

7-2014

# Teaching Attentive Reading and Motivated Writing through Digital Editing

Amanda A Gailey

*University of Nebraska - Lincoln*, [gailey@unl.edu](mailto:gailey@unl.edu)

Follow this and additional works at: <http://digitalcommons.unl.edu/englishfacpubs>



Part of the [Comparative Literature Commons](#), [Digital Humanities Commons](#), [English Language and Literature Commons](#), [Modern Literature Commons](#), and the [Reading and Language Commons](#)

---

Gailey, Amanda A, "Teaching Attentive Reading and Motivated Writing through Digital Editing" (2014). *Faculty Publications -- Department of English*. 172.

<http://digitalcommons.unl.edu/englishfacpubs/172>

This Article is brought to you for free and open access by the English, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications -- Department of English by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

AMANDA GAILEY

## Teaching Attentive Reading and Motivated Writing through Digital Editing

Though English departments, including my own at the University of Nebraska, have been teaching digital humanities (DH) courses for over a decade, hyperbolic claims about the perils and promises of using computers in the study of literature continue to appear in the press. A piece in the *Los Angeles Review of Books* likens the algorithms used by some digital humanities methods to fascism (Marche). Another, in *The Huffington Post*, compares the rise of digital humanities to “our uncritical acceptance of drone attacks” (Mohamed). On the other hand, digital humanists such as Franco Moretti, who famously promote “distant reading” as opposed to close reading, project a future for the humanities that radically departs from long-cherished methods. The controversial Digital Humanities Manifesto 2.0, put out by a group of scholars at UCLA, includes a lot of talk about DH’s “utopian core” and optimistic “democratization of culture and scholarship” (Presner). In my experience teaching dozens of courses and workshops on digital textual representation, the pedagogical value of digital tools—specifically TEI (Text Encoding Initiative)—is more complex and ultimately more rewarding than these caricatures imply.

Text encoding, at least the in-depth, student-conceived markup that I teach in my classes, is not free of complications. It is labor-intensive, time-consuming, and sometimes extremely frustrating to beginners. It requires reliable computers, preferably ones the students can take home, and a classroom license for software (I use Oxygen XML Editor) that costs several hundred dollars. In order to create projects that can be publicly displayed, students require access to a server with certain technological specifications. Also, TEI is far from an uncontroversial way of approaching texts even in the digital editing community—it is predicated on a theory of textuality that is open to an array of criticisms, and anyone who has worked with it for a long time will have a lengthy list of quibbles regarding various features. (Frankly, I’d like to see TEI become one of several widely accepted and well-documented ways we can approach digital editing.)

Notwithstanding its practical and theoretical hurdles, TEI is an invaluable tool for teaching literature. It makes a few pedagogical goals central to the work of the class: students must pay careful, consistent attention to the text; they learn to understand the cultural record as malleable; they feel a clear sense of purpose, audience, and expertise when writing; they leave with transferable technical skills. I’d like to offer instructors curious about

digital humanities some considerations for how digital text editing can augment important teaching goals.

## Background

“Digital humanities” has become such a broad term, a buzzword encompassing everything from stylometry to the future of scholarly publications and conversations about academic labor, that it is not a very clear indicator of what any class in digital humanities might entail. I find it more useful to think of my classes as courses on digital textual editing—that is, the continuation into the digital age of the disciplines concerned with the materiality and representation of texts, such as textual studies and editorial theory. Instructors of different kinds of digital humanities classes may have quite different goals and methods than I do and use different technologies to accomplish them.

TEI provides the technological touchstone for my classes, and a brief overview of it for the uninitiated may be helpful here. TEI is a vocabulary used in XML (Extensible Markup Language) encoding. Like its more familiar relative HTML (HyperText Markup Language), XML uses tags—descriptors in angle brackets—to label portions of text or points in texts. For example, a paragraph might be tagged in this way:

```
<p> Here he again fell faintly back. Again his mind wandered:  
but he rallied, and less obscurely proceeded.</p>
```

The first tag, `<p>`, marks the beginning of the paragraph, and the second tag, `</p>`, marks the end of the paragraph. The `<p>` means “paragraph” in HTML, and browsers know to interpret that tag to mean that the text inside of it should be displayed in a certain way—with a blank space separating it from preceding or following text, for example.

