Encoding the Discipline: English Graduate Student Reflections on Working with TEI

Maura Ives, Victor Del Hierro, Bailey Kelsey, Laura Catherine Smith and Christina Sumners
1. Introduction (Maura Ives)

In her keynote address at TEI 2012, Julia Flanders remarked that increasing numbers of humanities faculty and graduate students are attending TEI seminars, often for reasons other than to construct a digital archive.¹ These new participants seek out training in TEI because they know that TEI is an important tool within digital humanities, and/or because they believe that training in this (or any) aspect of digital humanities is a valuable academic or professional credential. While some of those participants will eventually create digital projects, many will use and think about TEI in other ways, incorporating it into their teaching and, more generally, into their ways of thinking about disciplinary objects and practices.

This set of short papers reflects the scenario Flanders identified. In Fall 2012, I taught an English graduate seminar titled Gender and Literature: Recovering 19th Century Women Writers. As part of the course's examination of the theory and practice of recovering writers excluded from the literary canon, we reviewed the representation of nineteenth-century women writers within print-based textual scholarship and digital texts and archives. We did not assume that digital remediation has resolved, or necessarily will resolve, problems that have proven intractable in print; instead, we approached these materials with the understanding that (to quote the syllabus) "digital textual scholarship complicates theory and practice, making old questions difficult in new ways, and creating a renewed sense of urgency around issues of canonicity and access."² To help students think through the complications of digital recovery work, the class worked together on a
project featuring the poem "Divided" by the nineteenth-century British writer Jean Ingelow, which culminated in encoding the poem using the TEI Guidelines. The primary goal of this assignment was not to teach students how to encode with TEI, but rather, to help them understand the trajectory of a digital textual recovery project by wrestling at each step of the way with the questions that normally emerge in the process of editing a literary text for a scholarly audience. At the same time, I also wanted to take advantage of the happy coincidence of the TEI conference taking place at A&M during the semester: the Ingelow project concluded before the conference, so that students had enough familiarity with TEI encoding to allow them to profitably attend conference sessions. To mitigate any anxiety that students unfamiliar with text encoding might experience, the TEI assignment was presented as a low-stakes endeavor. It counted for a very small percentage of the final grade, each student was only responsible for ten lines of text, and students did not begin the assignment until they had completed six weeks of background reading (in the history of textual recovery, editorial theory, and Ingelow's printing and reception history) to enable them to place the task of encoding Ingelow's poem into disciplinary context. One (three-hour) class period was allotted to an introduction to TEI using oXygen. The next class meeting was designated as a workshop day, during which the class worked together to encode their portions of the text and engaged in preliminary discussion of the kinds of details that would eventually form a project's encoding guidelines (what to do about indentation, stanza numbers, spaces before punctuation). Although the things that my students learned through TEI encoding sometimes had little to do with the TEI itself, their perspectives are instructive both because they are newcomers to TEI, and because they are the kind of newcomers that Flanders identified: they learned TEI encoding as part of their training as graduate students in English, and their experience with TEI was framed by the particular considerations of our seminar (which foregrounded considerations of access and inclusion within the digital realm), and by a larger, more varied matrix of disciplinary knowledge.

To prepare for their presentations, the three authors whose work appears here—Bailey Kelsey, Laura Smith, and Victor Del Hierro—decided to reflect both on what did or did not work in our classroom experience, and on larger questions of access and inclusion, the TEI and textuality, and TEI in relation to embodiment and the technologies of the book. In doing so, they had the benefit of reviewing anonymously submitted answers to a brief survey that was distributed to students in our class and in another graduate seminar, Professor Jennifer Wollock's Topics in Medieval Literature and Culture: The Ballad, in which a brief introduction to TEI was also presented. The survey questions—written by the student presenters, and distributed after the TEI assignment was completed—asked students about their previous experience with text encoding (if any), and their experiences with learning how to encode with TEI. Eleven students responded to the survey. Of those, only one had previous experience with TEI as part of professional training; five had no experience with any form of markup, and five had some familiarity with HTML.

