Courtney Rivard

Turning Archives into Data: Archival Rhetorics and Digital Literacy in the Composition Classroom

Using assignments drawn from a first-year composition course that centers the Southern Life Histories Collection, part of the New Deal’s Federal Writers’ Project, this paper argues for a pedagogical approach that teaches students digital literacy through archival rhetorics by converting archival texts into data.

The people, all the people, must be known, they must be heard. Somehow they must be given representation, somehow they must be given voice and allowed to speak, in their essential character . . . life histories can help with this job.

—W. T. Couch, These Are Our Lives

Grappiling with how to construct metadata categories from archival records, Jack Turner, a student in my first-year composition course (Writing in the Digital Humanities) posted this insightful comment on his weekly writing insights, a blog-style online comment: “What if I choose one set of data at the expense of another set of data, which I for some reason think is less important? Can I be trusted to do this, with all my biases (some of which I might not even be aware of)? Should I trust others to do this? Do I even

Copyright © 2019 by the National Council of Teachers of English. All rights reserved.
have a choice? . . . I have come to better appreciate that the data creators and presenters of our time are powerful people.”1 Jack’s appreciation for the power that “data creators” wield was born out of the process of turning personal stories from the Federal Writers’ Project created in the 1930s into metadata categories that could then be analyzed through data visualization. As discussed in further detail below, metadata constitutes the informational categories used to classify material, essential to describing and organizing archival collections as well as assessing trends in large data sets. Although the process of metadata creation and data visualization may at first seem out of place in a composition classroom, this article demonstrates how converting stories generated from archival material into data helps students develop digital literacies essential to navigating and intervening in the algorithmic and data infrastructures that now shape our understanding of the world.

The life history with which Jack chose to work was particularly powerful as it told the story of a woman known as Elsie who was forced into prostitution in order to provide a better life for her son during the Great Depression. The story of this young woman’s desperate struggles was captured nearly eighty years ago in a “life history” that was part of the newly emerging field of social documentary. William Couch, Southeast regional director of the Federal Writers’ Project at the time, explained the somewhat radical premise of the life history project in this way: “all the people . . . must be given voice and allowed to speak” (x–xi). Funded by the federal government as part of the New Deal with the goal of putting unemployed writers to work, the Life History Project (part of the larger Federal Writers’ Project) aimed to document the stories of everyday Americans by letting them tell their own stories. The idea that average people had important stories to tell that should be preserved in the historical record and that those stories should be told from the person’s own perspective laid the groundwork for social history and oral history methods that are dominant today.
The stories in this collection are both powerful depictions of what it took to survive such a difficult time in American history and extremely problematic as they are not dictations of the interviewees’ responses to a uniform set of questions, as is now the common practice with oral history methods. Instead the stories are what Jerrold Hirsch calls “conversational narratives” that tell the reader as much about the writer as they do about the interviewee. Because of the nature of the stories and the window that they provide into writing practices and struggles during the Great Depression, they constitute valuable teaching tools. However, little attention has been given to these documents largely because of the difficulty of searching through the collection. While the collection has been digitized, it has not been tagged with a specific metadata schema or marked up to make the text within the document machine-readable—a common situation for many newly digitized archival collections. Therefore, if a user wanted to, for example, find life histories that discussed African American education, they would have to individually download and read each of the 1,200 life histories. Because such a task is unmanageable, the collection has remained largely hidden. To remedy this situation, I asked Jack and his classmates to participate in this process of converting the life histories from the Federal Writers’ Project into data.

This work involved collaboratively generating a metadata schema from both digitized and analogue versions of the archival material and then using the data to construct visualizations to draw new insights into the Life Histories Project that students used to write a formal research paper for potential publication in our university’s undergraduate research journal. Moreover, their collaboratively generated metadata served two public goals as well. First, the metadata was given to the Southern Historical Collection to enhance their finding aids. Second, it was also used to further develop Photogrammar, a digital humanities project that combines the photographs from the Farm Security Administration with the life histories from the Federal Writers’ Project. Therefore, the metadata work that the students conducted had a significant impact on making these documents more accessible to a wider audience. Additionally, as the students made interpretations about the data at each step in this process, they also became keenly aware of the power of data creation and its impact on shaping the information in the archive, as evidenced by Jack’s comments, together
with learning traditional composition skills including research methods, collaboration, and argumentation.

