Assignment 1

Wesley Raabe

July 10, 2014

About "Digital Humanities: What is it?"

Our first course project is a collaboratively authored document entitled, "A Student Project: What is Digital Humanities?" The surface rhetorical purpose of this assignment is to engage with one another in regard to public definitions of digital humanities. All students will contribute to that effort, first, by reading assigned articles and blog posts, by writing a blog post, and by responding to one or two posts by other students. After the initial discussion is offered and revised in response to recommendations from other readers in the class (blog comments), the class's second effort will engage in a collaborative process to curate contributions and sharing the discussion with the public.

Because the public may consist of different types of readers, we will curate the content and present it in multiple formats, which may appeal to different readers. Those formats are a public blog, a printed document, and an ebook. However, to make the process more manageable, we shall break the class into two teams, each of which will curate half of the posts. You will both be challenged to use free open-access tools that can make this collaborative project workable—Markdown language, GitHub repository, Pandoc document converter, Zotero, WordPress blog software, LaTeX, or Calibre—and you shall be encouraged to think about the degree to which choosing a publication form (a blog, a print document, an ebook) should shape the way a document is created and published.

Is there such a thing as one ring to rule them all? Can students form a temporary fellowship in which Markdown, the language for simple document formatting, and Pandoc, the conversion tool, manage the magic ring effectively? More prosaically, can one team of work effectively to publish an extended collaborative document in multiple forms without walking all over one another. Furthermore, it will be necessary for each of us to see whether we can distinguish our own personal frustrations with learning a new technology from the limitations that are characteristic of the technology or the publication forms. Do some publication technologies lend themselves to certain types of document features (citation, links, keywords, commenting) that are more difficult or perhaps even pointless when presenting the "same" content with another publication technology? Many

advocates of technological solutions advocate a distinction between content (the ideas, the words, rhetorical markings) and form (conventions of typography, spelling), and this distinction is almost gospel among technology evangelists, who argue that the essential advances in document management that undergird technologies like Markdown and eXtensible Markup Language, which we will look at later in the course, are unthinkable without a powerful distinction between the essential content (which should stay the same) and the accidental forms (which can be manipulated as needed). Is a distinction between form (accident of being published online or in print form) and content (essential part of words and ideas) liberating and useful or constricting and misleading?

I should announce that the my choice to require you to use open access tools instead of commercial software reflects multiple factors: ideological leanings, pedagogical usefulness, and, hopefully, practical benefit for your future employment. In the ideological sense, digital humanities encourages an ethic of public sharing and collaboration. By asking you to use public software and work together, I encourage you to explore how this can be beneficial to your work. For example, when frustrated with aspects of writing this assignment, my Twitter colleagues Caleb McDaniel <https://twitter.com/wcaleb> and Lincoln Mullen <https://twitter.com/lincolnmullen> provided useful guidance. You should turn to fellow members of your team, to me, and to online guides also when you encounter technical difficulties. Now, for example, you could collaborate and put together most of same document formats with a combination of DropBox, Google Drive, Microsoft Word, and Scrivener. But it is more in keeping with the ethic of digital humanities to assume that the initial challenge of learning these open access tools will prove rewarding, that you will encounter alternate ways of thinking about what documents are. Instead of deciding that you need to pay a software vendor, decide to learn something new—an alternate way of working. The open access tools tend to force you to put yourself into control of the documents and not rely on software vendor choices. And everything is cheap, usually free. Finally, I worked as a technical writer for over a decade, so I can with some confidence tell you that you that your ability to adapt to alternate corporate cultures and documentation methods has the potential to improve your employment opportunities. You already know either Microsoft Office or Google Docs, so this is the opportunity to expand your personal toolkit. To know various methods of document management—and to be able to make recommendations or adapt either as a documentation manager or project manager—is a far more valuable skill than knowing the one software program that a company or organizations currently uses to produce documents.

Shared Processes, Week 3

Each student's individual contribution to this assignment will consist of the following parts:

1. Each student will offer a formal response to one or more of the readings

during the first two weeks of the course. The response will consist of a 2-to 3-page (500 words) blog post, with a thesis that engages one or more of the following general topics:

- digital humanities as a threat to literary studies (Kirsch; Marche);
- digital humanities as a set of institutional sites of study and professional organizations (Kirschenbaum);
- digital humanities as building things, such as computer applications, models, and online projects (Ramsay);
- digital humanities as sharing, developing a community (Sample);
- digital humanities as productive unease, grappling with tension between ideas of technological progress and the recognition that humanities research materials resist modeling and processing (Flanders).