HTML and XML (both of which descend from SGML, which stands for Standard Generalized Markup Language and was developed by computer scientists in the 1970s) are both based on a model of textuality called Ordered Hierarchy Content Objects (OHCO). OHCO is not itself any particular technology or language—it is an abstract way of thinking about text that is then realized in XML and HTML. OHCO holds that text is essentially composed of objects with content—a word, a sentence, a paragraph, etc. The order of these content objects is important: the second sentence should not come before the first. Finally, these ordered content objects bear a hierarchical relationship to each other: chapters may contain paragraphs, paragraphs may contain sentences, and so on.

XML and HTML implement OHCO. For example, consider this greatly simplified example of an XML-encoded book in Figure 1.

We can think of the text as divided into Tupperware containers. The largest container is the book, signified by the `<book>` tag and its “lid,” the `</book>` closure at the end. Inside that container are nested multiple chapters, and inside each chapter are a title and multiple paragraphs.

```

<book>
  <title>Ruth Hall: A Domestic Tale of the Present Time</title>
  <chapter>
    <title>Chapter I.</title>
    <p>The old church clock rang solemnly out on the midnight air. Ruth started... </p>
    <p>Ruth was not sighing because she was about to leave her father's roof...</p>
  </chapter>
  <chapter>
    <title>Chapter II.</title>
    <p>It was so odd in Ruth to have no one but the family at the wedding... </p>
    <p>"Well, I hope Harry will be happy," said Ruth's mother-in-law, old Mrs. Hall, as she untied her cap-strings...</p>
  </chapter>
</book>

```

Figure 1. Abbreviated XML encoding of *Ruth Hall*.

If we wanted to encode this book in HTML, we would use the tags provided by that language—probably `<body>` for the book, `<div>` for the chapters, `<h2>` for chapter titles, and `<p>` for paragraphs. Any browser would be able to read those tags and translate them into a display that would match our expectations. However, XML is crucially different from HTML, even though they syntactically resemble each other. XML does not provide a vocabulary, only rules for structuring the encoding. The labels given to different segments of text are entirely up to the user. This makes XML powerful: user communities or individuals can make up their own tags to note what is interesting to them about the data they are encoding—the tags used to mark up space exploration information will vary quite a bit from the tags used to mark up musical scores, for instance. TEI, then, was developed as a vocabulary of tags with which users can encode texts to note features of interest primarily to the humanities, social sciences, and linguistics. In its early years TEI offered only a few dozen tags, but today they number in the several hundred, many of which can be made more complex and specific by adding attributes and values—for example, `<hi rend="cursive">Bonjour!</hi>`, where “hi” indicates text that is graphically distinct, and `rend="cursive"` further explains that it is rendered in cursive.

The TEI tag set has been developed over many years by a user community who puts questions and suggestions before the group. It can be customized—no one uses all the tags provided by TEI, and some projects create a few additional ones that speak to their unique interests or materials. If a user decides to encode something in accordance with TEI, she inserts a line of code in her XML that points to a TEI-conforming schema, a computer-readable document that expresses the rules of TEI, and later validates her work against that schema.

XML, unlike HTML, cannot be read automatically by browsers. Because XML allows anyone to create any tags, no browser could possibly predict those tags or reliably interpret them for display. Consequently, TEI/XML requires the use of a stylesheet, a file written in XSLT (Extensible

Stylesheet Language Transformations)—a language I consider considerably more challenging and in some ways more rewarding than XML—to translate the XML into displayable HTML. Importantly, XSLT can also be used to do a great many other things with the XML, including culling all kinds of data from texts, splitting them up, and combining them.<sup>1</sup>

TEI is frequently used in large text projects to provide metadata and markup that is not much more complex than the basic structural tags one could find in HTML—using in-depth markup is very time-consuming, and the more complex and specific the encoding is, the less likely it is to apply to large numbers of heterogeneous texts. Many smaller “boutique” projects use much more specialized, descriptive TEI markup—often, such XML files include more markup than text as projects layer claims about formal structure, variants, historical context, and other topics of interest.