In what follows, I offer examples of class assignments that require the conversion of seldom-used primary archival texts, in this case from the Federal Writers’ Project, into data for analysis as a way to teach students digital literacy through archival rhetorics. Such an approach offers some important pedagogical affordances, including (1) the development of collaboration and argumentation skills, (2) the opportunity to learn and use feminist historiographic principles by directly intervening in archival structures, (3) the development of transferable digital skills that equip students with the technological know-how to intervene in our data-saturated, digital world, and (4) the ability to contribute to the advancement of the scholarly understanding of archival collections.

To demonstrate the value of this pedagogical approach, I begin with a discussion of recent approaches to the study of archives within composition and rhetoric and how teacher-scholars have used such theories and analysis to motivate their pedagogies, paying special attention to scholarship that focuses on the emerging and influential role that data play in archival rhetorics. I then move into an explanation of how I brought the strains of scholarship together to create assignments that required students to transfer archival information from the Federal Writers’ Project into data that they then visualized to make original arguments about the archival collection.

Archival Theory

In a special 2006 issue of Rhetoric and Public Affairs, Charles E. Morris III argued for “a ‘rhetorical (re)turn’ of the archive” in which he urged the field of rhetoric to analyze the construction of the archive itself, “chief among the inventional sites of rhetorical pasts” (113). He explained that the archive should be “understood not as a passive receptacle for historical documents and their ‘truths,’ or a benign research space, but rather as a dynamic site of rhetorical power” (115). Over the past ten years, many scholars in composition and rhetoric have taken up Morris’s call to understand archives as rhetorical constructs.2
Rhetoric and composition, however, is actually somewhat late to the archival scene as other fields of study, including history, postcolonial studies, and, of course, archival studies itself, have been grappling with the power structures embedded within archives for quite some time. Other scholars often credit poststructural theorists, namely Foucault and Derrida, with helping spark renewed interest in archives. Foucault’s *Archaeology of Knowledge* (1982) and Derrida’s “Archive Fever” (1994) are almost always cited in more critical readings of archives. These approaches unpack the power involved in constructing and using archival material, which Foucault argues amount to the “law of what can be said” (129). This interrogation of archival power helped fuel what Andreas Huyssen has called the “memory boom” during the 1980s. This boom grew as reaction against traditional ideas of historical methods that relied on a view of archives as “reliable repositories of truth, seedbeds of unabridged and veracious history. Open to inspection by all and preserved for all time . . . promising an authentic, untampered-with past” (Lowenthal 193). By focusing on the many voices consistently left out of this “authentic, untampered-with past,” scholars began questioning historical methods and how they relied on simplified notions of archival evidence. Many scholars within the history of rhetorics shared these same concerns, helping usher in a “critical shift from historical subjects to historical production itself,” though it took some time before the archive was seen as “a key site of that historical production” (Morris and Rawson 74).

While the field may have been a bit slow to position archives as objects of analysis, it is well poised to analyze the mechanisms at play in the construction of its power. As Cara A. Finnegan explains, archives “function as terministic screens, simultaneously revealing and concealing ‘facts,’ at once enabling and constraining interpretation” (117–18). K. J. Rawson’s work has gone far in interrogating the role of rhetoric in building such archival “facts.” He argues, “archives are rhetorical in the ways that they collect, preserve, classify, and make accessible historical materials. Furthermore, archives facilitate complex rhetorical interactions between researchers and historical materials through structural elements such as environment and proximity and through descriptive elements such as classification systems” (*Archiving* 44). It is Rawson’s last point about the descriptive elements of classification systems on which I want to dwell here, as much of the recent work on archives within composition and rhetoric has not adequately dealt
with the key rhetorical questions at play in developing and navigating these systems, questions in which digital literacy and information infrastructure are central.

