Algorithms in the Archive: Explorations of the Early Modern Digital Classroom
Sarah Connell and Julia Flanders, Women Writers Project, Northeastern University

In this workshop, participants will explore approaches to teaching early modern women’s writing that situate students as active editors, curators, transcribers, and annotators of early texts. The session will begin with a discussion of the opportunities for active student involvement made possible through digital resources such as those published by the Women Writers Project, including:

- Transcription and encoding activities that ask students to act as editors, making decisions about how to represent early modern texts in an online environment, determining which kinds of editorial information might be necessary for their audiences, deciding how to handle regularization of orthography and typography, and managing the many other considerations at stake in editing and publishing texts;
- Archive building activities that help students become familiar with the importance of metadata in accessing and working with texts, as well as the intellectual challenges in setting archival scopes and determining selection principles;
- Anthology and exhibit curation activities that ask students to locate and contextualize a set of early texts or other materials, providing hands-on experience with discovering texts in online databases, researching those texts, and sharing them with an audience through classroom presentations or online exhibits and anthologies;
- Keyword searching activities in which students work with a collection of texts, researching important concepts in the early modern world by studying the discourse for discussing those concepts; and
- Algorithmic analyses that rely on new methods in the digital humanities to discover patterns in large collections of texts like Women Writers Online, giving students technical skills that can be applied to a broad array of research questions.

In discussing these activities, we will use the WWP’s materials as examples, but we will also note where other web resources might be used. The pre-circulated readings will include several sample assignments, with discussion of how those assignments worked (or didn’t work!) in the classroom, as well as essential background information on the materials we will be treating as exemplars. Readings will enable participants to become familiar with these publications and with some of the ways they have been used in the classroom—they will also provide participants with specific instances to ground our discussions during the session.

Following this introduction, we will focus our conversations around an activity from the field of algorithmic analysis and machine learning, working with a web-based interface for exploring word associations in Women Writers Online. We will briefly demonstrate the interface to the participants and have a few minutes for experimentation; then, using a sample activity as a starting point, we will discuss how that activity might be modified for different kinds of classrooms. This will enable participants to share their own experiences in developing assignments and learn from those in other disciplines. We will open the conversation out to brainstorm a set of “dream” assignments, discussing in each case how these assignments draw
on both disciplinary and digital expertise. Finally, we will consider the broader questions raised by digital pedagogies, which include:

- What kinds of methodological questions (about transcription and editing practice, about research methods, about the data beneath the interface, etc.) can be effective in catalyzing student interest?
- Putting algorithmic approaches into conversation with more “traditional” digital practices such as transcription or curation, what might be some of the interesting dialogic moments between a view of texts as words and a view of texts as documents?
- How can we make the messiness and unpredictability of student interactions with texts and technologies a virtue rather than a challenge in teaching with these approaches?
- How can we turn routine “search” or “reading” activities with digital resources into opportunities for more active intervention?
- How can digital resources serve to mediate the distance that students often feel in working with premodern texts?

This workshop will be designed to bring in participants from a broad range of specializations, cohering around shared interests in teaching early women writers and exploring the potential of digital resources to make premodern materials accessible to students at many levels. Participants will have the opportunity to share their own experiences with selecting technologies and using them in the classroom—and to hear from others about which approaches work best in different contexts. The session will connect most directly with the conference theme on “Pedagogy and Public Humanities,” but our discussions will also foreground the importance of collectivity and collaboration in digital pedagogies, the potential for technology to confront intellectual hegemonies and disrupt disciplinary boundaries, and the centrality of choices made by both teachers and students in the digitally-driven activities we will explore together. Participants will come away with a more grounded understanding of the digital technologies that are on the forefront of the early modern classroom, as well as specific ideas for assignments, lesson plans, and pedagogical interventions—and how these might be adapted to suit different educational scenarios.
Readings:
For optimal formatting and to access all links, please follow the links below to read these texts online.
https://digitalpedagogy.mla.hcommons.org/keywords/reading/
[Please read the essay and the first four curated artifacts; the other artifacts are optional.]

Women Writers Project. “Methodology for Transcription and Editing.”
http://wwp.neu.edu/about/methods/editorial_principles.html

Women Writers Project. “Collaborative Annotation Assignment.”
http://wwp.neu.edu/wwo/teaching/assignments/annotation.html

Women Writers Project. “Word History Exploration.”
http://wwp.neu.edu/wwo/teaching/assignments/word_history.html

Suggested Readings and Additional Materials for Review:
https://jitp.commons.gc.cuny.edu/digital-close-reading-tei-for-teaching-poetic-vocabularies/

Payton, Jason. “Women Writers in Review Tag Investigation Project.”
http://wwp.neu.edu/wwo/teaching/reports/payton_wwir-tagging.html

Leslie, Marina and Sarah Connell. “Encoding the Archive.”
http://wwp.neu.edu/wwo/teaching/assignments/archival.pdf

Women Writers Online
http://wwp.northeastern.edu/wwo/
[WWO is subscription-based but we can provide attendees with login credentials.]

