ENG 39995: Digital Humanities Project Assignment

As noted in the "Digital Humanities Projects" section of the syllabus, the overarching criterion for evaluating a scholarly editing project is that "scholarly editions make clear what they promise and keep their promises" (MLA Committee on Scholarly Editions 23). This criterion—which I sometimes describe in paraphrase as "Say what you do, and do what you say"—is equally apt for digital humanities projects. Again according to the "Guidelines for Scholarly Editors," the end product (a scholarly edition) is evaluated according to five criteria: 1) accuracy, 2) adequacy, 3) appropriateness, 3) consistency, and 5) explicitness. See "Sources for Paragraph 1" below.

Because the MLA CSE Guidelines are intended for review of overall project at time of its completion, a portion of my evaluation shall consider other criteria during the process: project planning, scheduling, and group collaboration. At the project's conclusion, I shall review project in the following categories: introduction, transcription, textual notes, annotative notes (optional), images, and TEI-encoded text. I will collaborate with you on the publication of TEI-encoded text with the Drupal plug-in TEICHI. See "Sources for Paragraph 2" below.

The grade for your project will be based the following criteria:

- Proposal (10%)
- Interim Project Contributions (10%)
- Biographical and Historical Introduction to Alfred Chester (10%)
- Textual Introduction (10%)
- Accuracy of Transcription (10%)
- Images (10%): (*Note*: Reduced percentage if prepared by Gilgenbach, with half each [5% moves to Transcription and 5% to Textual Notes])
- Textual and Annotation Notes (10%)
- TEI XML-Encoded Text (10%)
- Peer Evaluation of Collaboration (10%)
- Short Reflective Project Essay (10%)

Proposal

The Proposal Draft is due on March 9.

The proposal should be a collaborative project among all contributors. I recommend that you write the proposal in Markdown and use GitHub to coordinate your work. The final draft should be submitted as a print copy and posted on the class blog. Though full sentences are important for introduction matter, parts of the proposal may be more effective as a series of bullet points, or a table. The proposal should be a 3–4 page document that addresses the following matters:

- Draft schedule for entire project (timeline), with deadlines for individual tasks that contribute to the project
- For each task listed on proposal, an assignee (or assignees) who are responsible for completing that task by the deadline. Please keep in mind my interim project deadlines on syllabus when assigning these tasks.
- For each task listed on proposal, draft principles for how that task should be performed.

The principles and procedures for the following tasks should be addressed in some detail on the proposal:
Schedule and Assigned Tasks

Please list tasks that need to be completed. Please list deadlines by which each task will need to be completed for the project to be complete by the end of the semester. Please list person (or persons) responsible for each task.

Information that you will need to complete this task include the number of letters, an estimate of the length of each letters. For proposal, the schedule and assigned tasks should include contributions from all three students.

This schedule should be re-evaluated over the remainder of the semester, and the inability to meet deadlines (whether because personal matters, whether because of unanticipated technical difficulties, or whether because initial estimates were unreasonable) should prompt you to reconsider range of tasks performed.
Biographical and Historical Introduction

Who is Alfred Chester? Who are his parents and siblings? Where did he live? What are his dates of birth and death? What did he write? Where are his papers? With whom and where did he live? With whom did he correspond? Who are his well-known contemporaries? When researching information on Chester, you should consult standard reference works for writers, such as *Dictionary of Literary Biography* and *American National Biography*. Consult the reference entries that have entries on Chester. In addition, consult the Kent State catalog entry and WorldCat. With basic information established, other valuable sources would include Ancestry.com (see Kent State library list of databases). Are Chester's publications collected? If Chester is identified with any literary movements, can you identify details about those with reference entries in *Oxford Bibliographies*? To identify additional references for Chester, consult Harner's *Literary Research Guide* and the American Library Association *Guide to Reference* at `http://guidetoreference.com`.

For proposal, provide brief bibliography of major reference sources and plans for acquiring other biographical information. Please write a 100-word introduction and list of a minimum of 4 reference sources and a brief note for each on what type of information that source provides.