Today, one’s encounter with the archive is almost always through some digital medium. Whether one “googles” where an archive is located, searches ArchiveGrid to find information about collection holdings, uses an online catalogue, or accesses digital material from a digital archive, the digital is always present—always shaping what can be found. Janine Solberg explains that such “digital tools and structures . . . increasingly support our research efforts” and have significant “material and epistemological implications for how we discover, access, and make sense of the past” (55). While Solberg and other scholars call for different strategies for how to deal with the influence of digital technologies in the archival research process, more attention is needed in interrogating precisely how these technologies function rhetorically to shape the archives. Such interrogation necessarily begins with a discussion of metadata and information infrastructure.

While some composition and rhetoric scholars may cringe at the term metadata, it is, in reality, a term in which all researchers are already well versed. Metadata refers to data about data and is used “when design[ing], creat[ing], describ[ing], preserv[ing], and us[ing] information systems and resources” (Gilliland 1). While the term often gets used to discuss digital information infrastructure today, it really refers to any instance of describing some characteristic of data. As scholars, we constantly use metadata to navigate libraries and databases to find relevant resources for research. For example, when researching articles on a database such as JSTOR, scholars may limit the date of publication to a particular time period, exclude book reviews, and click on relevant subject terms. Each of these choices used to narrow the search relies on metadata. The article must be classified by date, type of article, and subject, and each of these categories must have a set amount of possible descriptors within it. Therefore, metadata is essential to finding relevant material in the vast sea of information. Moreover, as Nathan R. Johnson explains, “since information is concomitant with knowledge, criticality, and awareness, the form of infrastructure has real consequences for the forms of public communication, knowledge, and political life that are already being studied by rhetoricians” (1). Acknowledging these consequences, Tarez Samra Graban argues that metadata and also “institutions, citation patterns, location(s) . . . [have] become the new
genres of information around which we should build our historical models and tools” (“Ripple” 7). Extending this call for greater attention to the “new genres of information,” I argue it is also incumbent upon us, as teachers, to instruct our students in how to critically engage and understand the rhetorical nature of this information infrastructure. As I demonstrate, archives offer one avenue for such instruction.

For archives, generating metadata is essential for organizing, cataloguing, and indexing extensive holdings and usually focuses on describing the original context of the collected material. Because archives are understood as storehouses of authentic historical material, great care is given to proving authenticity and situating the material in its original historical context. Information studies scholar Anne J. Gilliand explains, “elucidating and preserving context is what assists with identifying and preserving the evidential value of records and artifacts in and over time; it is what facilitates the authentication of those objects, and it is what assists researchers with their analysis and interpretation” (3). However, determining which specific metadata to use to describe and classify an object is perhaps one of the most important acts of rhetorical power in archives—the very site of invention that Morris referred to in his opening to *Rhetoric and Public Affairs*. For these reasons, teaching students how to create archival metadata not only teaches them important digital literacy skills concerning information infrastructure but actually places the rhetorical power in their own hands.

**Archival Pedagogies**

Composition and rhetoric scholars have also been bringing renewed attention to archival rhetorics into their teaching in order to achieve a number of pedagogical goals. Pamela VanHaitsma identifies two strands of scholarship on archives and pedagogy that have recently emerged: “first, on how archival work enables student participation in scholarly inquiry; and, second, on how digital technologies afford even greater participation by allowing students to also build new online archives” (36). This first strand of scholarship is generally acknowledged within the field as many have demonstrated how archival research holds immense weight for teaching
students how to conduct original research, gather evidence, and support arguments (Buehl et al.; Enoch and Jack; Greer; Hayden). Additionally, as more and more archives digitize their holdings, together with the quickly growing number of archives that are born-digital, teacher-scholars are turning to digital archives to achieve these same goals. Moreover, many teacher-scholars have begun to acknowledge that digital archives also offer pedagogical affordances for teaching students digital literacy.