Women Writers in Review
http://wwp.northeastern.edu/review

Women Writers in Context
http://wwp.northeastern.edu/context

Women Writers Project Teaching Materials
http://wwp.neu.edu/review/teaching
CURATORIAL STATEMENT

In the beginning, we learn to read; after we are literate, we read to learn. In this received wisdom of early childhood education, reading is only temporarily difficult, material, intractable; afterwards, it recedes into the background, becoming the transparent skill through which we access worlds of knowledge. But we are in fact always learning to read and always learning about reading as we encounter new languages, genres, and forms, and mediums.

How we learn about reading’s past, present, and future seems especially important right now. Reading has been transformed by digital media and computation, by the computer screen and the e-reader—just as it was by the advent of machine printing, by the Gutenberg press, by advances in manuscript production, and by the shift from scroll to codex, as well as by social and historical changes in everything from the economics of book production to literacy and access to education. According to the most anxious reports, the advent of the screen is destroying our ability to read deeply, slowly, and linearly. “It is a cliché universally acknowledged,” writes Rita Raley in the introduction to her “Distracted Reading” syllabus, that we (where “we” means, variously, the Twitter generation, the Millennial generation, participants in the Network Society) no longer read, or if we do read, we read poorly, with insufficient attention and affect. Reading, by which is meant literary reading, is said to be a “lost art” and certainly “at risk.” We multitask and thus cannot sustain the kind of focus and attention required for a long, complex narrative.

Of course, as students of reading hastily reply, we have always read discontinuously, extensively, and intermittently as often as we have read in long uninterrupted stretches and intensively. Accordingly, all of the syllabi and assignments featured and referenced here—nonlinear reading via hypertext, commonplacing in digital form, the digitalization of nineteenth-century marginalia—alight on our current moment in the history of reading as an opportunity to rethink reading’s history. Each seeks to take the measure of the continuities and discontinuities between the codex’s page and the page on the screen. In Book Was There: Reading in Electronic Times, Andrew Piper offers one compelling description of how we imagine this continually-ramifying difference between the book and the digital text:

If books are essentially vertebral, contributing to our sense of human uniqueness that depends upon bodily uprightness, digital texts are
more like invertebrates, subject to the laws of horizontal gene transfer and nonlocal regeneration. They, like jellyfish or hydra polyps, always elude our grasp in some fundamental sense. What this means for how we read—and how we are taken hold of by what we read—is still far from clear. (2-3)

The fundamental uncertainty over what it means to add the distributed electronic text to an ecosystem of literary reading previously defined by the “vertebral,” bounded book (not to mention the printed form, the magazine, the newspaper, the manuscript, and the typescript) leads to a new pedagogical challenge for teachers whose fundamental charge is to teach students “how to read”: how do we begin to teach students to understand the relationship between page and screen if we have only begun to grasp it ourselves? Thinking about the power dynamics of reading in Jane Austen and the Secret of Style, D.A. Miller notes the dangers of “being read reading”(2). Reading in the digital age reduces none of these complications of reading’s relation to identity and survival, but it raises yet another specter: the possibility that we are “being read” as data by machines as well as by humans.

Machine reading has a longer history than we sometimes assume. Work on the nineteenth- and twentieth-century history of talking books and other reading machines shows how reading technologies developed to convert text into sound for blind readers questioned assumptions about the nature of reading while also contributing to the development of machine reading technologies like optical character recognition (OCR). In the realm of literary interpretation as “reading,” of course, Stephen Ramsay has famously emphasized the continuities between human and machine. For Ramsay, literary criticism already contains “elements of the algorithmic.” Both machine and literary-critical reading have in common a tendency towards strict repetitive protocols:

Any reading of a text that is not a recapitulation of that text relies on a heuristic of radical transformation. The critic who endeavors to put forth a “reading” puts forth not the text, but a new text in which the data has been paraphrased, elaborated, selected, truncated, and transduced. This basic property of critical methodology is evident not only in the act of “close reading” but also in the more ambitious project of thematic exegesis. In the classroom one encounters the professor instructing his or her students to turn to page 254, and then to page 16, and finally to page 400. They are told to consider just the male characters, or just the female ones, or to pay attention to the adjectives, the rhyme scheme, images of water, or the moment in which Nora Helmer confronts her husband… (Reading Machines: Towards an Algorithmic Criticism, 16)

Many of the assignments collected here show us this continuity in action. Whether they ask students to keep a commonplace book on a Pinterest board or teach close reading through
Markdown encoding, each reveals how the protocols of historical readers and reading communities find affinities with or form the origins of computational or digital forms.