2. Each blog post must have at least one formal citation of another piece of writing, either from list above or from another blog post or article.

3. By Monday of week 3, each student's original blog post, which shall be written using Markdown language, will be posted on the class's Markdown-enabled blog AND shared on the class's GitHub repository.

4. By Wednesday of week 3, each student shall respond to two other students' blog posts. Each response must refer to a specific passage or idea in the student's post and compare or contrast with another idea from another student's post or from the assigned reading (i.e., quote from and cite source). The response posts should be posted on the Markdown-enabled blog.

Splitting Up Labor, Week 4

After initial posts are complete, the class shall collaborate to produce posts in multiple forms. Editorial work shall be collaborative, and one member of class will take responsibility for one of the following processes: 1) print document, 2) public blog post, and 3) ebook. You are encouraged to collaborate with other members of the class to manage or troubleshoot any difficulties that you encounter. Whenever possible, please work together (in IBE, in reserved library room, at same computer, with online collaboration tool, or with professor) to achieve solutions to problems.

1. During week 3, each student will contribute GitHub fork for team members' original blog posts. Based on commentary received on the blog, posts should be corrected and revised.. Furthermore, the editorial group by end of week shall organize the student contributions (section headings and introductions, shared bibliography), and shall provided editorial approve of the final project for publication. When the other teams (print, public

blog, ebook) do their work during week 4, the editorial group shall remain responsive and accommodate revisions to GitHub repository documents to better serve needs of print, blog, or ebook team.

2. During week 4, the person responsible for print document shall use Pandoc or Pandoc and LaTeX to convert blog posts from its team members to a printed LaTeX document (report or article in PDF format) that gathers all student contributions and has a shared bibliography for all citations. The final product should be based on the most current and updated GitHub fork.

3. During week 4, the person responsible for public blog shall develop an independent public blog of all group posts with keywords and citations. The final product should be based on the most current and updated GitHub fork.

4. During week 4, the person responsible for the ebook shall use Pandoc or Pandoc and Calibre to generate an ebook in ePub format and (optional) publish the book on Google Books. The final product should be based on the most current and updated GitHub fork.

At every stage, students are encouraged to raise the bar above minimal assignment requirements. Ask yourself, for example, would post or printed document benefit from a table of contents or index? Should blog posts have keywords? Should contributor notes be published also? To earn full credit for excellence, you need to raise the bar.

General Notes: I encourage an iterative process. Test procedures to see if they work. Try to do everything, but develop reasoned processes to scale back ambitions if it becomes overwhelming. The final step of the process will be a personal reflection on collaboration in the assignment (500 words), which will be posted on the course blog. The groups that are responsible for individual publication formats may make recommendations to alter source documents. The editorial team is responsible for assessing needs of other teams, coordinating any document changes, and reviewing any proposed changes.

Editorial Working Group

Your group is responsible for proper grammar and spelling and consistency of citation and formatting. Revise and mark up individual posts for revision. Maintain control over the source Markdown documents in GitHub. Review and return posts with editorial markup by Tuesday.

I recommend marking up individual markdown documents with invisible comments. But after individual edits and suggestions are made, notify the writer who authored the post to review the edits. After writers make their edits (preferable on GitHub source), it is your final review that provides consistency of formatting style and editorial style.

You should install Zotero, the software on which you will create entries for all cited materials https://www.zotero.org/. One of the features of Zotero is shared public library. And I have placed the initial references for "What Is DH?" assignment at the following location: https://www.zotero.org/groups/274025. If not accessible, send me a reminder. I will open group membership in Zotero library to all students.

The job of the editorial working group is to ensure that all are cited consistently and that no citations are duplicated. When the Zotero library is consistent, other working groups will need the .bib file from the Zotero library to create a bibliography. To export the bibliography, highlight all entries in "WhatisDH" folder, right-click (PC) or control-click (Mac) to launch menu and click "Export Items." When selecting format, choose "BibLaTeX." Place the bibliography (.bib) in the GitHub repository with markdown documents from blog.