James Purdy argues, “literacy in a networked, digital world will increasingly involve the ability to ethically, critically, and effectively create, navigate, evaluate, and use digital archives” (36). Picking up on this affordance of digital archives, Jessica Enoch and Pamela VanHaitsma argue for teaching students a concept of archival literacy that focuses on the relationship between rhetoric and the digital. They make a call to teach “students to analyze digital archives for their rhetorical properties with the goal of assessing the ways these properties affect and inflect the research and knowledge-building process” (218). Using existing archives, they offer a number of practical exercises to develop archival literacies, including “selection, exigence, narrative, collaboration, and constitution” (219). Although their focus is not on archival construction, they do acknowledge that teachers could delve “more deeply into the digital component of these archives to consider how tagging, metadata, search capability, design interface, and other technological features shape the researcher’s online experience” (233). Extending their notion of archival literacy in order to contribute to the growing scholarship that advocates for student-led construction of digital archives (Comer and Harker; Enoch; Gatta; Norcia; Solberg), I argue that by giving students the skills and opportunity to build the digital infrastructure of an archive and then use visualization techniques to analyze the data, they learn how these digital components work by directly contributing to the rhetoricity of the archive.

Photogrammar and Life Histories Collection
I developed this pedagogical approach that centers archival infrastructure within digital literacy pedagogy while teaching Writing in the Digital
Humanities, a first-year composition (FYC) course set within a writing-in-the-disciplines (WID) framework that I designed at the University of North Carolina, Chapel Hill. The course did not require any prerequisites or prior technological skills, fulfilled specific general education requirements, and brought in a variety of students, some with skills in data and programming, but most with only basic technological skills such as social media and word processing. Because this was a WID course, it aimed to teach students the rhetorical and stylistic conventions that govern professional and academic writing in the digital humanities (DH). Although there is little scholarly work on the ways in which to bring DH into the first-year composition curriculum, Olin Bjork argues that DH “provides a rationale and opportunity for composition instructors to expose their students to aspects of technical writing processes alongside the argumentative and expository writing processes practiced in the discipline of English studies.” In his chapter “Digital Humanities and the First-Year Writing Course,” he makes an excellent case for teaching students both qualitative and quantitative methods from DH, arguing that such an approach allows students to “become not only ‘close readers’ but also ‘distant readers,’ no longer content with supporting their insights on culture and rhetoric solely with examples from individual texts.”

With his insights in mind, I designed the course to include two major projects, one that required qualitative methods and “close reading” and another that required quantitative methods and “distance reading.” In order to do this, I crafted these assignments to focus solely on the Life Histories Collection so that students’ close reading of selected life histories in their first unit would drive their analysis of the entire collection in the second unit. Additionally, because I wanted students to understand the real-world possibilities that such distant reading methods offer, I brought them into a collaborative DH project, Photogrammar, I had recently joined with Yale University and the University of Richmond.

Photogrammar provides metadata and mapping visualizations of all the photographs taken by the Farm Security Administration/Office of War Information (FSA/OWI) during the Great Depression and World War II, many of which are now iconic in our understanding of life during that time period (see Figure 1). One of Photogrammar’s major innovations is that it goes beyond providing a database of the over 170,000 photographs by using metadata from the photographs to give users tools to visualize and analyze the collection in ways that have never been done before (see Figure 2). The
metadata extracted from the photos includes the photographer’s name, location where the photograph was taken, data about the photograph, subject heading, and caption if given. Therefore, users can track where and when a particular photographer traveled and what subjects he or she chose to give to certain photographs. Tools that provide such metadata exploration offer new “ways of visualizing relations among texts, topics, and subjects,” allowing researchers to make connections previously unseen (Graban et al. 239).

I joined the Photogrammar project with the proposal of adding the life histories from the Federal Writers’ Project as another layer to its visualizations capabilities, so that users would be able to read the stories alongside