When I teach classes on digital editing, I address TEI/XML, XSLT, and also basic HTML and CSS (Cascading Style Sheets). Students select materials, scan them, develop an editorial methodology informed by theory we read in class, decide upon TEI tags that best accommodate that editorial methodology, encode the text in TEI, write XSLT that harvests desirable data from the texts and translates the XML into displayable HTML, and compose some simple CSS to make the HTML aesthetically satisfying. Finally, they write editorial materials to guide readers through their material.<sup>2</sup>

### **Pedagogical Benefit: Attentive Reading**

Many practitioners of TEI, including me, disagree with OHCO at least to some extent and can find some of its attendant technological limitations frustrating. However, it offers a rigorous, systematic, and somewhat flexible way for students to inscribe a view of the text onto the text itself. TEI offers a refreshing alternative to writing a term paper, which, for all its pedagogical value, nonetheless often tempts students to cherry pick textual evidence and wait until the last minute, circumventing the goals of extended, thoughtful engagement with the text. To properly encode a text in TEI in a digital editing course, students must first read the text in order to determine what features are of interest. Some of these are bound to be formal and would at least include noting the basic structural features of the work. From there, students can develop any number of interests in the text, including continued study of the formal properties—TEI includes tagsets for noting meter and scanning each line of a poem, for example—or more content-based interests, such as tracing the gender of speakers, the tribal identity of characters, the locations of places mentioned in the text, etc. The students, usually working in small groups, agree on a focus (or foci) they wish to bring to the text, then write a rationale explaining why that focus is important and what kind of critical lens informs it, as well as how they will perform that focus with encoding, including developing customized tags if appropriate.

Students then typically divide the text into equal segments and encode their portions, communicating frequently to discuss unexpected difficulties and to clarify approach. They trade their portions for quality control and to ensure continuity. This process requires each student to read and reread the text for any features that warrant encoding based on their criteria, and during this work they frequently encounter conflicts in the text when it fails to conform to their prescriptions—perhaps a meter is irregular, or gender is uncertain, or the author uses racist or outmoded terms to describe tribal identity, or places in the text are imprecise or fictitious. They are forced to account for these disruptions by adjusting their description of the text, all of which enriches their understanding and pushes them to think about textual difficulties or inconsistencies that could be easily skirted in a brief analytical essay.

In the spring of 2013 I taught a class called Digital Archives and Editions at the University of Nebraska (now a recurring part of our curriculum). It combined graduate and undergraduate students, most of whom had no prior experience with XML. It was stunning to see how some of these students developed over the semester, feeling motivated to master the technology to communicate their view of the text, then, when the technology drove them deeper into the words, developing even richer views of the text. For example, one group of three students found a nineteenth century book on beekeeping, L. L. Langstroth's *The Hive and the Honeybee*, which is a seminal text in the field. One of the students had previously worked as a beekeeper, another had a background in theology, and the third was a practicing poet. They were drawn to the book for its historical significance at a time when the fate of bees is very much in question. The student with the theology background was interested in how bees and beekeeping have historically been used as religious metaphors and how the author, who himself was a clergyman, drew on those metaphors to explain the science. The poet was intrigued by how this historical scientific book relies heavily on literary allusions and includes snippets of poetry throughout. The beekeeper appreciated it as foundational to that profession. So the group decided to bring their respective views to bear on this work. They typed all 384 pages of it, encoding structural features throughout, and as they went, they layered in three systems of notes: one from the perspective of a beekeeper, one from the perspective of a theologian, and one from the perspective of a poet. They ended the project with almost 400 notes woven into their text, annotating everything from references to Shakespeare to details about bee subspecies.<sup>3</sup>

Sometimes in public discussions digital humanities gets conflated with studying social media or “distant reading,” both of which are valuable endeavors in themselves, but prompt poorly informed teeth-gnashing about digital humanities heralding the death of close reading and textual primacy. One recent example of this is Stanley Fish's piece in *Opinionator*,



the opinion pages blog of the *New York Times*, in which he describes digital humanities as seeking to “Digitize the entire corpus” so “you can put questions to it and get answers in a matter of seconds.” Fish and several commenting readers seemed intent on pitting digital and traditional humanities against each other. One reader responded that digital humanities is “duping” students with “sub-literacy”; another—no kidding—claimed that we are going to “finish off civilization as well.” On the contrary, one of the oldest strains of digital humanities, text encoding, is fundamentally an exercise in close reading. I have never seen the duped youth in my charge exhibit as much careful attention to textual detail as they do when they have been asked to create a digital edition of a text. Several years ago at the University of Georgia, one of my undergraduates, upon completing his first digital humanities assignment, said quite eloquently that he was surprised to find that the assignment was really “an extended meditation upon a poem.” This seems to hold educational value for digital and non-digital humanists.

