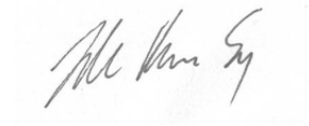


Lloyd Sy

Digital Humanities Dossier

Practicum Statement

I certify that I have completed all 60 practicum hours required for the certificate.

A handwritten signature in black ink, appearing to read 'Lloyd Kevin Sy', is written over a light blue grid background.

Lloyd Kevin Sy

Practicum Details

I am enrolled in the DH Practicum Course (DH8000) in the Spring 2023 semester.

Information about the various projects I worked on during my time at UVA can be found in the rest of this portfolio. An estimate of the amount of time I have worked on each, however, is below:

- Collective Biographies of Women: approximately 750 hours over five years as RA/Project Manager (Fall 2017 - Spring 2023)
- For Better For Verse: approximately 50 hours over the course of one semester (Fall 2017)
- Wikipedia, School of Data Science: approximately 100 hours over the course of two semesters (Spring 2022 - Fall 2022)
- Direct Discourse Lab: approximately 100 hours over the course of several years, especially Fall 2022 - Spring 2023

These total to approximately 1000 hours over the course of my graduate career. Beyond those projects, I have attended GIS workshops and various DH mixers. I also gave a talk about my DH autobiography in Fall 2021. Several times throughout my program, I have also been called on to teach in different classes, instructing students in the grad course Women Writers on XML and coming into the DH core course to discuss my experience. Outside of those projects, I have probably spent some 25 hours at other DH events.

Elective Coursework

As I started my doctoral program in 2017, the DH certificate was formed while I was mostly out of coursework. I am fulfilling my coursework requirement through two classes from the bookends of my years here—an eighteenth-century fiction class in which I performed computational text analysis, and an independent study in Spring 2023.

ENEC8600

This course, taken with Professor Cindy Wall in the English department, looked at broad questions of space and time in the eighteenth-century novel. The final project for the class was a traditional, 10-12 page essay.

My final project was born when I read Jane Austen's *Northanger Abbey* and recognized that Catherine Morland, its protagonist, was an extremely terse speaker. I found that, in approximately 95 conversations throughout the novels where she is involved, she is only the primary interlocutor (the person who speaks the "most" during the conversation) in seven of them. I found this by doing a rudimentary first step towards a more robust computational text

analysis: I made an Excel spreadsheet and kept track of Catherine's conversations per their chapter, her interlocutor, and the amount of words each spoke. The resulting paper commented upon conversation in the novel generally, while profusely drawing on data from the spreadsheet. For instance, I read each of the seven conversations in which she *did* speak more and pointed to the data behind them to point to their *exceptionality*. Read as a series of these exceptional conversations, we can see Catherine's oft-noted critical development as a character as reflected quantitatively, by word counts. I also wrote a Python script to more thoroughly sift through these conversations, though did not find a way to incorporate the script's findings into the final paper. The paper (and my overall performance in the class) yielded an A+.

Most importantly, the essay became the immediate springboard for launching a more extensive dataset of *all* conversations in Jane Austen's novels, which formed the basis of my Direct Discourse Lab. Over the course of the next few years, I built a database in my spare time, eventually marking up nearly 7000 speech acts in her six major novels. Down the line, this set up my next elective course for the certificate.

ENGL 8993

In the Spring 2023 semester, I completed an independent study to further develop and bolster the aforementioned digital project, the Direct Discourse Lab. By this point, I had built up the apparatus for computational textual analysis in several ways. As just mentioned, the database of Austen's conversations was heavily augmented by 2022. During that year, I more formally created the lab by finding three undergraduate students to serve as research assistants. Two of them—Neeka Samimi and Maddie Stokes—were connected to me through the Jefferson Foundation's research assistant program; the other, Alexa Kamm, was my former student. While

the former two worked on performing computational text analyses—querying the database and writing up thoughts about collections of quotes from different speech acts—Alexa worked exclusively on a Heroku-platformed, django-based website that I had built up over the last few years. When Heroku updated to a new stack in the spring of 2023, the website broke. I mention this fact because it represents a fundamental—but difficult—step in my understanding of DH projects: they are messy, and upkeep, especially in the long term, requires extremely careful consideration.

Nonetheless, the most important part of the project—the relational dataset—lives on, available in a .dump file on Github and easily queryable by any SQL query tool. We are composing a paper in which we perform a Naive Bayesian classification on the quotes. Put briefly, the essay tries to determine what’s going on when Austen uses the verb “said” vs. when she uses the verb “cried.”

In my independent study, Alison (an expert in nineteenth-century British literature) worked closely to hone this essay into a potential article. Beyond the sheer intellectual excitement of performing this task was the more contextual excitement—collaborating with co-authors is a new experience for me, but is extremely common throughout the sciences. Thus among other things the project hopes to investigate the possibility of a science, lab-based model in the humanities setting, in which people make modular, well-defined contributions to a project, and a primary investigator (myself, in this case) ensures that the parts fit together.

Reflections on DH Community

As I always tell people, my digital humanities experience at UVA was one of the most formative intellectual experiences in my life—how can anything that someone spends 1000

hours doing not? It was also an important personal experience, in some ways, not least for reasons laid out in this document. Above all, in the extremely confining world of the ivory tower, the collaborative nature of digital humanities taught me how our research might be distributed, might arise out of merging different strengths and energies.

One incredible consequence of this collaborative nature is the interdisciplinarity that my work has taken on: I've worked closely with students from fields as diverse as History, Computer Science, and Nursing. United by a common methodology, I find that we have some of the most interesting conversations I have had at UVA.