The study of reading’s history spills outside traditional boundaries of discipline and department, and takes place in special collections library reading rooms and digital studies labs as well as classrooms. The course on the history of reading in the west from scroll and codex to the ebook and computer screen is a genre in its own right; examples are Ann Blair’s History of the Book and of Reading, Timothy Burke’s The History of Reading, and Adrian Johns’s A History of Reading. But the pasts, presents, and futures of reading make their way into a much wider range of syllabi, appearing in classes on topics ranging from literature, book history, and digital humanities to user experience and web design, cognitive science, and library and information science. I should note that I have adhered fairly strictly the the “digital pedagogy” rubric here, thus omitting important history of reading assignments like Sarah Werner’s “Your book’s early readers and users” and materiality of reading assignments like Jeffrey Makala’s “Reading By Candlelight” as well as crucial guides like How to Read in College and Methods for Studying Online Materials.

The following exercises, projects, and resources bring the continuities and disjunctions between print and digital forms to the fore, and ask students to interrogate cultural messages and social assumptions about reading—their own and others’—by historicizing reading, reflecting on their own modes of engaging with texts on and offline, and speculating on the futures of reading. These assignments rise to the challenge of grappling with reading today by taking hold of one of the most longstanding, traditional aspects of humanities pedagogy: the classroom’s role as a place where students and teachers both create and transmit new knowledge together. Just as digital technology and social media have given the collective aspects of reading a new visibility (if also a new ephemerality)—we can turn on social highlighting on our e-readers, comment on blog posts, or socially bookmark—so too has digital pedagogy brought a new kind of visibility to the longstanding practices through which students and teachers have created knowledge together in the classroom. Whether they imagine reading as distracted, discontinuous, materialized, black-boxed, close, distant, human, or mechanical, these assignments work to reveal that reading is social even when it feels most solitary. Whether we read alone in front of a screen, aloud in a mosque, or silently in a crowded library, they show us, the practices of reading are always both material and collective.
CURATED ARTIFACTS

Topic: Discontinuous Reading

Cut/Copy/Paste: Remixing Words

- Artifact Type: Syllabus
- Artifact Permissions: CC BY-NC-SA 3.0
- Creator and Affiliation: Whitney Ann Trettien, University of North Carolina at Chapel Hill

Whitney Ann Trettien’s “Cut/Copy/Paste: Remixing Words” (Duke University, Spring 2012) unsettles the apparent opposition between writing and reading, explaining to her students that “you cannot read the texts of this class without, in some sense, writing them, and we’ll spend a good deal of time doing both in and out of class.” Her syllabus traces the history of readers who have remixed old texts to create new experimental writing, inviting her students to from Latin cut-up poems to seventeenth-century German paper instruments to generative computer programs. Trettian’s readings and exercises insistently emphasize the materiality of both page and screen, expertly preempts the inscription of a material/dematerialized divide between the printed and the digital book, an especially tricky aspect of digital pedagogy. The brief, specific exercises – including instructions to excerpt, extract, cut up, collage, and auto-generate writing via both human and machine – could be integrated into anything from a first-year writing class to an upper-level history of reading or literature class to help students think about the materiality of reading and writing.

Theory and Practice of Hypertext

- Artifact Type: Syllabus
- Source URL: http://www2.iath.virginia.edu/courses/ensp482/syllabus.html
- Artifact Permissions: contacting author, will update
- Creator and Affiliation: John Unsworth, Brandeis University

Hypertext was revelatory to scholars in the humanities because it held out the possibility of rendering visible the interconnected webs of citations scholars see when they read. Courses on hypertext genres proliferated in the 1980s and 1990s—from the several still (precariously) preserved online, it’s likely that scores of these courses existed. John Unsworth’s ENSP 482: Theory and Practice of Hypertext (University of Virginia, Spring 1999) introduces students to the possibilities of discontinuous, non-narrative, online reading and writing. The course begins with a brief history and how-to, moves through hypertext classic texts, and includes several weeks of
project-based work before finishing with an investigation of emerging hypertext genres (critical, creative, autobiographical). Searching for hypertext course syllabi provides a sharp reminder of how badly preserved and quickly lost the artifacts of digital teaching often are, and of how important it is that we work to preserve and contextualize our digital pedagogy materials if we want them to be around for future students, teachers, and scholars. Most citations or links to hypertext syllabi or assignments I could find were broken, and the student projects linked from course websites seem to have an even lower survival rate. Unsworth’s syllabus offers a useful survey of hypertext theory and practice that would be a useful addition to courses on electronic literature, the history of narrative, or the introductory or history weeks of an intro to digital humanities class; it would make compelling reading for any methods of literary study class examining citation practices.

