Society for Textual Scholarship, NYU, March, 2009
Erica Fretwell and Matt Cohen
Duke University, Department of English
enf3@duke.edu

Interface Design for Static Multimedia Documents

[**SLIDE 1**] In the field of scholarly editing, there are few projects that gather and display an author's manuscript marginalia. Annotations made by a writer in the margins of printed texts or images are crucial sources for analysis in literary, philosophical, and historical study because they are rare evidence of direct interaction between a reader and his or her wide-ranging influences. But the visual dimensions of annotations--including the fact that they often either contain or appear on images, not just text--no less than problems of attribution and scale, have made it difficult for print editions to become common. Even in electronic form, annotations are difficult to handle; recent discussions on the TEI listerv have shown the heterogeneity of approaches to them.

With support from an NEH Digital Humanities Start-Up grant, we have created a set of software technologies and encoding practices that allow for the tagging, displaying, and searching of static documents that mix print, manuscript, and visual images--documents such as printed texts or images bearing handwritten annotations. These technologies include a suggested approach for encoding coordinates in XML transcriptions so that search engines can visually display results of user searches for manuscript words and phrases; web-based software for linking XML editing programs to an image display to allow encoders to relate bitmap images to XML text; and model stylesheets capable of displaying transcriptions of annotated documents together with digital images of those documents. We have kept Peter Robinson's warnings about the tendency of previous markup interfaces to be difficult to use--warnings based on his

experiences with his own software, Collate--in mind. "If everyone who wanted to make digital editions was required to use [existing] tools," Robinson writes, "very few digital editions would ever be made." The same is true, Robinson says, of TEI. We don't want to steer entirely clear of what Robinson terms the "intense theological disputes" that can ensue upon becoming invested in TEI as a basis for text encoding, but here we do want to focus a bit on the ways our tool tries to make some encoding easier and faster.

And there are tools out there that are inspirational, such as one similar to ours being developed by Doug Reside for the *Our Americas Archive Partnership*. Then there is the ARCHway Project, which features a suite of powerful tools for relating texts to images and for managing multiple hierarchies. An important goal of the ARCHway Project is to make markup simple enough to be used by transcribers with little familiarity with information encoding and portable enough to work in multiple computing environments for widely different kinds of archival projects. This goal is significant for us, too. Given that automation of manuscript transcription seems distant on the horizon, an efficient user interface for coordinate designation of words and images is crucial for making possible large-scale manuscript editing projects like that at the Whitman Archive (there are literally thousands of extant Whitman-annotated documents). Unlike ARCHway, however, our interface is web-based, and allows encoders to mark space as a structural entity.

Our test documents come from Duke University's Special Collections' holdings of documents and manuscripts annotated by Walt Whitman, which highlight the intersections among authorship, textuality, and electronic editing. A particularly unusual class of documents that Whitman created has posed difficulties for electronic rendering.

We call them "flipbooks," to denote a text that, while supported on a single sheet is, in essence, a scrapbook with multiple leaves glued on top of each other. That is, "flipbooks" are documents that are annotated and then layered on top of each other so that one can flip each clipping or page over. Sometimes there are a small number of permutations [**SLIDE 2**], sometimes a large number [**SLIDE 3**]. These documents share some of the problems Marta Werner discusses in her edition of Emily Dickinson, Radical Scatters.

Using a set of Whitman's "flipbooks," we attempted to design an interface that will maximize audience and utility while addressing theoretical issues (and, we hope, raising new questions). What exactly constitutes marginalia, and how might one design guidelines flexible enough to include its many morphologies? How would one digitally render (both in XML and on-screen) marginalia, a kind of inscription in which positionality and size are as important to meaning-making as literal content, and through which a document's hierarchies often shift or become undefined? To what extent is publication, both digital and non-digital, more broadly a spatial act in ways that marginalia help throw into relief? Today we'll try to show how we tried to make a virtue of the spatial challenges offered by marginalia by sketching out adaptations of TEI P5 guidelines that would address the layered relationship among space, transcription, and tagging characteristic of printed or photographic documents bearing manuscript annotations.

Other recent projects dealing with marginalia or with collage-like documents have departed from a strictly TEI-based approach. For example, the William Ewart Gladstone marginalia collection at St. Deiniol's, for example, puts marginalia

transcriptions and references into its catalog system, GladCAT. Citing the time involved in fully marking up marginal notes in Gladstone's substantial corpus, the GladCAT builders take "the individual … as the smallest unit of reference." "One of the principles of GladCAT is that it is a *guide* to the annotations, but," its makers specify, it "cannot be a substitute for consultation of the physical volume." This approach has many advantages, particularly for annotation sets that are comparatively systematic. Gladstone wrote a key to decoding his annotational marks in the back of one of the volumes; including a couple of marks for which he didn't offer explanations, about eleven different marks dominate the marginalia. A disadvantage of this system is that this simplified capture system precludes including some important things, such as very long marginal commentaries or Gladstone's indication of accent marks. Paul Olsen-Smith's project *Melville's Marginalia Online* doesn't use TEI; nor do Paul Dyck and Stuart Williams, in their edition of a biblical concordance made for King Charles the First by the Ferrars of Little Gidding. Dyck and Williams point out that:

