. The full evolution of the

markup structure--the building of the DTDs for all text and image documents--had not been matched by a corresponding development in Inote, at least for those who would want--as I did--a tool that could function within the SGML marked database (texts as well as pictures). This discrepancy arises because the first version of Inote, unlike the DTDs, was not mapped to the logical (DTD) structure of the files in the Archive. It would be formally integrated with the SGML marked database only when it could summon its materials within pre-established indexable categories. Furthermore, an adequate integration would require some kind of mappable relation between those indexable forms and the SGML marked database.

To address these problems I suggested that we limit our consideration, at least initially, to textual images--i.e., images of manuscripts, proofs, and printed documents--since these are far simpler than pictorial images. We began by posing the question "what is the formal structure of a text page?"

This initial query rises through a pair of presuppositions implicit in The Rossetti Archive. The first reflects the Archive's practical delivery of its images, which the Archive manipulates as units of either single pages or single page openings (i.e., a pair of facing pages). That procedure flows from a second assumption about texts in general. We assume that a "text" is a rhetorical sequence organized by units of page, with each page centrally structured in terms of a sequence of lines commonly running from top to bottom, left to right and within some set of margins (which may be reduced to nil (practically) on any side).

In marking up the formal structure of the text image, these general conventions defining the shape of the page will govern the markup. Consequently, I proposed the following: that the page be formally conceived as a structure of different spatial areas. I initially proposed four marginal areas (left and right margins plus header and footer) and a central text area stacked into four equal horizontal sections. This design was found to be more complex than necessary, and we eventually settled on a page design of three stacked horizontal areas with no mapping at the margins.

The essential point of this structure is to permit SGML marked textual materials to be mapped directly to digitized images. An indexable code is supplied to digital materials so that a formal relation can be established between the two conceptual states of every text (i.e., texts conceived as linguistic fields and texts conceived as bibliographical fields). SGML marked texts have nothing to say about the physical status of marked materials because the markup is not conceived in terms of spatial relations. Even if a set of SGML fields were to be defined for bibliographical features of text, no formal structure would exist to connect the digital images to the SGML marked texts--because the latter have not been conceptually defined in relation to the former. In the case of textual materials, this formalized representation of the bibliographical field would serve primarily to facilitate the study of documents with "irregular" textual conditions (e.g., documents with many additions, corrections, and erasures; or documents with nonlinguistic elements, such as Blake's illuminated texts). At least that was the initial imagination for the scheme.

Inote has now been developed along these lines and its functions have been applied and adapted by the editors of [The Blake Archive](#). The results can be seen in *The Blake Archive's* recent release of its first installment, an edition and study tool for [The Book of Thel](#).

But not all the results. The practise of the theory of Inote revealed some interesting ideas about computerizing textual materials in relation to a database of images. For instance, it is apparent that in such cases one should define the basic textual unit as the page (as is done in *The Blake Archive*) rather than the work (as is done in SGML and--alas--in The Rossetti Archive). Only if the basic unit is the page (or the page opening) can the lineation in the digital image be logically mapped to the SGML markup structure. Of course if SGML software were able to handle concurrent structures, this consequence would not necessarily follow.

*The Blake Archive's* work conforms to the original thought about Inote that it be shaped to integrate the meta-data in an SGML-marked text to the direct study of the digital images that constitute that meta-data. As the tool was being adapted by *The Blake Archive* editors, however, their work exposed more severely than ever the problem of analyzing the data of digital images. Blake's work lent itself to the idea of Inote because that idea was fundamentally a textual one; and while Blake's works are profoundly iconological, they are also, at bottom, texts, not pictures.

We still do not have means for carrying out on-the-fly analyses of the iconological information in pictures (let alone pictures that are aesthetically organized). Our work with Inote shows how far one might go--and it is pretty far, after all--to integrate an SGML approach to picture markup and analysis. But the limitations of such an approach are also painfully clear.