What can be learned from the observations of a small group of graduate students concentrated within a single humanities discipline? To begin with the most obvious point, their comments vividly demonstrate why it is important to incorporate TEI in the graduate curriculum, how that incorporation might best be approached, and what happens when TEI is viewed through a disciplinary lens. Despite the trend that Julia Flanders noted in her presentation, the initial responses of students in both of the
graduate classes surveyed indicate that we have much work to do to foster a broad yet meaningful engagement with digital work in humanities disciplines. To be specific, Texas A&M has a strong, and very visible, commitment to digital humanities, and the University-wide Initiative for Digital Humanities, Media, and Culture housed in the English department, and yet, as Christina Sumners noted in her presentation, many students surveyed were unfamiliar with TEI, uncertain of its value, and puzzled as to why it would be included in their coursework, which led to initial impressions that were often negative (student survey responses included words such as "apprehensive", "intimidated", and even "terrified"), which explains why one student commented that TEI training "was not something I would have sought out on my own."

Although my students seemed quickly to grasp why TEI was important and found the experience of encoding "Divided" to be not only painless, but interesting in unexpected ways (as seen in Victor Del Hierro's meditation on hand-writing the TEI tags on a paper copy of Ingelow's poem), their apprehension underscores the fact that many students in our discipline remain wary of text encoding in particular, and digital humanities in general. If this is the case at an institution like ours, it seems safe to say that a significant portion of the next generation of humanities scholars are not prepared to interact with digital scholarship, and digital humanities scholars, in meaningful and productive ways. Moreover, if we want TEI to work in the service of diversity—including, but certainly not limited to, an expanded canon of digital content and an expanded community of digital humanities scholars—we need to acknowledge and attempt to remove the barriers that keep that from happening. My experience, and that of my students, strongly indicates that we must continue to work hard at presenting TEI in ways that bridge the distance (real or perceived) between unfamiliar activities like text encoding, and the more customary activities, values, and intellectual frameworks that our students have assimilated as part of their academic training.

We might start, as Bailey Kelsey and Laura Smith argue, with questioning our assumptions about what texts (especially digital texts) are, and about how we understand interdisciplinarity and the boundaries between technical and humanistic knowledge. We must take to heart common-sense advice such as Kelsey offers when she explains that the traditional gambit of explaining TEI by likening it to other markup languages may cause more harm than good; rather, as Sumners suggested in her presentation, we should direct students to instructional websites such as TEI by Example (Van den Branden, Terras, and Vanhoutte 2010). We must be struck not only by the visual image of Victor Del Hierro's TEI as graffiti tag, but by his insight into the importance of aligning text encoding with other writing practices, and as Smith suggests, we must be open to partnerships across the curriculum that extend TEI's usefulness as a means to enhance collaborative teaching and learning. But most of all, we must situate TEI within the classroom and within disciplinary conversations, especially conversations about such issues as gender and race in the academy, disability and the digital humanities, and the nature of texts and/as technologies.

2. Reflecting on Pedagogical Approaches to TEI (Bailey Kelsey)

Since the beginning of 2012, I have worked on two literary projects rooted in TEI encoding. Last spring, I worked with Dr. Amy Earhart to encode the Alex Haley papers...
available in Texas A&M's Cushing Library; this fall, I worked with Dr. Ives to encode ten lines of Jean Ingelow's poem "Divided". Both these TEI opportunities were part of my graduate coursework, and preparation for these projects involved extensive reading about scholarly editing and practical skills. Out of these two experiences, I have found a compounded misconception about digital texts that is central to the 2012 "print versus digital" debates.

1. The first misconception is this: digital texts and digital archives merely replicate physical texts and physical archives in a non-material environment.

2. The second misconception is that all possible digital representations of a text are created equal.

In response to the open-ended survey question ("Is there anything else you would like us to think about?") one student described the experience of being interrupted by a colleague while encoding a ballad to TEI standards. The colleague pressed this student about the usefulness of TEI against an HTML page or a PDF scan. The student was able to speak to the potential preservation advantages of TEI, but not to the actual function of TEI encoding within the larger, and very complex, world of markup languages and web pages. I want to pause here a moment to note that the student's defense of TEI as a stable preservation option for digital texts is an interesting one considering that once, not many years ago, the instability of digital texts and their non-material forms was at the center of "print versus digital" debates. But what I believe this personal anecdote from this graduate student highlights is the necessity of grounding all TEI work in graduate courses in discussions about what 'text' is and has been, and what 'digital texts' are and are not.