> The interesting way that our project bumps against the limitations of existing encoding standards…is that we are trying to encode cut-and-paste, something no one has previously needed to do. […] [T]he TEI has not been conceived to describe a collage of texts. Thus, the most challenging and inspiring quality of this project is to imitate the Little Gidding community in using common structures to construct a particular meta-structure.

The cut-and-paste problem is one we share with Whitman's flipbooks. Our focus was less on the problem of attribution, though, than on the problem of tagging dimensionality--cases in which clippings aren't pasted down completely, for example.

The question of how to tag marginal annotations had, in the months leading up to our grant, seen some discussion on the TEI listserv. First, it was clear that the issue involved both TEI markup and stylesheet handling, since the para-hierarchical nature of annotations meant that how a comment was rendered on the screen would impact the degree to which it expressed the intellectual relations in the source text. Second, it was clear that different projects had different solutions to this, of varying complexities. None of the projects posting were, as we are, dealing with three-dimensional documents, and many were not dealing with multimedia documents (usually manuscript-only or print-only ones). We had hoped at one point to offer an all-purpose solution for this heterogeneity, but we ended up thinking that a variety of approaches is a good thing. Some annotations stand in a comparatively uncomplicated relationship to document hierarchies, and existing approaches handle these well. With the graphical interface design we'll demonstrate in a moment, however, we made decisions that we believe would work for many projects that handle such documents.

We conscientiously committed to P5 early on, and attempted first to use the <facsimile> element group--as described in section 11.1 of the P5 guidelines--to relate the spatial specificity of annotations or multiple media presences to TEI-encoded transcription. The <facsimile> element allows encoders to designate hierarchies of surfaces with zones, underneath them--each, conveniently, potentially having its own image file. We found ourselves, however, needing some way to indicate <zones> that crossed two <surfaces>--a common feature of three-dimensional documents. After creating a milestone element, <vb> (standing for "vertical break") as a kluge for handling this problem of the relationship between surfaces, and after consulting with the graphic

interface designer, we instead decided to table <facsimile> and extend the existing *Whitman Archive* schema such that it could handle pages -- surfaces -- as intellectually significant structural units. As a contributor to a discussion last month of the <facsimile> element on the TEI list put it, "In some types of physical documents, there is a clear rhetorical structure as well as the page. … Currently our encoding model doesn't actually capture the knowledge that we bring to … a set of facsimiles" (Dan O'Donnell, TEI-L, 18 Feb 2009).

The basic conceptual difficulty is simply that the <facsimile> scheme imagines the object it is representing as a digital surrogate of the original text--and a digital scan is relentlessly two-dimensional. As a hierarchical markup language, XML can capture three-dimensionality, but it can't do so easily. And TEI alone, as Jerry McGann and others have pointed out, cannot preserve the proliferating hierarchies that a literary text is capable of spinning out. Add physical dimensionality as a function of those hierarchies, and it gets even more difficult.

Concurrent markup, as described by Andreas Witt, et al, and others, may provide a solution in some cases for capturing conflicting textual hierarchies. An approach that foregrounds the process of relating competing hierarchies to each other may also be possible; instead of producing a single compound XML file, an interface could draw together different marked-up versions of the same text, layering them for user interaction. Using standoff markup together with an interface designed to show articulation and conflict points would make TEI more viable for documents such as the Whitman flipbooks, but the <facsimile> element still wouldn't quite work. As Alex Dekhtyar and Ionut Iacob have observed, there are no "complete answers for the problems of

management of concurrent XML data." Attempting to capture a text's physical dimensionality as a meaningful part of its interpretation quickly shows the persuasiveness of Jerry McGann's observation that "digital markup schemes do not easily…map to the markup that pervades paper-based texts" ("Making Texts").

Our current implementation utilizes a <page> and model that allows content to be encoded hierarchically. In this system document entities are represented as s with attributes used to differentiate the various physical and conceptual structures that they encode. Physical ordering is expressed through the hierarchy of the elements with the inner-most element within a stack of elements having the highest or top-most position on the page. Occasionally, intellectual units will cross surfaces intact, even as the surfaces impact their interpretation or potential field of reference. In these cases, we chose to join s using the @prev and @next attributes and an ID system. This seems clunky, but we ended up liking the way it calls attention to hierarchies that cross each other.