The Commonplace Book Assignment

- Artifact Type: Assignment
- Source URL: http://jitp.commons.gc.cuny.edu/the-commonplace-book-assignment/
- Artifact Permissions: CC BY-NC-SA 3.0 US
- Creator and Affiliation: Vimala Pasupathi, Hofstra University

This assignment reimagines the short close reading paper as a commonplace book. Early modern readers often copied out passages from their reading into notebooks to make a personal repository of extracts for future reference. Vimala Pasupathi writes in her write-up of this assignment that “it seems to me that there is no better way to teach students about the consumption of literary works and textual forms than to have them assemble their own Commonplace books.” Connecting early modern commonplace to contemporary uses of Pinterest and Tumblr, Pasupathi asks students to create their own commonplace book by copying out weekly passages from the assigned Shakespeare play and writing paragraph-long close readings of each passage using in “a hard-copy diary of sorts, or on the web as a blog, Tumblr, or Pinterest page.” This assignment asks students to read work by Adam Hooks and Alan Jacobs on commonplace books. Another version of the assignment might also ask students to look first at digitized commonplace books like those in the “commonplace books” section of the online exhibit Reading: Harvard Views of Readers, Readership, and Reading History as examples. The assignment could be easily adapted to any class taking up questions of collection, quotation, or close reading, whether centered on early modern culture or not, though it gains extra meaning in the context of a class on modern reading practices. In an upper-level undergraduate or a graduate class, this assignment could scale all the way up to asking students to build and maintain a commonplace book of research or open research notebook on the model of Whitney Ann Trettien’s, W. Caleb McDaniel’s, or Shawn Graham’s.
Cassey Dickerson Friendship Album Project

- Artifact Type: Assignment
- Source URL: http://lcpalbumproject.org/
- Artifact Permissions: Creative Commons Attribution 4.0 International License
- Creator and Affiliation: Cassey Dickerson Friendship Album Project Team, the Library Company, Swarthmore College, Bryn Mawr College, Rochester Institute of Technology, and Rutgers University

This site includes a collaboratively-created digital edition of the nineteenth-century friendship album of Amy Matilda Cassey, a prominent member of Philadelphia’s black middle class. As the site notes, Cassey is a significant figure in part because she was an active member of “the Philadelphia Female Antislavery Society, local black literary and debating societies, and other reform movements;” her album “is filled with calligraphed poems, inscriptions, essays, sketches, and watercolors which illuminate and further document the active and intimate connections of Philadelphia’s black leaders to a larger network of activists and reformers” including contributions from Sarah Mapps Douglass, Margaretta Forten, Frederick Douglass, William Lloyd Garrison, and Wendell Phillips. The extensive annotations on the Cassey album were created by the Early African American Print Cultures class of Swarthmore professor Lara Cohen in collaboration with Library Company’s Curator of African American History Krystal Appiah. The work of metadata creation, transcription, and annotation prompted the students to think in sustained ways about the remediations involved when the historical records of private reading and writing are transformed in order to allow, in the words of the students’ unpublished research narrative, “public readerships to find significance in private networks.” An example of project-based learning, the digitized, transcribed, and annotated albums could also form the basis of assignments on the history of reading, print circulation, and the digitization of archival materials. In addition to serving as a model for project-based work at the undergraduate level, this project could engage secondary school students. The [Abolition Seminar] includes a lesson plan for middle and high school students asking then to explore the Cassey album and use it as a model to create their own fictional album for a historical figure.

METHODOLOGY FOR TRANSCRIPTION AND EDITING

GENERAL PRINCIPLES

The WWP encodes works in English, or in English translation, by women before 1850.

In general, we encode entire texts rather than excerpts, and our transcription includes all front and back matter, including paratextual material which may not necessarily be by the author. We excerpt only in cases where the desired text is a very small part of the total published document, and where transcribing the entire document is currently impractical.

Our transcriptions are encoded in XML, following the specifications of the Text Encoding Initiative (TEI), with documented TEI extensions to accommodate the needs of our particular corpus and approach. We produce a transcription of each text that preserves the original spellings, typographical errors, lineation, hyphenation, and other details of the text. The transcription also records corrected readings for typographical errors, expansions for uncommon abbreviations, and regularized versions of old-style typography (such as the use of “i” for “j” and “u” for “v”). These alternative readings are represented through the TEI markup so as to support flexibility in display: using an appropriate stylesheet, we can display the text with or without original lineation, typography, errors, and so forth. These practices are described in more detail below.