In his blog post, "Text: A Massively Addressable Object", published in the 2012 anthology Debates in the Digital Humanities, Michael Witmore defines texts as objects that are "massively addressable at different levels of scale" (Witmore 2012, 325). For Witmore, addressable means that the text object—whether it is a book, a digital text that has been TEI encoded, or a manuscript of a ballad—can be queried at a "certain level of abstraction" (325). He also argues that digital texts are still material because of the human labor that must go into marking them up to TEI standards, making them accessible and searchable online. This human labor is as real as the human labor of publishing a print text, or of scribes copying and recopying manuscripts (324–25).

Further, though, TEI must be understood as a digital scholarly editing tool. This requires understanding scholarly editing historically and presently, as well as understanding how TEI functions in the larger history of markup languages. But how we bring these histories into the classroom and how we relate TEI to other editing theories or computational practices must be improved upon.

The articles and manuals I have been assigned in the classroom as an introduction to TEI start at the very beginning of markup language history and text encoding practices, which can be overwhelming with their technical jargon. Common to these articles are discussions of both HTML and TEI, which then prompt comparisons between HTML and TEI in classroom lectures and discussions. These comparisons are particularly problematic because TEI is not the same kind of markup language as HTML. These comparisons, then, do not form the intended "concept bridge" from the known (HTML) to the new (TEI). From our survey responses, it also became clear that the HTML as the "known" markup language is an assumption; several of our classmates had never done any markup prior to TEI.
From our class and the ballads class, we can begin to see how the misconceptions being formed about "print versus digital" in 2012 are sometimes perpetuated by our own ways of bringing these new technological tools and theories into the classroom. Of course, we should still bring TEI into the classroom, as often as we can, but we should always be reflecting on our pedagogical strategies.

3. Considering TEI’s Placement in Writing Curricula (Laura Smith)

As a writing center tutor, I am interested in how twenty-first-century media platforms function pedagogically. In her 2005 keynote address to the Computers and Writing Conference at Stanford University in Palo Alto, California, Andrea Lunsford advocated for writing pedagogy that shifts from a hegemonic production of printed texts to a form of technology that enacts performative and embodied discourses and evokes physical, imagined, and geographical spaces ("Writing Technologies" 170–72). While designing an innovative writing course at Stanford University, Lunsford facilitated a five-year study that examined the intersection between performance art and writing practices. Asserting that "all writing is a performance" (Fishman et al. 2005, 235), Lunsford determined that verbal acts (e.g. slam poetry, role playing) allowed for "students' live enactment of their own writing" (Fishman et al. 2005, 226), and improved revisionist strategies. In this hybrid space, the questions become: how can this epistemic aspect to writing extend to include digital humanities, specifically TEI practices? How can we, as educators, use these virtual embodied spaces to reconfigure monolithic constructions of writing curricula?

Perhaps the first barrier to acknowledge is, as one graduate student noted, the "selling" of TEI to faculty, administrators, and students. Misconceptions of TEI and digital humanities (DH) make it more difficult to integrate DH projects into curricula. In our survey of two graduate classes, one student responded to the question: "What were your initial responses to learning about TEI?" with the statement "Actually, when I first saw 'text encoding' in the syllabus, I couldn't understand what I would do. And later when I found out what the text encoding is, I was a little embarrassed, because I thought this work is not for a literary scholar but for computer technicians." In her 2012 blog post, "Some Things to Think About before You Exhort Everyone to Code", Miriam Posner examines what she considers a pervasive attitude within the digital humanities field: to consider oneself a DH scholar, one must know how to code. This student response returns to Posner's question—does considering oneself a DH scholar require an understanding of coding (or, in our case, text encoding)?—but also brings into conversation how programming functions in the field of humanities and the assumption that it is situated firmly within a mathematical/scientific realm.

One approach to reconfiguring the notion of TEI as equivalent to computer coding and technical jargon is to emphasize how TEI forces users to interact beyond the physical text. Last spring at Texas A&M University, a graduate English course titled The Emergence of Digital Literary Studies, taught by Dr. Amy Earhart, encoded Alex Haley's notes and revisions of his text, The Autobiography of Malcolm X. In this course, these students created a TEI encoding in the form of genetic markup that allowed them to differentiate the colors of ink Haley used in his revisions. Rather than defining this information solely as "revisions" in the TEI hierarchy, the students' decision to deliberately note this penmanship change suggests a way that TEI can extend beyond the
scope of the factual, the presentation of text as text. Students interacted with Haley's meta-text, writing that reflects upon itself, creating a revisionist map for users to follow. Thus, implementing TEI into classroom environments engages students in conversations on the multi-dimensional understanding of 'text' and 'representation', while simultaneously contributing to scholarly research.