---

Figure 1. Photograph from Marion Post Wolcott. Metadata included: location—Carrboro, Orange County, North Carolina; date—September 1939; caption—“Mrs. Lloyd, ninety-one year old mother of Miss Nettie Lloyd, who is a pellagra victim. Mrs. Lloyd was born and reared in Orange County; has lived on this spot since her marriage sixty-nine years ago. It is on a new road leading off from left of Route number 54, about four miles west of Carrboro, Orange County, North Carolina.”
Figure 2. Photogrammar (photogrammar.org). The image on the top uses location data from each photograph to position them on a county map of the United States. The darker hues indicate a greater amount of photographs. Here one can easily see the widespread nature of the photographs as well as which counties received the greatest attention. The image on the bottom shows how metadata can be compared to push analysis even further. Here a user can click on a county in California, and view how many photographs were taken by each photographer, the year the photographs were taken, and the subject of the caption that was used to describe the photograph. Seeing the data visualized next to each other allows for quick comparisons often not seen otherwise, which can spark new research questions.
the visual images of the Great Depression. Having already worked with the Life Histories Collection for a number of years as a way to introduce students to archival research, I knew that the collection offered many possibilities for extracting metadata that could enrich our understanding of the field of social documentary that emerged at the time. According to William Stott, “social documentary deals with facts that are alterable. It has an intellectual dimension to make clear what the facts are, why they came about, and how they can be changed for the better. Its more important dimension, however, is usually the emotional: feeling the fact may move the audience to wish to change it” (26). While many have noted the ways the FSA photographs hoped to elicit a particular affective response from their audience to spark social change, few have acknowledged how the life histories used personal narratives in a similar way.

In 1936, William Couch began his tenure as the assistant state director of North Carolina for the Federal Writers’ Project (FWP) and moved to the director position of the entire Southeast Region by 1938. He set out to capture the life histories of southerners from all walks of life, believing that these “readable and faithful representations of living persons . . . [would] give a fair picture of the structure and working of society” (Couch ix). He hoped that these life histories would challenge dominant conceptions of southern life and identity that often painted southerners as backward, lazy, and a key reason the United States could not move out of the economic depression. Though his project began with great interest, the start of World War II focused the nation on other pressing issues, and the many life histories that were collected under his administration have largely been ignored for the last eighty years.

While serving as an important precursor to oral histories with the intention of giving voices to ordinary people to share their complex lived realities, these life histories do not hold up to the methodological rigor of oral history practices today. Instead, these “conversational narratives” gave the writers great license in interpreting the interviewee’s life. As a result, their quality and form vary considerably. Some read as journalist accounts, others like the opening of a play with thick descriptions that paint a vivid picture; still others contain a problematic and bewildering written dialectic that clearly emanates from stereotypes concerning race, gender, and class. Because of the complicated composition of the life histories and their status as historical evidence, the collection was ideal for teaching students about
archival research methods. Moreover, the physical condition of the histories and their readily available metadata also made them an excellent choice for inclusion into the Photogrammar project. Each life history provides important information about its creation (in other words, metadata) similar to the FSA photographs, including the name of the writer and editor, the location and the date of the interview, and the name, gender, race, and occupation of the interviewee. This information is usually printed at the top of the life histories (see Figure 3). Additionally, the life histories are typed, digitized, and organized with one life history per archival folder. Knowing that the collection would be useful to the Photogrammar project and held potential for teaching students about metadata, I joined the collaboration with the understanding that I would use the collection as a pedagogical tool to teach students what I saw as an underdeveloped area within digital literacy—data construction and information infrastructure. Writing in

Figure 3. Life history of Georgia Crockett Aiken (Stanley Combe, interviewer).
the Digital Humanities, which I first taught in the spring of 2017, is an outgrowth of this goal.

In the following sections, I give specific examples of assignments I used in this course to demonstrate how constructing and visualizing archival metadata can teach students about collaborative composition, the power involved in information infrastructure and the benefits of feminist and social justice–oriented metadata indexing, and transferable digital skills. Although I use a single archival collection that was integrated into a public-facing digital humanities project, I demonstrate that metadata construction and indexing can be done with any archival collection. Moreover, it is the act of transferring archival text into data that can then be visualized that yields the most benefit for students. Therefore, this pedagogical approach can also be used with any selection of archival material, whether analogue or born-digital, as long as the material contains accompanying information to generate uniform metadata categories.