THEORY OF THE TEXT

We treat the text as a document more than as a work of literature: hence our approach emphasizes transcription of the full document rather than only the “work,” and preservation of renditional details, original spellings, and errors, rather than their effacement. In addition, each document is treated as a circulating cultural artifact, whose historical specificity is part of its value. As a result, we do not emend the text or create critical or synthetic editions; each encoded text is a transcription of a particular physical object.

XML and the TEI tend to imply a theory of the text which emphasizes its structure as an important ontological fact about the text’s existence. The WWP believes that this kind of encoding provides an intuitive and productive way for scholarly users to read and navigate the text; however, we do not insist on it as the only possible theory of the text.

INCLUSION AND EXCLUSION CRITERIA

The WWP’s inclusion criteria in principle are very broad: over the (very) long term, we seek to capture a substantial subset of printed texts by women in English dating from 1850 or before. Within those broad constraints, the question of which texts to focus on, and which to transcribe first, has been challenging and our approach has changed somewhat over time. In the early phases of the project, we gave priority to texts based on a number of factors including unavailability in print, scholarly and pedagogical demand, and our ability to handle the text
within our operational constraints. With the advent of large-scale digitization efforts such as the Internet Archive, EEBO and ECCO, and Google Books, simple scholarly access is of diminished concern, and we have also increased our attention to questions of cultural and geographic diversity. We assess current scholarly interest in periods, genres, and issues, and seek to develop the collection in broadly representative directions, while giving priority to texts which respond to the current and developing needs of the scholarly community.

Our steering committee meets regularly during the course of the year to discuss priorities for transcription, which often take the form of a thematic cluster of texts (to be accompanied by one or more related exhibits in Women Writers in Context). These choices are driven by the research interests of our community of users (including the WWP itself). We actively seek nominations of texts and authors from the public. Our choices and prioritizations necessarily represent a balance between aspiration and practicality; given our limited resources for transcription, the value of very long and very complex texts needs to be balanced against their cost in time and effort. When possible, we partner on transcription with other projects and with external collaborators, but our encoding system is complex enough that it requires fairly substantial training and unfortunately does not lend itself to crowd-sourcing or short-term volunteer effort.

The collection has thus far focused almost entirely on printed works; the early phases of the project deliberately postponed the encoding of manuscript materials until we could devote attention to the methodological issues involved. However, we do plan to include manuscript texts by women during the same period. We are currently engaged in an extended pilot to explore these issues. At the conclusion of this pilot, we will probably seek external funding for a broader initiative involving a team of scholarly experts and a wider range of manuscript texts. In addition to English texts written by women, the WWO textbase includes texts co-authored by men; texts of doubtful authorship, where the WWP feels there is good reason to believe the author was female; texts translated into English by women (the original author may be male, although these texts would have a somewhat lower priority); texts written by women in other languages and translated into English by men (again, with a slightly lower priority); historical accounts of trials or other events which claim to report women’s words more or less directly; narratives dictated by women to male transcribers (even where it seems likely that the transcription is not verbatim); texts written under a female pseudonym which have circulated as women’s writing (whether or not the author is actually female; again, with a somewhat lower priority).

The aim behind these choices is to give an inclusive cross-section of the written culture as it circulated at the time, by authors who represented themselves (or were represented) as women, with significant diversity of theme, geography, topicality, genre, format, and audience. For example, allowing the inclusion of dictated texts such as some slave narratives makes it possible to include texts by illiterate women which would be excluded if we insisted upon a strict construction of authorship. Similarly, historical reports of women’s words (for instance, in the context of a witchcraft trial) give a view of women’s discourse which would otherwise be
inaccessible. Texts in translation have circulated within the culture of English women’s writing and represent an important component of that culture. Categories like these need to be distinguished from writing which is straightforwardly “by women” (for instance, for purposes of linguistic comparison), but this can be accomplished by appropriate identification.

**CHOICE OF EDITION**

The WWP always transcribes from a specific copy of an early edition, contemporaneous with the author unless particular circumstances dictate otherwise (e.g. posthumous publication). Where possible and appropriate, we use the first edition. In cases where a later edition is of equal or greater scholarly importance (because of authorial revision, censorship, etc.), we also aim to encode the later edition, although we may not be able to do so immediately.

As a rule, the copy chosen for transcription is generally the only source of information for that transcription; the WWP does not provide a record of variants, emendations, etc. from other copies or editions. However, in cases of illegibility, the transcription may be supplemented with readings from other copies of the same edition. In very rare cases we may consult other editions where there is only a single flawed copy of the chosen edition available. The source of such readings is always explicitly documented. A possible future project (perhaps developed as a specialized exhibit) might be to publish a cluster of texts that survive in significantly varying versions as linked transcriptions, allowing for easy comparison between editions. See Principles of Transcription, below.