Partnerships with various disciplines may prove to be another strong step. In their article, "Collaboration and Concepts of Authorship", Lisa Ede and Lunsford vocalize the need for collaboration throughout academia, mentioning hypertext projects, such as the Orlando Project (http://www.srtsrn.ualberta.ca/orlando/), the Victorian Web (http://www.victorianweb.org), and the now defunct British Poetry 1780-1910 Hypertext Archive, as three examples (Ede and Lunsford 2001, 364). Ede and Lunsford argue that these forms of collaboration allow for a "plethora of theoretical perspectives characteristic of literary studies [and] also the qualitative and quantitative methods used across varying fields" (Ede and Lunsford 2001, 364). When considering incorporating DH projects into college curricula, partnerships with Writing Across the Curriculum (WAC) programs, writing centers, and other pedagogical networks may be useful.

4. Tagging Over Everything: Thinking About the Rhetorics of TEI (Victor Del Hierro)

After having encoded my section of Jean Ingelow's poem, I was struck by the simplicity of the encoding process and concept. Although the assignment only called for encoding 10 lines of a poem, the experience of a moment in the life of a TEI encoder opened a discourse with the literal way encoding text was talked and thought about. Initially it was interesting to reflect on the idea of imposing this digital structure onto a written text, while at the same time remaining conscious of the way encoding had made apparent the structures within a printed text. Influenced by a section in our class on Emily Dickinson, we are reminded that there are differences in texts when we move from written to printed and obviously now on to digital, not to mention the complications that western canons face because they usually privilege written alphabetic texts. With this in mind I decided to do my own tagging:
Performed in this way, the encoding—or more appropriately the tagging—creates an interesting visual contrast. Handwritten, the tags become less technical and much more descriptive. Instead of command or a direction, a tag like `<lb/>` is describing a line break as opposed to performing a line break.

When I was doing this, I did not know what I wanted this to be or what direction I wanted this to go. I did know that tagging, to me, meant graffiti. The origins of graffiti in New York City were a rhetorical practice that expressed agency but often served as a metaphor and avenue for exposing other discourses. Simply stated, tagging makes you aware of spaces that you did not notice before. For me, TEI encoding forced me to tag space that was always present and read but not always consciously accounted. This also forced me to privilege anything but the actual text because when I encode, I am encoding the space around the printed text and now the printed text just inhabits the space in between my tags. This then shifts the rhetorical implications of our tags. Moving the conversations about TEI encoding towards more familiar writing practices creates bridges of accessibility across disciplinary divides. The stigma around encoding as a highly technical skill creates a barrier for many first-time users that is unnecessary and easily traversed. Actively thinking about the rhetorics of encoding as both descriptive and performative texts enables encoding to be read as its own rhetorical practice. This shift also represents the consciousness of the physical embodiment of the encoder when they encode.
In "Wampum as Hypertext", Angela Haas complicates the distinction between technology and high technology by discussing the way wampum belts function the same as Western hypertexts. Haas contends that while there are many similarities, they differ when we understand the way wampum is reliant on cultural practices and memory:

Consequently one could argue that wampum is limited in relation to contemporary Western hypertexts in that it requires human intervention to remember the intent and content of the original message; however, one could also posit that such interaction encourages continuous civic involvement instead of an over-reliance on technology.

(Haas 2007, 93)

Haas goes on to discuss the relationship of memory to the body:

The body remembers the weaving and the performance of wampum. Regular performances of wampum hypertexts suggest that Western hypertexts are relegated to dormancy until the moment we need to recall them. Both conceptions of hypertext require human interactions, but Western hypertext does not require a conscious effort to remember the message encoded in the technology. Thus human memory (physiological, emotional, mental, and bodily) and material memories are connected—in alliance to foster hypertextual memory.

(Haas 2007, 93)

Drawing from Haas's discussion and comparing it to our discourse on TEI and print, we can begin to flesh out the ways in which a digital transcription becomes an embodied practice as well as a re-embodiment of printed text.