The coordinate-capture interface allows an encoder to insert a transcription of a document, draw layers on a digital image or images, and then quickly arrange and designate individual word-spaces [**SLIDE 4; SLIDE 5; SLIDE 6; SLIDE 7**]. Each coordinate space becomes a div, and the tool is extensible to perform more elaborate relating of element-spaces as individual projects desire. For this prototype, we focused principally on easing coordinate capture for individual words and on designing an interface that could handle documents with multiple layers attached to single supports.

The encoding tool maintains a direct association between the image surfaces and the markup expressed by the user, either as the encoding of structural elements or the

transcription of document content. The tool also understands the basic structure of the transcribed text (line and word breaks, paragraphs, etc.), allowing it to build a spatial model of the word positions on the page with minimal user input. This internal model can be saved to and restored from an XML document, allowing the user to work visually while at the same time seeing the encoded output. The search and browsing interface uses the same basic framework as the capture interface.

The generation of a text-based interface, through which researchers can quickly read documents and copy them into other applications for use in publications, is still in progress. Brian Pytlik Zillig, the *Whitman Archive*'s stylesheet designer, has created a basic stylesheet rendering of the text content of the XML files, paired with the source images. But this rendering can't do much more than rely upon the image browsing interface to suggest to users the specific spatial points at which different components of the document are related to each other. So we have started work on a more sophisticated implementation that will pass the coordinate information to ImageMagick to allow for a more dynamic visual presentation of the text transcription. This way the text representation can draw on the spatial relations of the text-chunks in the original document. With this interface, users will be able to "flip" through different components of the document, displaying combinations of text that more closely approximate the reading options of the originals.

With this interface, then, we've tried to make transcription and coordinate capture of manuscript elements more efficient, and with our underlying markup scheme, to treat spatial elements, contra TEI's impulse, as (optionally) intellectually significant. This approach allows the user to look at different layers of a document in different

configurations. In the longer run, it might also allow searching by coordinates--that is, delimitation of content or structure searching by, say, a specific region of the page such as the top-left-hand corner. Different projects can, implementing our scheme and using this free tool, define the role of surfaces or zones differently. Digitally implementing Coleman Hutchison's insights about the significance of page breaks in Shakespeare's sonnets, for example, might imply a different sort of markup strategy than our corpus of texts did. So we've tried to make our schema as generic as possible with respect to divs and coordinate encoding, rather than to prescribe a specific intellectual framework for marking up mixed static media texts. But there is a hermeneutics implied here: we are suggesting that encoding with the notion that texts are three-dimensional is important in some cases, and that allowing for the maximum potential flexibility for users to see all the various sides and articulation points of documents might shape the planning stages of both encoding and the interface design for a text.

WORKS CITED

Bradley, Matthew. "GladCAT: An online catalogue of the books and reading of William Ewart Gladstone at St Deiniol's Library." St. Deiniols Library Web Site, January 2009. Accessed 1 Mar 2009. <http://www.st-deiniols.com/cms/goto.asp?id=142>

Dekhtyar, Alex, and Ionut Emil Iacob, "A Framework for Management of Concurrent XML Markup." Data and Knowledge Engineering 52.2 (2005): 185–215.

Dyck, Paul and Stuart Williams, "Toward an Electronic Edition of an Early Modern Assembled Book," Computing in the Humanities Working Papers A.44, (July 2008). <http://www.chass.utoronto.ca/epc/chwp/CHC2007/Dyck_Williams/Dyck_Williams.htm>

Email correspondence between Matt Cohen and Gabriel Bodard.

Hutchison, Coleman. PMLA piece.

Kiernan, Kevin, et al. "The ARCHway Project: Architecture for Research in Computing for Humanities through Research, Teaching, and Learning." Literary and Linguistic Computing 20, Suppl 1 (2005): 69-88.

McGann, Jerome. "Marking Texts of Many Dimensions." Companion to Digital Humanities. Eds. Ray Siemens, John Unsworth, and Susan Schreibman. Oxford, UK: Blackwell, 2004. 198-217. <http://www.digitalhumanities.org/companion/>

Our Americas Archive Partnership. Rice University and the University of Maryland. <>

Werner, Marta. Radical Scatters: Emily Dickinson's Fragments and Related Texts, 1870-1886. Ann Arbor: U of Michigan P, 1999. <http://www.hti.umich.edu/d/dickinson/>

Robinson, Peter. "Current issues in making digital editions of medieval texts—or, do electronic scholarly editions have a future." Digital Medievalist 1.1 (Spring 2005). <http://www.digitalmedievalist.org/journal/1.1/robinson/>

TEI P5 guidelines.

TEI-L posts, April 2008. <http://listserv.brown.edu/archives/tei-l.html>

Witt, Andreas, et al. "Unification of XML Documents with Concurrent Markup." Literary and Linguistic Computing 2005 20.1: 103-116.