Texts which were encoded at the project’s inception were occasionally chosen on other grounds, since the project’s editorial goals have changed somewhat over the years. Transcriptions which do not currently follow the principles outlined above will be updated over time to conform to our current editorial practice.

**PRINCIPLES OF TRANSCRIPTION**

**Treatment of textual variants**

Textual variants from other editions than the one being transcribed are not included except supplementally in cases of illegibility.

**Hyphenation**

Line-end hyphens are preserved. Soft hyphens are distinguished from hard hyphens, and are recorded using the TEI’s mechanism for recording end-of-line hyphenation. They may be displayed or suppressed depending on whether original lineation is expressed or not. In cases where it is unclear whether a line-end hyphen is hard or soft, we follow the hyphenation for that word used elsewhere in the same text; if the word does not appear elsewhere, we record a hard hyphen.
Typographical errors

Typographical errors in the original document are recorded, together with a corrected reading, using TEI’s mechanism for recording error.

Regularization

The WWP regularizes intraword spacing to a single space. We regularize space between words and any following punctuation to zero spaces.

Original typography and spelling

Old-style typography (including the early modern use of i and j, u and v, and vv and w) is preserved, together with a normalized reading, using TEI’s mechanism for handling original readings.

The WWP does not currently record modernized spellings. Doing so poses a number of challenges, not least of which is financial; it would be a very large undertaking and one which would require special funding. In addition, however, there are conceptual challenges such as the frequent difficulty (especially in our oldest texts) of determining the correct modern equivalent. Modernization for many texts is closer to translation than to spelling correction; a word which appears to be a direct modern equivalent may in fact have a rather different meaning, and to substitute it may create a misleading impression of the text — particularly for readers less familiar with early texts, whom modernization is intended to help. Although offering a modernized reading may in some cases make a text more accessible, our experience shows that novices are usually able to adjust, and may even learn more from contact with a less mediated version of the text.

Special characters and Unicode

The advent of Unicode has made it possible to represent nearly all of the printed characters in early modern printed books without difficulty. However, we occasionally encounter characters (such as alchemical symbols or inverted characters) that are not included in the Unicode standard. For these we use the TEI’s mechanism for representing characters not in Unicode (see the TEI Guidelines for more detail).

Handwritten additions and deletions

Handwritten additions which are roughly contemporary with the text are transcribed in full. Deletions are encoded, with the original printed text being transcribed as content. If the deleted text is illegible, that fact is also encoded. The presence of modern handwriting (such as shelf marks or other inscriptions) is represented as a gap using the TEI’s mechanism for capturing omissions.
Illegibility

Illegible passages are captured using the TEI’s mechanisms for representing illegible text; characters or words which are unclear but can be tentatively deciphered are captured using the TEI’s <unclear> element; if multiple readings seem possible, all are captured as alternatives. If the text is completely illegible in the transcribed copy, but legible in another copy that meets the standards for inclusion described above, we may include the legible reading using the TEI’s <supplied> element to document the source. Text not legible in any copy is recorded as a <gap>.

Features omitted from transcription

The WWP’s approach to transcription focuses on the linguistic text, and while we also provide some basic information about non-linguistic features of the text, we do so in a simplified way. There are also a number of features which we do not transcribe. These are omitted largely to enable us to encode more efficiently, and to focus on making more texts available rather than on giving exhaustive visual detail about correspondingly fewer texts. Finally, we understand that no transcription can ever capture strictly visual or graphical detail with sufficient accuracy to replace the original for certain kinds of study. Scholars who need information of this sort will need to consult the original in any case, and for us to attempt to duplicate that information here would be wasteful.

The WWP records the presence of illustrations, together with a brief description of the illustration and a transcription of any text which appears within the illustration. We also encode the presence of ornaments and ruled lines. We do not distinguish between different kinds of ornaments or rules. For our purposes, an illustration is any graphical feature which contains representational content; an ornament is any purely formal or abstract graphic (e.g. a border of acanthus leaves).

The WWP does not transcribe running headers and footers, with the exception of page numbers, signatures, catchwords, and press figures.

The WWP does not transcribe bookplates, modern handwriting, or modern library stamps. The omission of these features is indicated in the transcription using the TEI’s <gap> element.

The WWP does not transcribe smudges, foxing, dead insects, or other non-textual, non-graphical marks. No indication at all is made of their presence, unless they render text illegible or unclear.