Textual Introduction

For your final project draft, you should provide a textual introduction with the following:

- Explains the principles for choosing a set of documents to edit
- Explains the principles for transcribing the documents
- Explains the principles for proofreading the documents
- Explains the principles for encoding the text
- Explains the principles for annotating the text
- Explains the methods for acquiring and rendering images for user
- Explains the technology used for publication

For proposal, you should decide who will write the textual introduction and what sources will be consulted to identify principles and methods.

Transcription Methods and Principles

Below are some guidelines for transcription:

- No silent emendations during transcription.
- Do not expand abbreviations during transcription.
- Note alterations in hand: ink color, pencil, typewriter.
- Note written-over corrections in square brackets: `[this over that]`
- Place comments and interpolations in square brackets: mis-spelled words, items in need of annotation, etc.

Items in square brackets (notes to yourself) will need to be dealt with systematically after the transcription is complete and corrected. You may want to use regular expressions to replace square brackets in source text with comment elements in XML. Some square brackets may need to be encoded in XML as draft notes. Finally, if there are multiple versions of the text (uncorrected or corrected), you may need to encode corrections or multiple versions using witness tags and notes. The eventual choices that you make on these matters will depend on editorial principles chosen in textual introduction. But the most systematic way to assess potential issues is during transcription.

How shall the project communicate to Cara Gilgenbach our needs?

Below are some principles for proofreading:

- The usual guideline for scholarly editing is a total of 3 silent proofreading passes. A double-keying with digital file comparison is equivalent to two proofreading passes.
- Whether proofreading by double-key file comparison, by oral reading, or by silent reading, print document and mark errors for correction.
- In oral proofreading, read punctuation and printing marks aloud.
- When you find an error in a line, mark it. Then, proofread the entire line again, from the beginning.
- If original documents are avaialble to you, at least one proofreading pass should be against original document.

Please draft preliminary choices for acquiring text, with the following items addressed:

- How will you transcribe? From original documents? From photocopies? From digital facsimiles? (**Tip:** When transcribing, I recommend that you transcribe line by line and use a plain text editor set to UTF-8 encoding.)
- Will you use optical character recognition (OCR) to acquire text? Which software?
- Will you use oral or side-by-side proofreading (original or facsimile copy alongside transcription) to mark the transcription for correction?
- If you will double-key the text and compare the two transcriptions, which software will you use? (**Tip:** XCode FileMerge, JUXTA, other?)
- When you have a transcription that needs to be proofread for correction, how will you mark errors in the text? And how will you correct the errors and ensure that corrections are complete and that you do not introduce other errors during corection? (**Hint:** Keep 2 versions of file, uncorrected and corrected: after initial pass at corrections, compare with FileMerge to ensure all marked corrections are made.)
- How will you name files? (**Tip**: I recommend naming each Chester letter by its date.)
- How will you manage multiple versions of each file: 1) initial transcription, 2) second transcription, 3) initial OCR version, 4) corrected transcription. (**Tip**: Do not just copy over original file with corrected file. Devise a file naming convention for each type of file.)

Image Acquisition and Processing
Please address the following:

- Who shall be responsible for acquiring images?
- Which hardware and software packages will be used to acquire images? To change image format from, say, TIFF to XML?
- How long will it take to acquire images? Count the number of images, and estimate number of minutes to acquire a single image, the entire process. Pull from folder, scan, save digital file, replace in folder.
- Assuming approximately 1.5 hour image capture sessions with a 10-minute setup and clean-up, how many image-capture sessions would be necessary to capture all images?
- At archival standards, 600 dpi and 24-bit color, what is the approximate file size of each image? What is the amount of space that will be needed to store all images?
- What method (software, settings) will be used to convert high-resolution archival TIFF images to display resolution JPEG images?
- What file naming convention shall be used for images?

Textual Notes and Annotation Notes
How shall the following be recorded and/or displayed in your edition?