### **Pedagogical Benefit: Discovery and Empowerment**

In my experience, students of literature tend to think of literature as abstract and immaterial—that is, they have been trained to think of the lexical text as stable and its material and editorial history as irrelevant. *Leaves of Grass* or Dickinson’s poems seem to them to be sequences of words, but they have usually not been asked to consider the material history of those words—which are messy and fluid and pose many challenges to the clean anthologized versions received by students. Further, students tend to think of literature as a fixed field in which all the important decisions about what is included and what is excluded have already been made, usually on principled and objective grounds, by experts in the past. They sometimes know that the canon has undergone changes, mostly in response to the progressive movements of their parents’ or grandparents’ generation, but the work of deciding what is important or beautiful seems to strike them as now complete. (I am generalizing here, to be sure, but this has been my sense from many classroom conversations on these topics.)

Digital editing offers an opportunity to foreground the material histories of texts and to invite students not only to consider literature as a malleable and interpretive selection of texts, but also to suggest revisions and expansions of this selection. When students are confronted with documents (or scans of documents), sometimes drafts or published variants of each other, and they are asked to suggest a way of representing the documents, they are forced to closely consider the kinds of editorial choices that were previously obscured from them. If Dickinson actually wrote four possibilities for a word in that line, on what principle do you select one to prepare a reading text? How does the meaning of the poem change based on these decisions? Would it be better to leave the ambiguity open to read-

ers? Should we try to preserve irregularities in her punctuation and spelling, even though they often don't translate neatly into print, or should we normalize them, which may do damage to her intentions? Or would she have intended for these to be corrected if they were published? Or should we respect her decision to not publish? Students must consider these questions and develop reasoned positions on them simply in order to complete the assignment. By the end of this work, those clean anthology versions seem to be conveniences of publishing that obscure a much richer cultural past and a creative process that is often lost in the exclusive teaching of final product.

When students are asked to select texts for their digital editions, they must go exploring in the documentary record to find interesting, under-examined materials. The assignment asks them to locate materials that are not readily available in print or online, or, if they are readily available, the student must make a case about how her project will contribute a new view or understanding of that text. (I further constrain these choices by requiring the texts to be in a language I can read and free of copyright restrictions.) This work requires students to think of the voices that have been left out—it invites them to think of the canon as a fairly small and exclusive body of texts, and as they justify their selections for the assignment, they often encounter texts that complicate the view of cultural history suggested by canonical texts, either by offering alternative views or by providing a background for canonical works that makes them seem less exemplary. Sometimes students are drawn to work by recognizable authors that appeared in complex publication contexts: for instance, two students edited four Sherlock Holmes stories as they appeared in *The Strand*, understanding that the periodical context, which mingled journalism with Doyle's stories, would likely have informed the way readers received various textual details.

Asking students both to think about texts from a material or representational perspective and to contribute creatively to cultural knowledge is the hallmark of many digital humanities classes. Paul Fyfe has described this “felicitous disorientation,” being jarred out of old reading habits, and not the presence of digital media per se, as what defines the ethos of digital pedagogy—the computers are really incidental, and simply serve as an invitation to create research projects that ask students to think differently, discover, and create. Likewise, Julia Flanders, one of the most influential shapers and theorists of text encoding, has described “provocative friction” as one of the most important qualities of work in the digital humanities, and argues that “an awareness of the representational significance of medium” is foundational to DH work.

Material research, discovery, and contribution are not new to the humanities—they are essential to bibliographical studies, for example, once a core part of graduate study in literature. It also has been integral



to the study of history. In fact, the prominent (even if scandal-embroiled) historian Stephen Ambrose credited this kind of work with luring him into the field of history. He describes an assignment in which his professor asked students to delve through materials at a state archive in order to produce a biography of a neglected figure. "And that just hit me like a sledgehammer," he wrote. "It had never before occurred to me that I could add to the sum of the world's knowledge" (Ambrose). Digital humanities, in particular digital editing, offers a way for us to reintroduce this valuable tradition into our curriculum, updated with new methods and new critical perspectives.