TREATMENT OF DOCUMENT RENDITION

With electronic texts, there is a large difference between the information that is recorded and the way the text is displayed (on the screen or in print output). The WWP records a great deal of renditional information, both directly (in a renditional attribute which records many details
of the document’s original presentation) and indirectly (in the use we make of renditional information in deciding what a given textual element is). However, in displaying texts electronically, or in creating printed output, we are guided not only by the document’s original appearance but also by considerations of readability in the new format. Our aim in displaying the document is to present the same information that the original document conveyed — for instance, the presence of paragraphs and stanzas — but without necessarily using the same means of conveying it. Thus while different documents may use indentation or line spacing to show a paragraph break, our standard display regularizes the display of paragraphs to a single style. Similarly, different documents may use varying amounts of space to separate poetic stanzas; our standard display makes it clear where the stanza breaks are, but does not seek to reproduce the exact amount of original spacing. Over time, we are developing additional viewing options that will give the reader some control over these features (for instance, to choose whether to display page and line breaks, or to choose between original and regularized spelling). However, we do not plan to offer anything approaching a facsimile view of the text; the increasing availability of page images makes this less useful, and the level of detail at which our encoding represents formatting information does not support a truly realistic display.

That said, the WWP seeks to capture as many as possible of the meaningful renditional features of the text, for their informational value. By “meaningful” we mean features which affect the reading of the text, where “reading” is understood to mean all aspects of reception, not simply the absorption of strictly denotational meaning. Meaningful renditional features are those which affect the way the reader knows what kind of textual feature she or he is looking at and understands its relationship to other textual features.

Our transcription records most (though not all) significant details of the appearance of the source, including:

- font shifts (roman, italic, blackletter)
- capitalization and use of small capitals
- text alignment with respect to page margins: left, right, center relative indentation
- line, column, and page breaks
- rough positioning on the page (for marginal notes, annotations, and the like)
- end-of-line hyphenation
- wrong-font letters
- turn-unders and turn-overs in verse
- significant use of relative white space to delineate textual structure (stanzas, paragraphing, etc.)
- inverted letters
- rotated text
- underlining
- superscripts
• the presence of dropped, raised, or decorated initial capitals

We do not record:

• absolute or relative type size
• font of punctuation
• specific type face information
• absolute line spacing or vertical white space
• baseline irregularities
• broken type
• running heads (except page numbers)
• kerning and word spacing irregularities (except where these may be significant to the determination of word boundaries)
• swash characters
• ligatures (except digraphs such as æ)

DOCUMENTATION AND METADATA

In addition to transcribing the full text of each document, the WWP also records certain kinds of metadata, or information describing the document and its transcription. This information is recorded in the TEI header. For more detail on TEI headers, see the TEI Guidelines.

Information about the source copy

The WWP records the author’s name (if known), the facts of publication about the original text, the location of the source copy used in our transcription, including library catalogue number where possible, and the Wing or STC number where applicable.

Language

The WWP records the main language of the document and any other languages used in the document.

Genre

The WWP assigns a rough genre classification to each text.

Keywords

The WWP will over a period of time record topical keywords for each document or (in the case of multiple works published together) for each textual unit.
Details of encoding and editing

The WWP records information about the general editorial practices used in preparing the textbase, and also information about the specific practices used for the individual document, if it requires special treatment.
COLLABORATIVEANNOTATION ASSIGNMENT

Description: This assignment provides an opportunity to experiment with different approaches to annotation and contextualization of primary sources. It could be conducted as an in-class exercise over several class sessions, or as an out-of-class assignment (perhaps with the final discussion in class, or presented as an in-class report from each group).

Pedagogical goals:
• Learn to write scholarly annotations on a primary source text
• Work collaboratively to annotate a text
• Work collaboratively to decide on a rationale for annotation
• Learn about the ways in which primary texts are presented and contextualized for a modern readership
• Learn about the relevance of audience and readership to textual presentation
• Gain critical awareness of annotation processes as part of scholarly textual study
• Gain critical awareness of the historicity of texts

In preparation, divide the class into small groups and have each group of students choose a short text (or a few sections of a longer text) to work on. (Some suggestions are given below.) After reading the text through carefully, have each group discuss the following questions in preparation for the annotation process:

• What aspects of the text most need explanation for a modern audience? (For instance, unfamiliar names, references to places and events, unfamiliar words, historical and political background, information about the author’s life, etc.) What kinds of information would a contemporary have had which modern readers no longer possess? What kinds of information might we want to have that a contemporary would not have had access to?
• What would be the most important things to explain for a novice reader? What would most contribute to a productive reading of the text? What do you not need to explain?
• Identify the specific details you plan to comment on in the annotation process, and describe why you chose to focus on these. For instance, if your group decided to identify individuals and events named in the text, explain the rationale for your decision. What kind of reading and research will your annotations support?
• How should your annotations be presented to be most effective? (As footnotes, endnotes, marginal notes, some other format?) What difference does this make to the reader’s experience of the text?