- **An authorial error, uncorrected text displayed:** If you wish to display an uncorrected reading in case of an obvious error, will you encode both versions and display only uncorrected version? Or will you not mark the uncorrected version? Note choice in your textual introduction.
- **An authorial error, editorial correction displayed:** If you wish to display a corrected reading in case of an obvious error, will you encode both versions and display only corrected version? Note choice in your textual introduction.
- **An authorial error, both versions displayed:** Does your display system TEICHI offer an option to display both the uncorrected and corrected text, in a display option or with a note?
- **An authorial abbreviation:** Do you wish to display an expanded reading or the original authorial abbreviation? Will you encode both versions and display only original abbreviation or expanded version? How will reader be notified of presence of abbreviation or an expanded abbreviation, editorial note or display option? Note choice in your textual introduction.
- **An authorial correction or interpolation** If author corrects during initial transcription (overwrites or makes an editorial interpolation that corrects an error), will you encode both versions and display only original assay or corrected version? How will you notify reader? Editorial note or display option? Note choice in your textual introduction.
- **An authorial revision:** If author revises, will you encode both versions and display only original draft or revised draft? How will you notify reader? Editorial note or display option? Note choice in your textual introduction.

- **Multiple hands:** If there are multiple written hands on the document (author, author's friend, letter recipient, librarian mark), will you encode non-authorial hands? How will you notify reader? Editorial note or display option? Note choice in your textual introduction.
- **Explanatory Annotation:** What types of annotation are needed to aid assumed reader? Some to consider include the following: proper names, geographical places, publications, unfamiliar words, foreign words. What you annotate in addition to these will depend on characteristics of individual documents. After a review of documents, can you identify classes of annotation that seem necessary for your assumed reader?

Please note that it is not necessary to encode, devise ambitious display options, and serve all these possibilities. You can treat some or many with editorial policies that include silent authorial correction, silent editorial correction, silent expansion of abbreviations, and ignoring alternate hands. If you provide, for example, a digital image, you may choose to only provide a reading text for your user. But these choices (which you include, which you exclude)should be addressed in your textual introduction, regardless of what you choose. Also, you may "encode" but choose not to display. In that case, you may want to allow some option for user to download encoded text.

XML Encoding
What procedures shall you follow when preparing XML encoding?

- If initial transcription is made in plain text, how will you transfer corrected text into XML document?
- If initial transcription includes simple marks like square brackets to designate items to be annotated, abbreviations, hand (typed, ink color), authorial corrections or interpolations, how will you encode those items in TEI XML, from among following choices: 1) update transcription encoding so encoded formally according to TEI in published document, 2) update transcription encoding to XML comments, which do not display, 3) Strip out initial encoding, and leave out of finished version. (**Tip:** Regular Expressions are your best bet here.)
- What type of XML encoding does the publication environment TEICHI support for publication? Does it support the full range of TEI tags? Or only a limited set?
- Is a desirable encoding practice too difficult to realize in XML? Should you reconsider editorial standards because of limitations of encoding?

Sources for Paragraph 1

The advice and criteria on scholarly editing are defined in greater detail in *Electronic Textual Editing* (see pgs. 23–46 in print textbook), on the web site of the MLA CSE (see `http://www.mla.org/resources/documents/rep_scholarly/cse_guidelines`), and on the preview version of our textbook that is hosted by TEI Consortium (see `http://www.tei-c.org/About/Archive_new/ETE/Preview/#body.1_div.2`). I encourage you to consult any one of these versions of the "Guidelines for Scholarly Editors," when planning the project. These general principles will be applied during my evaluation of any aspect of the project.

Sources for Paragraph 2

For the portion on planning and scheduling, my guidelines are based on the National Endowment for Humanities Scholary Editions and Translations Grant Guidelines (see `http://www.neh.gov/files/grants/scholarly-editions-dec-9-2014.pdf`). Unlike for previous paragraph, I do not recommend that you consult that guideline document—I only acknowledge its influence on this document. The evaluation of queries and coordination is based on my own experience with multiple digital projects and our class work with coordination tools (including GitHub). Criteria for evaluation also have been influenced by the online guidelins from Professor Ann R. Hawkins at Texas Tech University (see `http://hawkins.writingstore.com/index.php?option=com_content&view=article&id=133&Itemid=105`).