I have no idea how or when this nexus of problems will be overcome, though I do have some thoughts on experimental avenues that

might be explored. Nevertheless, electronic textuality is so intimately bound to the manipulation of images that the issues must remain at the forefront of attention. Logical markup through schemes derived from linguistic models, powerful though they are, cannot even serve the full needs of textual scholars, much less those of musicologists, art historians, film scholars, and artists in general. The Pentagon and the Infotainment industry have committed large resources to research into these problems, whose importance is clear to them. While scholars and archivists lack those kinds of financial resources, we would be wrong to stand aside while the issues are being engaged and theorized. Indeed, scholars like ourselves typically possess a phenomenological understanding of such materials that is obscure to techno-scientific researchers. So it's extremely important that traditional scholars and critics experiment with the study of digital images, and--perhaps even more useful--set their students to play and experiment with these materials. If there ever was a situation calling us to imagine what we don't know, this is one.

#### IV



figure 1

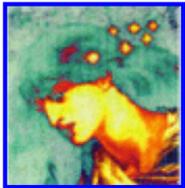


figure 2

An artist friend of mine and I recently stumbled into an imagining of that kind. We were fooling around with a digital image of Rossetti's famous painting *The Blessed Damozel*--filtering the image in a series of playful and random ways. At a certain point we generated an image that startled me. The arbitrary distortion had suddenly clarified a chromatic organization I had never noticed--familiar as this painting was. Here are two images--a facsimile of the damozel's head in the original painting in the Fogg Museum (fig. 1), and a deformed version (fig. 2) we generated with our image editor. The chromatic relationships exposed here were so interesting that I wanted to try some more disciplined experiments of a similar kind when I had the time. In the initial series we did not keep track of the distortion sequences. Besides, the images were of limited use because of the smallness of the cropped picture area.



figure 3

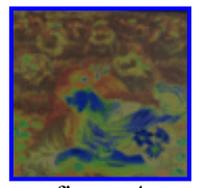


figure 4



figure 5

Here are two recent deformations made on a larger area of *The Blessed Damozel* (original, fig. 3). The images were produced using the standard filter protocols in Adobe Photoshop 4.0, with some adjustments to sharpen the features that were generated (figures 4 and 5). The rhetorical power of Rossetti's art is such that its intellectual beauties, which are always primary concerns for him, get thrown out of relief (as it were) when we encounter them in their natural (so to speak) condition. That is to say, Rossetti works hard to ensure that his arguments will register as primary apparitions rather than as abstract ideas. He works this way programmatically--indeed, much of his importance as an artist depends upon the way he explores the liminal moment between the conventions of pictorial illusionism, on one hand, and modernist abstraction on the other. One may well recall here his famous dictum on the relation of spiritual and material presences, ideas and phenomena:

Thy soul I know not from thy body, nor  
Thee from myself, neither our love from God.

The formal relations between the flesh of the damozel, the stars in her crown, her symbolic flowers, and the glimpsed world behind the heaven of embracing lovers is exposed in these distorted images: in one case through the patterns of blue/green, in the other through their gold equivalents. Those patterns help to explain how Rossetti manages to evoke, in this voluptuous Venetian exercise, the presence of a transnatural or divine order of reality. He has concealed within the painting's decorative opulence a subtle allusion to a commonplace feature of primitive religious pictorialism. We glimpse the allusion in the two previous deformations, but it leaps to our awareness when we pass the original image through a black and white emboss filter (fig. 6). In this deformation we realize that an iconic gold ground is being subtly evoked in the broken, sinuous line of interstitial moments near the emparadised lovers, and that the line explicitly relates to her crown of stars and to the central triangle of her face, her hands, and the lilies she is holding.



figure 6

Looking again at the first two distorted images reveals another key feature of the painting: the composition by rhyming circular forms, with the damozel as the central form (circling clockwise) and the embracing lovers as repetitions (circling counterclockwise). This pattern is exposed in another distortion, we can also see that six of the eleven pairs of lovers are arranged in an ellipse echoing the damozel's crown of stars.



figure 7



figure 8

These images call attention to important elements in Rossetti's picture. Here are two other interesting deformations I show without comment (figures 7 and 8). Such images recall those developed by art historians to demonstrate pictorial formalities--e.g., when full-colour paintings are analytically recast in black and white with edges or other forms highlighted in bold. What is important--and new--about our electronic deformations, however, is their arbitrary character. In Blakean terms, they come to us like the "dictations" of "Eternity". We register the formal exposures they execute because we possess an a priori understanding of formal relationships. Nevertheless, because aesthetic forms like *The Blessed Damozel* always pursue orders of appearance, their intellectual beauties are wrapped in veils, as Shelley acutely observed. Art works lead us to respond equally to the veils and to their informing ideas. The deformed images suggest that computerized art editing programs can be used to raise our perceptual grasp of aesthetic objects.