Developing and thinking through the memory and rhetorics of TEI means that we are weaving meaning into our encoding while creating meaning from our encoding. Upon a second inspection of the image of my tagging, you can see where I remember how indentation functions for encoding. Furthermore, as I reflected on the experience, I was able to think about encoding in more personally accessible ways. The texts we work with are already encoded. We know this; that is why they carry meaning and stories. Through TEI, are we aiming to expose those stories? How are we using this technology to work for us? There needs to be a conscious discussion of what this re-embodiment means: are we simply trying to reproduce what we already have? Specifically in a recovery context, we are aiming to keep certain texts alive because they have been neglected and will not last. Constantly asking "Why?" is imperative, in my opinion, because it keeps us from reproducing what we already have in print and pushes encoders to think critically about that reproduction.

Returning to my metaphor of graffiti tagging, I am drawn to the work of Jean Michele Basquiat, originally known by his tagger name SAMO. S-A-M-O stood for Same Old Shit. Basquiat found a way to re-invent what writing could be and how it functioned. What is inviting about TEI is that possibility, the same possibility of writing that Cherrie Moraga describes: "Our writing can help take us there if we require the most of it and ask it the right questions. Our journey of return is not romantic; it is ordinary. It is the dusty road of our own pitiful colonized preoccupations, which I have come to call the 'mundane.' The marvelous mundane of our lives, where the barest truths are revealed" (Moraga 2011, 85).
BIBLIOGRAPHY


Ives, Maura. 2007. "Her life was in her books': Jean Ingelow in the Literary Marketplace." Victorian Newsletter 111 (Spring): 12–19.


NOTES

1. Julia Flanders, "TEI and Scholarship" (TEI and the C(r|l)o(w|u)d: 2012 Annual Conference and Members' Meeting of the TEI Consortium, November 8, 2012, College Station, Texas).

2. The complete description reads: "The idea of "recovery" of writers omitted from the canon is not new, and the emergence of digital archives seems at first blush to render recovery a moot point, if not a fait accompli. In reality, digital textual scholarship complicates theory and practice, making old questions difficult in new ways, and creating a renewed sense of urgency around issues of canonicity and access. Since recovery spans several fields of inquiry (literary history, feminist and gender studies, bibliographical and textual scholarship, the history of the book, archive studies, and digital humanities), I imagine this course both in terms of a set of core questions about the project of recovery, and in terms of a persistent focus on the structures and practice of recovery, especially through digital means." (Ives 2012)

3. Ingelow (1820–1897) was primarily known for her poetry and for short fiction, much of it written for children. Despite her popularity in Britain and the United States, her reputation suffered in the twentieth century, such that her work is rarely included in literary scholarship or in classroom anthologies (see Ives 2007).

4. A fourth participant, Christina Sumners, was unable to revise her presentation for inclusion here. We have incorporated, with attribution, a number of her insights.

5. The questions were: 1) Do you have any previous experience in text encoding (such as HTML, or previous TEI experience)? 2) What were your initial responses to learning about TEI (your response to the concept of text encoding when you saw it on the syllabus, and/or preparing to encode a document)? 3) How would you characterize your overall experience as a newcomer to TEI? 4) If you know of any sources that we might want to consult (articles or books you've found during your research on your paper), please let us know about them! 5) Is there anything else that you would like us to think about as we work on the presentation?

6. For discussions on the digital canon, and on the lack of reflection and representation of women and people of color in digital humanities, see the essays by Jamie "Skye" Bianco, Amy Earhart, and Tara McPherson in Gold 2012.


8. Although we do not address these issues in detail in this paper, they were frequent topics of discussion in class, given our overall concern with exclusionary cultural and institutional practices. We would like to draw special attention to George H. Williams's observation that "[d]igital knowledge tools that assume everyone approaches information with the same abilities and using the same methods risk excluding a large percentage of
people. In fact, such tools actually do the work of disabling people by preventing them from using digital resources altogether" (Williams 2012, 202).

9. Dr. Earhart's project was conducted within her graduate seminar, The Emergence of Digital Literary Studies.


11. The page image is from the first United States edition of Ingelow's Poems (Boston: Roberts Brothers, 1864). The image source is Open Library (http://openlibrary.org/books/OL23325911M/Poems).

ABSTRACTS

Students in two graduate English seminars at Texas A&M were offered an introduction to TEI encoding and were surveyed as to their experiences. Four students presented short papers to the TEI meeting that incorporated both the survey responses and their own perspectives as relative newcomers to TEI. These papers demonstrate why it is important to incorporate TEI in the graduate curriculum, how that incorporation might best be approached, and what happens when TEI is viewed through a disciplinary lens.

INDEX

Keywords: English studies, graduate education, access, inclusion, rhetoric, Jean Ingelow, graffiti