Collaboratively Creating Metadata

At a recent conference on invisible labor in the digital humanities, Mark Algee-Hewitt explained that a collaborative digital humanities model “requires ‘buy-in’ from all project participants to an equal extent.” Keeping this in mind together with the important elements outlined in UCLA’s Center for Digital Humanities Student Collaborators Bill of Rights (DiPressi et al.), I began the course with a discussion of goal setting for the collaborative project. Because we would be working collaboratively in class and then sharing our labor with the Photogrammar project and the Southern Historical Collection, I wanted students to be clear about what they desired to get out of this project. The model that they came up with was to conduct research and data construction together, but to write individual papers. In order to facilitate reflection on this process of research and writing, I also asked that they write what I called a “weekly writing reflection,” in which they could discuss successes, frustrations, or curiosities relating to each week of work. These reflections were then posted on our class digital platform, which were viewable to all students. I encouraged students to read and respond to their peers’ reflections in order to share sources and ideas, which Wendy Hayden argues facilitates networking and helps students “get past the image of the lone researcher and see their research as a social act” (409).
Additionally, students asked to spend time at the end of the semester coming up with examples of how to transfer the labor and skills completed in the class to their résumés, as well as to have their names noted on the Photogrammar site as “research assistants.” In order to create a system that adequately credited students’ labor on the project, I worked closely with two of my Photogrammar collaborators, Lauren Tilton and Taylor Arnold. As Tilton and Arnold also have undergraduate students working on other parts of the project in a lab setting (University of Richmond’s Digital Scholarship Lab), we were mindful of attributing student labor while also making sure that there was sufficient uniformity of the data to make it useful. In order to give students credit for their labor and have control over how their names were represented, we created a Github page linked to the Photogrammar site that gives the profiles of all student contributors, where each student is a site editor with control over how their profile is represented. Github is an open-source development platform widely used in the digital humanities community as a way to share both code and rationale for digital projects. Because students have the ability to edit how their profiles are displayed on this site, they can change and update their profiles over time to fit their career needs.

Additionally, Tilton, Arnold, and I outlined some pieces of metadata that we would need to build the infrastructure of the Photogrammar, including the basic demographic information that was printed on most life histories—name, race, and gender of the interviewer, interviewee, and editor, together with the date and location of the interview, as well as the title of the life history. We told students that this basic metadata would be incorporated into Photogrammar, but that they would need to determine other more nuanced forms of metadata to fuel class analysis, and possibly to influence future iterations of the project. Having laid the groundwork for our collaboration followed by activities that helped students understand how metadata worked in archival collections, we moved our focus to the Life History Collection (LHC).

As scholars such as Wendy Hayden and Susan Wells have noted, part of the allure of using archival materials in the classroom is the connections and traces that are found there, often sparking strong affective responses from the researcher-students. For this reason, I began our digital project in the reading room of the Southern Historical Collection. After briefly
discussing the origins of the LHC, emphasizing its situation with the larger Federal Writers’ Project and the New Deal together with the ambitious goals William Couch had for the project, I asked students to select five life histories that would be pulled for them when we visited the archive. Of course, in order to choose these life histories, students first had to move through the digital by going to Southern Historical Collection’s website, which contains the finding aid for the LHC. As previously mentioned, the LHC is not indexed, so this meant the finding aid consisted of a long list of the 1,200 life histories only organized by arbitrary folder (see Figure 4). I knew that students would likely feel some frustration in the limitations of this design, which I later used as a way to spur a conversation about information infrastructure and user experience and expectations to begin our metadata assignment. As students were selecting their life histories, I encouraged them to choose individuals, subjects, occupations, or locations with which they felt a strong connection as they would be working with these five stories for the entirety of the semester. Such an approach allowed students to become invested in the stories and the people represented in them so that they understood the stakes of the metadata construction, which, as J. James Bono et al. explain, “situate ‘knowledge,’ include or exclude different actors, dictate hierarchy, [and] reveal ideological and/or institutional politics and biases.” Therefore, I reasoned that if students

Figure 4. Federal Writers’ Project, Life History Collection finding aid, http://finding-aids.lib.unc.edu/03709/.
felt a connection to the individuals represented in the life histories, they would be more attuned to these issues of hierarchy and exclusion that are imbricated in metadata indexing. For this reason, I designed the first unit project to entail a close reading of one individual life history that students used to write a digital biography with Adobe Spark, a free digital storytelling platform that requires little technological expertise. As Adobe Spark represents a low-bridge, point-and-click technology that students quickly mastered with little instruction from me, I was able to focus on teaching them basic research and composition skills while exposing them to archival research and the Life History Collection.