There are critical opportunities to be exploited in the random use of these kinds of deformation. For instance: take a given set of paintings and run them through a series of deformations--let us say, a set of Rossetti's paintings, or a set of works chosen from a known corpus of Pre-Raphaelite, Aesthetic, and Symbolist art. Every time an altered image is generated that is judged interesting or revealing in some way, save the settings by which the image was produced. Then use those settings randomly on all the pictures in the chosen set and compare the results.

The images I've just shown are the first of a group whose settings we have saved and that we mean to try in an experiment of this kind. I have no idea what its outcome will be. Perhaps it will give us some new views of the intellectual organization or stylistic conventions of Symbolist art, perhaps not. Perhaps it will suggest other critical moves that might be facilitated by computerization. Artists are of course already using these tools in analogous ways. Critics and scholars are perhaps less imaginative beings, but even we needn't commit ourselves only to imagining what we already know.



figure 9

Let me close by showing two final distortions of Rossetti's famous damozel. These come from the set I and my friend initially made when we were simply playing around with his image editor (figures 9 and 10). Most people take immediate delight in both these images and talk volubly about what they mean. These meanings range from suggestive redescriptions to abstract discussions of critical methodology. Some doubt that the images have any interpretive relation to Rossetti's original picture, but many see the matter very differently. My own response, at first, was that these kinds of distortions, while charming, lacked the critical function I saw in the images developed with the Adobe filters. In the latter, one could readily appreciate how the filtered forms clarified important aspects of Rossetti's famous painting. They typically revealed not what we didn't know, but what we didn't know we knew. I think it's also true that the filterings

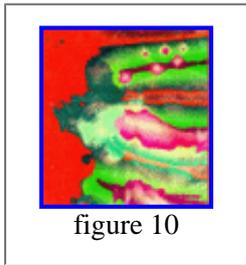


figure 10

delivered some things we didn't know at all, or at any rate that I didn't know. I didn't know, for example, that the spaces separating the background of kissing lovers had been chromatically organized so that they would relate simultaneously to two different, and in other respects two contradictory, forms of pictorial order. This startling fact about the picture only emerged when two of the distorted images were set beside each other. But that revelation was unusual, and it seemed clear to me that the deformations largely functioned in a pedagogical way. Insofar as these images brought an imagination of the unknown, they were pointing to the image editor as a critical and interpretive tool. One is impressed to realize that random manipulations of an indeterminate set of filter settings can perform useful critical operations.

Conversations with students, however, continually returned to the last two deformed images, and I have come to think very differently about them. The critical force of the Photoshop deformations develops from their ability to expose matters that will be generally recognized, once they are seen. We might differ on how we evaluate the formalities exposed through the Adobe filterings, but not about the structures and forms as such. What arrests us in the final two images is their idiosyncrasy. These images generate what we would elsewhere and otherwise call "subjective" interpretations. They came to being through a series of obscure intuitive operations, a series of transformations and transformations of transformations. The series ended when the image proved satisfying.

I don't as yet have a vocabulary for talking about this process of developing unknown images in relation to familiar aesthetic images. The process invites us to reconsider the critical authority of so-called subjective aesthetic engagement. These strange images evoke our interest exactly because they don't pretend to supply us with a generic response to the picture, something every viewer could agree with (which is what we are offered in iconographical and formal redescriptions). The distortions arrest our attention, moreover, only because we already know the original, which comes back to us through them as if from an unimaginable world or point of view. As a form of interpretive action, these images recall Galvano della Volpe's idea that interpretation develops a "quid" against which the unique qualities of an art work can be measured. The "quid" is critically useful not insofar as it mirrors its object but insofar as it fails to do so--or perhaps rather, insofar as it dramatizes the fact that an art work abides no equivalences even as it seems poised to generate the shapes of equivalence forever. In this respect the distortions suggest the usefulness of thinking about art--at any rate, certain art works--as if they were informed by an idea, or an inertia, that has not been exhausted in the executed fact of the work we think we have and we think we know. This is an ancient way of engaging art that was revived in symbolist and surrealist practise. Not suprisingly, it is a view that Rossetti shared.