Print Working Group

For the print working group, you must use Pandoc and LaTeX to generate a PDF and a printed document. The print working group shall use all approved editorial posts to create a print document. The initial method for creating the document shall be as follows:

1. Install Pandoc and LaTeX for your operating system.

2. After the editorial working group's work is complete, use GitHub to download the final updates of blog posts and bibliography (but do test with sample date from GitHub before attempting the final version).

3. Review the following routines for managing elaborate Pandoc documents (multiple files with citations).

   - Managing Multiple files in markdown. https://github.com/akmassey/markdown-multiple-files-example

   - Adding name of the .bib file to YAML metadata block. http://johnmacfarlane.net/pandoc/demo/example19/YAML-metadata-block.html
   - Setting up appropriate Citation Style Language in GitHub, which you add to Pandoc file directory. https://github.com/citation-style-language/styles. The Chicago style names begin with "chicago-". The Modern Language Association style names begin with "modern-language-". Place them in main document file folder for convenience.
   - Download the `default.latex` template, and place it in the main document file folder for convenience. https://github.com/jgm/pandoc-templates/blob/master/default.latex.

- Review the `pandoc-cite-proc` extension documentation if you wish to generate bibliography entries through Pandoc http://johnmacfarlane.net/pandoc/demo/example19/Citations.html.

After the setup is complete, see the Pandoc documentation for instructions on generating LaTeX document from Markdown with bibliography. You essentially have two choices:

- Use the `pandoc-cite-proc` extension to generate the bibliography simultaneously with one Pandoc command. If you do this, then you will need to alter the `default.latex` template to change the output format. Pandoc embeds the bibliographical information in the document as text (from the .bib source). I have not spent much effort learning how to alter the `default.latex` template.
- Do not use the `pandoc-cite-proc` extension. Instead, you may edit the generated LaTeX document. After you adjust the format of the LaTeX document how you wish (narrower margins, perhaps) you may generate the bibliography with BibLaTeX (from the .bib source). To generate a bibliography from within LaTeX, you must run LaTeX, then run BibTeX, and then run LaTeX twice more. Many LaTeX editors have macros or commands so that you can run them from toolbar. You can also do it from the command line. To generate a bibliography from within LaTeX, you must run LaTeX, then run BibTeX, and then run LaTeX twice more. I can help you with this, as I know LaTeX quite well.

Below are some other recommendations recommendations to consider:

- If you want to manually alter the LaTeX document, one dissatisfaction that new users often feel is with margins. For information on page layout, see http://en.wikibooks.org/wiki/LaTeX/Page_Layout.
- After Pandoc generates your LaTeX document, you may want to consider uploading the document to an shared LaTeX editor, on which you can collaborate for your LaTeX design. The two that I would recommend you consider are http://www.authorea.com and https://www.sharelatex.com/. Do note that default for both these tools is to only allow one collaborator at a time for free accounts, though Authorea says it allows two.

Public Blog Working Group

For the public blog working group, you must set up a WordPress installation with a Markdown plugin. I am using Reclaim Hosting, and we can use the shared class blog. The person who is managing the blog should contact me, and we will work together on the setup.

1. (Administrator): Install WordPress on the domain, http://portal.reclaimhosting.com/knowledgebase.php?action=displayarticle&id=2.
2. (Administrator): In consultation with fellow members of your working group, create a URL and title for blog that offers approximate representation of assignment question (that is, no "adorablekittens.com").
3. (Administrator): Invite other students to your blog as administrators. On Users menu, click "Add Users," and invite other class members via their official Kent State address. For Role, select "Administrator."

Students who are invited in Step 3 are set up as administrators of the WordPress account. Subsequent steps may be performed by anyone in work group who has an Administrator account in WordPress, so all subsequent steps should be performed collaboratively.

1. When logged into WordPress as an Administrator, select Plugins menu and search for "JetPack." Install it.

2. After Jetpack appears in your list of plugins, click "Activate."
3. On the Jetpack menu item, click "Get Started" and "Authorize Jetpack."
4. Create a test post with Markdown language (you may need to log out and log back into the blog after JetPack is installed). For details on the flavor of Markdown that is supported, see <http://jetpack.me/support/markdown/>.
5. When GitHub repository is updated with revised blog posts, develop a procedure to publish posts on your blog. Please attempt Markdown, but you have the option to decide against Markdown should it prove unworkable (uninstall JetPack).