Next, have each group copy and paste the text they’re working with from the WWO site into a wiki, word processor, or other software environment for writing and editing. This will serve as the basis for the new, annotated version of the text. Working as a group (or dividing the text
into sections so that each student can work independently on a separate section), the students should go through the text and add annotations following the rationale developed by the group in the earlier discussion.

When the annotation process is complete, have each group exchange materials with another group (so that each group is now looking at an unfamiliar text). Ask each group to compare the annotated version with the original, and then in discussion consider the following questions:

- How much difference did the annotations make to the comprehensibility of the text? What insights were possible with the annotated version that were not possible with the original?
- What kinds of annotations were most helpful? Which ones were least helpful?
- What is the overall effect of the annotation on the text? How does it alter your impression of the text?
- How did the annotations address you as a reader? What knowledge did they assume you had? Did you feel comfortable in that role?

SUGGESTED TEXTS

The following short texts from the earliest period in the WWO collection are easily divided into manageable segments and also contain ample material for an annotation exercise:

- Margaret Roper, *A Devout Treatise upon the Pater Noster* (opening address from Richarde Hyrde)
- Anne Askew, *The First Examination* (each student take a single page)
- Elizabeth I, *The Accession Speech and Prayer*
- Isabella Whitney, *The Copy of a Letter* (each student take a single poem)
- Isabella Whitney, *A Sweet Nosegay* (each student take a single page)
- Various authors, *A Monument of Matrons* (each student take a single chapter from the main section, starting on page 1)
- Elizabeth Cary, *The Tragedy of Mariam, Fair Queen of Jewry* (each student take a single scene)
- Elizabeth Poole, *A Vision*
- Anne Bradstreet, *The Tenth Muse* (each student take a single poem)
- Elizabeth Grey, *A True Gentlemewoman’s Delight* (each student take a small number of recipes)
- Elizabeth Grey, *A Choice Manual, or Rare and Select Secrets* (each student take a small number of recipes)
**WORD HISTORY EXPLORATION**

**Description:** In this assignment, students take a close look at how word meanings, context, and usage change over time. This assignment could be conducted as a series of two in-class sessions (in a computer-enabled classroom) or as an assignment with an in-class and a take-home component. It could also be conducted, with slight modification, as a group exercise for the entire class, with the instructor projecting the WWO interface from his/her own computer and students guiding the search and discussing the results.

**Pedagogical goals:**
- Learn to use WWO interface features to gather information about word usage
- Work collaboratively to frame a research problem
- Gain an increased understanding of how language changes over time
- Gain an increased understanding of how cultural themes are expressed in language

In preparation for the assignment, have each student identify a theme or topic of special interest (preferably in advance). These could arise from class discussion, or could be selected from a set of central topic areas associated with the course. Students could also work in small groups, each one focusing on a different theme or topic.

Next, ask each student or group to come up with a list of half a dozen words that are most closely associated with their theme. For instance, if the theme is “authority”, related words might be things like “power”, “rule”, “ruler”, “leader”, “authority”, and so forth. Have the students look up the current meaning of each word in a modern dictionary, and note any current range of meaning.

Next, have students search for each word in turn. For each word, have them look at how the word is distributed by genre (using the genre facets under the "Filter" heading on the left of the screen) and by time period (using the date facets, or using the vertical timeline ribbon). If the word only appears in later periods, consider whether there are alternate spellings that should be explored. Then ask them to look through the list of search results for examples of how each word is used in specific genres and time periods.

Looking at these results, ask the students to consider the following questions as a group:

- What is the earliest usage of this word in the collection, and what can you tell about its meaning from the surrounding text? Is the meaning very different from modern usage, or essentially the same?
- What is the latest usage of the word in the collection? Is its usage the same as modern usage, or different?
• Looking at the results in between, how many different meanings of the word can you identify? In what contexts is it most often used? For instance, does it appear chiefly in texts on certain subjects in particular (religious, political, domestic, etc.)? Does it appear in any specific genre more than others (poetry, fiction, drama, etc.)? Looking within the text, is the word associated with particular contexts such as dialogue or footnotes? Is it associated with speakers of one gender more than another?

• Can you identify any key points in the word’s history where its meaning seems to shift significantly?

After repeating this process with several words from a specific theme, ask students to consider the following broader questions:

• Which of the terms you initially identified seem most importantly associated with your theme in the texts represented here? Did any of your initial terms not appear at all?

• How has the language associated with your theme changed over time?

• What might these changes reveal about the significance of your theme for understanding women’s writing?

This project can be given a more specific focus by narrowing the chronological period; to do this, include a date range in the initial search, using the bibliographic search fields.

If your class has access to a historical dictionary like the OED, it could also be interesting to look up the words there and see whether the WWO texts’ usage is consonant with the attestations given in the dictionary.