After students were familiar with a number of life histories through their work on the first unit, we began the second unit on indexing and then visualizing the collection, which required much more class time on building technological skills. In this unit, each student was responsible for indexing the five life histories that they selected. I then brought their metadata together with existing data from the Photogrammar project, to total three hundred life histories, which constituted a sufficient sample size for data visualization. The first step in this project was for students to decide on a metadata schema to index the collection; in other words, a list of potential categories that could be used to organize the collection. In order to create this schema, they had to think through a number of fundamental rhetorical questions: What information is actually available in the archival collection? Who is the intended audience that will search for these documents? What will these users want when searching the LHC? Because students knew that their metadata would be used as part of the already existing Photogrammar project, they were able to envision a number of potential users: academic researchers, students, history aficionados, or family members conducting genealogy searches. Having specific users in mind together with close knowledge of a number of life histories helped students generate potential metadata categories.

As previously mentioned, the federal writers who originally wrote the histories already recorded a fair amount of information involved in the creation of the document, including the name of the interviewer, the location and date of the interview, and the name, race, and occupation of the interviewee. Students quickly and intuitively chose these categories, which were also necessary for the Photogrammar project, and then added them to our class Google spreadsheet (see Figure 5). As Peter Kittel and Troy Hicks
note, Google documents and spreadsheets foster collaboration as students can write simultaneously, and if all users are signed in, they can also see who is making changes and track those changes. Because spreadsheets are the primary way to write and manage metadata, students learned important and marketable data management skills through this collaborative work.

While the initial categories borrowed from the life histories came quickly, others proved more difficult. In a weekly writing insight, one student puzzled over this work of creating metadata categories by asking, “how can keywords be unified? How do we cope with missing information, or anomalies/exceptions to the rules we set? At what point do we consider adding a new 'column' (in our case) or information?” After significant discussion students settled on the following metadata categories: interviewer’s name, race, and gender; interviewee’s name, race, gender, occupation, education level, religion, marital status, number of children; editor’s name, race, and gender; location and date of interview; use of dialect in writing; writing style; and extent of revisions. However, the more difficult step of tagging or

<table>
<thead>
<tr>
<th>Title</th>
<th>Interviewee 1 Name</th>
<th>Interviewee 1 Race</th>
<th>Interviewee 1 Gender</th>
<th>Interviewer 1 Name</th>
<th>Interviewer 1 Race</th>
<th>Interviewer 1 Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life History of Mrs. Jim Shelton</td>
<td>Louise Abbit</td>
<td>white</td>
<td>female</td>
<td>Claude Dunnagan</td>
<td>white</td>
<td>male</td>
</tr>
<tr>
<td>The Holiness Preacher</td>
<td>Louise Abbit</td>
<td>white</td>
<td>female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cotton in the Blood</td>
<td>John Abner</td>
<td>white</td>
<td>male</td>
<td>George Andrews</td>
<td>white</td>
<td>male</td>
</tr>
<tr>
<td>The Photographer</td>
<td>John Abner</td>
<td>white</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>white</td>
<td>male</td>
</tr>
<tr>
<td>A Quiz Kif? A Casket?</td>
<td>John Abner</td>
<td>white</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>white</td>
<td>male</td>
</tr>
<tr>
<td>The Knitter</td>
<td>John Abner</td>
<td>white</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>white</td>
<td>male</td>
</tr>
<tr>
<td>Three Bibles</td>
<td>John Abner</td>
<td>white</td>
<td>male</td>
<td>William Couch</td>
<td>white</td>
<td>male</td>
</tr>
<tr>
<td>Up And Down</td>
<td>John Abner</td>
<td>white</td>
<td>male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tom Levine, Tailor</td>
<td>John Abner</td>
<td>white</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>white</td>
<td>male</td>
</tr>
<tr>
<td>A Gambier's Philosophy</td>
<td>John Abner</td>
<td>white</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>white</td>
<td>male</td>
</tr>
<tr>
<td>Through A Glass Darkly</td>
<td>George Anderson</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>white</td>
<td></td>
<td></td>
</tr>
<tr>
<td>We Have No Slums</td>
<td>Loretta Bailey</td>
<td>white</td>
<td>female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Life at Sea</td>
<td>James Beaman</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkey Raising</td>
<td>James Beaman</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>white</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Cabinet-Maker</td>
<td>James Beaman</td>
<td>male</td>
<td>Edwin Messergill</td>
<td>white</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5. Student-generated spreadsheet of life history metadata.
assigning a preselected list of possible descriptors (known as tags) for each category came next. For example, if the story explained that the interviewee left school in the second grade, which tag should they write under the “education level” category? To answer this question, students had to collaboratively decide on the number and specific tags they would allot for each category. In this example, they decided that they would give five possible tags to this education category: none, elementary, high school, college, and vocational. Additionally, they assigned the level of education based on whether the interviewee actually finished the designated division. Therefore, because the interviewee left school in the second grade, they gave a tag of “none” for the education category. It is at this step, as students began to collaboratively make their own rules and standards for the archival metadata, that they really began to see the power of interpretation that was in their hands.