Below are some design recommendations to consider:

- What should home page of blog look like? Should certain posts be more prominent than others–such as most recent first? Or should you select a design that places equal weight regardless of when post published? Review difference between behavior of "page" and "post".
- Should documentation (works cited) be associated with individual posts, or should entire blog have a shared bibliography? Which makes more sense to user, Chicago style notes and bibliography or MLA-style parenthetical references? Can you use Pandoc to generate bibliography or notes? Is Markdown the best way to format posts, or could you use Pandoc to generate another format, like HTML, that would work better?
- For student posts that refer to other student posts on the same blog, can you devise a standard system for cross-referencing between posts, such as a naming convention for blogs that can be used also in cross-references?
- Could your blog be better with information in the About section or other widgets? Might you promote blog on Twitter or Facebook?

- Would tags or keywords offer a better means of sorting posts into groups?
- Would blog posts be better served with individual bibliographies for each post or a shared bibliography for entire blog.

EBook Working Group

As you will read on <http://chronicle.com/blogs/profhacker/make-your-own-e-books-with-pandoc/39067> from Lincoln Mullen, creating an eBook is "almost trivially easy." Mullen, I should note, does not include a bibliography in his example. Therefore, as the ebook working group, you have the special responsibility of improving the quality of the final product beyond the source documents.

Below are the basic steps that Mullen recommends, revised to emphasize our class documents.

1. Download the updated GitHub versions of our class blog posts.
2. Review the following routines for managing elaborate Pandoc documents (multiple files with citations): managing Multiple files in markdown. <https://github.com/akmassey/markdown-multiple-files-example>.
3. Create a title page named `title.txt` and a copyright page named `metadata.xml`.
4. Review the Pandoc options to handle bibliography entries and citation style.

- Review the `pandoc-cite-proc` extension documentation to generate bibliography entries through Pandoc http://johnmacfarlane.net/pandoc/demo/example19/Citations.html.
  - Review adding name of the .bib file to YAML metadata block. <http://johnmacfarlane.net/pandoc/demo/example19/YAML-metadata-block.html>.
  - Review setting up appropriate Citation Style Language in GitHub: place `.csl` file in Pandoc file directory. <https://github.com/citation-style-language/styles>. The Chicago style names begin with "chicago-". The Modern Language Association style names begin with "modern-language-".

5. Using Pandoc, stitch all the document parts together, including citation entries with the `pandoc-cite-proc`.
6. Inspect the book in Calibre.
7. Publish the ebook on a free service, such as Google Books. See <https://www.lib.umn.edu/faq/22714>

A helpful tutorial on ebook options with Pandoc and Calibre is available at <http://puppetlabs.com/blog/automated-ebook-generation-convert-markdown-epub-mobi-pandoc-kindlegen>.

However, you should carefully consult the Pandoc documentation as well, which lists all options. See the basic tutorial at <http://johnmacfarlane.net/pandoc/epub.html> and the advanced list of options in the user's guide http://johnmacfarlane.net/pandoc/README.html>

Below are some design recommendations to consider:

- Should your book have embedded fonts?
- Should your book have a cover page and a TOC?
- Is the default style acceptable, or should you consider a .css file to make the format more pleasing?

A series of more technical methods for creating ebooks is available from IBM on its developer network at <http://www.ibm.com/developerworks/xml/tutorials/x-epubtut/index.html?ca=drs->. Though I recommend that you use Pandoc, the IBM documentation may help you to understand other customization options.

Bibliography

Flanders, Julia. "The Productive Unease of 21st-Century Digital Scholarship." *Digital Humanities Quarterly* 3.3 (2009): n. pag. Web. 9 July 2014.

Kirsch, Adam. "Technology Is Taking over English Departments." *The New Republic* (2014): n. pag. Web. 9 July 2014.

Kirschenbaum, Matthew. "What Is Digital Humanities and What's It Doing in English Departments?" *ADE Bulletin* 150.2010 (2010): 1–7. Web. 27 Oct. 2013.

Marche, Stephen. "Literature Is Not Data: Against Digital Humanities." *The Los Angeles Review of Books* (2014): n. pag. Web. 9 July 2014.

Ramsay, Stephen. "On Building." *Stephen Ramsay* Jan. 2011. Web. 26 Oct. 2013.

Sample, Mark. "The Digital Humanities Is Not About Building, It's About Sharing." *Sample Reality* May 2011. Web. 26 Oct. 2013.