### **Pedagogical Benefit: Writing with a Purpose and for an Audience**

Ambrose adds to his description of that assignment, "I'll never forget the feeling I had when I finished that work, and, and wrote the 10 page bio of this guy: 'I know more about Charles A. Billingham than anybody else in the world!' I just thought that was marvelous." When students in digital editing courses find materials that are interesting but neglected, when they read the materials and inscribe a painstaking understanding of their form and content onto an edition of those materials, they sometimes emerge as unparalleled experts on the topic. When they write introductions or explanatory notes for these materials, they don't worry about reaching a page requirement, nor do they rack their brains for something to say—they are filled with observations about the material and excited to offer their discoveries and interpretations online.

Digital humanities classes can significantly contribute to efforts to teach writing in such a way that students feel motivated and purposeful. The discussion of purpose in the teaching of writing goes back for decades, and recently has been explored by notable scholars of pedagogy such as Gerald Graff and Cathy Birkenstein, who have written about how we can equip and empower students to write about literature. Graff and Birkenstein argue that providing students with motivation—for instance, by asking them to summarize and disagree with a piece of criticism—allows them to produce the kind of writing that we actually want to see (1–15). In-depth textual encoding produces motivation because it forces such familiarity with the materials that students feel confident about their insights, aware that these insights are new, and aware that their writing has the purpose of making these insights public. This sense of acquired expertise on a topic can be a powerful impetus for purposeful writing.

### **Pedagogical Benefit: Transferable Skills**

Finally, TEI offers professional currency to students—more so than other technical knowledge we routinely spend time on, such as MLA citation style. TEI requires significant training time, especially if it includes learn-

ing basic XSLT, HTML, and CSS, though it is possible to introduce these skills in one semester. In exchange for this time investment, students learn technology that is not only valuable to the goals of teaching reading and writing, but is itself a professional skill. It is difficult to get an accurate count of projects that use TEI, but currently the TEI email list has over a thousand international subscribers, and seventy institutions help fund the organization. When students learn TEI/XML, they emerge familiar with a professional standard that is used around the world.

Rigorous digital editing in the humanities is currently based on TEI. This may not always be so, as the market gives rise to new technologies and digitally inclined researchers develop competing standards. However, the field itself will almost certainly always be defined by an interest in close textual attention and the creation of usable final products, which ought to appeal to a range of literature instructors and not only those who identify as digital humanists.

*University of Nebraska-Lincoln*

#### Notes

<sup>1</sup> For example, see my colleague Brian Pytlyk Zilling's project, TokenX, which uses XSLT to visualize all sorts of information about XML-encoded texts: <jetson.unl.edu/cocoon/tokenx/>

<sup>2</sup> A syllabus for my digital editing class is online at <segonku.unl.edu/digitalediting/?page\_id=9>

<sup>3</sup> The group's XML can be viewed online at <segonku.unl.edu/digitalediting/class-files/beekeeping/HHB.xml>

#### Works Cited

- Ambrose, Stephen. Interview by Academy of Achievement. *Academy of Achievement*, 22 May 1998. Web. 2 Dec. 2013.
- Fish, Stanley. "The Digital Humanities and Interpretation." *New York Times Opinionator* 23 Jan. 2012: n. pag. Web. 2 Dec. 2013.
- Flanders, Julia. "The Productive Unease of 21<sup>st</sup>-Century Digital Scholarship." *Digital Humanities Quarterly* 3.3 (2009): n. pag. Web. 1 Dec. 2013.
- Fyfe, Paul. "Digital Pedagogy Unplugged." *Digital Humanities Quarterly* 5.3 (2011): n. pag. Web. 1 Dec. 2013.
- Graff, Gerald, and Cathy Birkenstein. *They Say, I Say: The Moves that Matter in Academic Writing*. 2nd ed. New York: Norton, 2009.
- Marche, Stephen. "Literature is Not Data: Against the Digital Humanities." *Los Angeles Review of Books* 28 Oct. 2012: n. pag. Web. 10 Dec. 2013.
- Mohamed, Feisal G. "Can There Be a Digital Humanism?" *Huffington Post* 28 Dec. 2011: n. pag. Web. 3 Dec. 2013.
- Moretti, Franco. *Distant Reading*. London: Verso, 2013.
- Presner, Todd, et al. "Digital Humanities Manifesto 2.0," *Humanities Blast*, n.d. Web. 1 Dec. 2013.