## Bibliography

### Printed Materials

- Dekker, Thomas. *The Dramatic Works of Thomas Dekker*. Ed. Fredson Bowers. Cambridge: Cambridge University Press, 1953.
- Dickinson, Emily. *The Manuscript Books of Emily Dickinson*. Ed. R. W. Franklin. Cambridge, MA: Belknap Press of Harvard University Press, 1981.
- Della Volpe, Galvano. *Critique of Taste*. Trans. Michael Caesar. London: New Left Books, 1978.
- Holderlin, Friedrich. *Samtliche Werke: Frankfurter Ausgabe [historische-kritische Ausgabe]*. Ed. D. E. Sattler. Frankfurt am Main: Verlag Roter Stern, 1975-.
- Langland, William. *Piers Plowman: The B Version*. Ed. George Kane and E. Talbot Donaldson. London: University of London, Athlone Press, 1975.
- Patterson, Lee. "The Logic of Textual Criticism and the Way of Genius: The Kane-Donaldson *Piers Plowman* in Historical Perspective." *Textual Criticism and Literary Interpretation*. Ed. Jerome J. McGann. Chicago: University of Chicago Press, 1985. 55-91.
- Samuels, Lisa. "Introduction". *Modern Language Studies* 27.2 (Spring 1977). 3-6 (Special Issue of the journal edited by Samuels and devoted to the subject Poetry and the Problem of Beauty).
- Shakespeare, William. *The Parallel King Lear, 1608-1623*. Ed. Michael Warren. Berkeley: University of California Press, 1989.

## Online Materials

- Kirschenbaum, Matthew. "Lucid Mapping and Codex Transmissions in the Z-Buffer."  
URL: <http://web.archive.org/web/20041123091955/http://www.iath.virginia.edu/~mgk3k/lucid/>
- McGann, Jerome. "The Rationale of HyperText."  
URL: <http://web.archive.org/web/20041123091955/http://jefferson.village.virginia.edu/public/jjm2f/rationale.html>
- McGann, Jerome. *The Rossetti Archive*  
URL: <http://web.archive.org/web/20041123091955/http://www.press.umich.edu/bookhome/rossetti/index.html>
- *The William Blake Archive*. Ed. Morris Eaves, Robert Essick and Joseph Viscomi.  
URL: <http://web.archive.org/web/20041123091955/http://www.iath.virginia.edu/blake/>
- *The Canterbury Tales Project*. Ed. Norman Blake and Peter Robinson.  
URL: <http://web.archive.org/web/20041123091955/http://www.cup.org/Chaucer/ctptop.html>
- *The Dickinson Archive*. Ed. Martha Nell Smith, Ellen Louise Hart and Marta Werner.  
URL: <http://web.archive.org/web/20041123091955/http://www.iath.virginia.edu/dickinson/>
- Inote. Software from the Institute for Advanced Technology in the Humanities.  
URL: <http://web.archive.org/web/20041123091955/http://www.iath.virginia.edu/inote/>
- "Language Visualization and Multilayer Text Analysis." Visualization Group, Theory Center, Cornell University.  
URL: <http://web.archive.org/web/20041123091955/http://www.tc.cornell.edu/Visualization/contrib/cs490-95to96/tonyg/Language.Viz1.html>
- *The Piers Plowman Electronic Archive*. Ed. Hoyt N. Duggan.  
URL: <http://web.archive.org/web/20041123091955/http://www.iath.virginia.edu/piers/>
- *Who's Who in the Metamorphoses of Ovid: The Analytical Onomasticon Project*. Ed. Willard McCarty, Burton Wright and Aara Suksi.  
URL: <http://web.archive.org/web/20041123091955/http://www.princeton.edu/~mccarty/Onomasticon/>
- *The Wittgenstein Archives at the University of Bergen*. Dir. Franz Hespe.  
URL: <http://web.archive.org/web/20041123091955/http://www.hd.uib.no/wab/>

▶ ABOUT IATH

▶ FELLOWSHIPS

▶ OUR RESEARCH

Copyright © 1997 by [Jerome McGann](#), all rights reserved

Document URL: <http://jefferson.village.virginia.edu/~jjm2f/chum.html>

Last Modified: Monday, 15-Dec-1997 11:38:57 EST