Feminist and Social Justice–Oriented Metadata Indexing

In her weekly writing insight, Alex Andrews explained, “with my life histories, I struggled with categorizing the major issues that they dealt with because many problems are not cut-and-dry, but fluid and hard to pin down. I have to resist the urge to speculate by digging too deep into the manuscript.” Also sensing this difficulty of assigning tags, Delane Dixon wrote, “The difficult portion of this week was figuring out how to best describe the metadata but in a manner effective for people to research in the future . . . I had to think deeply about what was being said and what it implies for the metadata.” Naomi Van Horn explained how issues of audience proved to be a challenge for her; “Our wide range of an audience makes it more difficult to come up with specific terms for our metadata. Scholars and teachers will have larger and more specific vocabularies than students and everyday people that just want to learn more about history.” As these three comments evidence, by being in charge of crafting the rules and standards that shaped the archival infrastructure, students were able to see the importance of both meeting their audience’s expectations and maintaining historical authenticity. Also noting the importance of focusing on “standards in the making,” Bono et al. argue that such an approach
“stresses information’s ecology, or how classification systems—as cultural formations—resonate, clash and create space for ambivalence, exclusion and inclusion.” Information and library scholars Marisa Elena Duarte and Miranda Berlde-Lewis explain this knowledge-making effect of classification systems as resulting from “the power to name.” Students felt the weight of this power especially when crafting indexing standards for race and gender within the life histories.

The issue of tagging racial categories for historical documents is, of course, a particularly powerful one that reflects the complicated and uneven relationship between history and identity in the United States. As a result, such tags are not simply “a description of racial identity (one of many which must be submitted to the normalizing impulses of political-correctness),” but “an organizing principle which frames the context and ‘uses’ of the concept of race itself” (Bono et al.). In other words, because these terms are the words that a user must use to access, navigate, and understand the collection, they shape the knowledge of the material itself. In order to explain the profound weight of how metadata shapes knowledge production, I tell students about my work with two Native American tribes located in Louisiana (the Pointe-au-Chien and the Isle de Jean Charles Band of the Biloxi-Chitimacha Confederation of Muskogees, or BCCM) who are currently petitioning the US government for federal recognition. As I have written about elsewhere (see Rivard and Cushman et al.), the federal recognition process in which tribes must petition the federal government for formal acknowledgment as a sovereign tribe and “domestic dependent nation” (Quinn 356) requires extensive archival research in which tribes must prove they existed as a political entity from the present until historical times, defined as the point at which European colonizers set foot on their land. While the colonialist implications of this system of recognition are obvious, the role that archival infrastructure and metadata play may not be. During my first research trip for the Pointe-au-Chien and BCCM, I was tasked with investigating the papers of the French colonial governor of the region, most of which were in French. Before entering the archive, tribal members gave me a list of keywords to use to retrieve relevant materials for their case. The list not only included their ancestors’ names and important locations, but also the tip for me to use the word *sauvage* in conjunction with the other names. This savvy research strategy, which proved quite helpful in retrieving relevant material, represents profound
epistemological violence. I ask students to think about what it means to type in this brutal word as a way to access history both for non-Native and Native researchers, and how such a metadata category already shapes the knowledge to be found within it.

Together with this work, I also discuss K. J. Rawson's writings on the Digital Transgender Archive. Rawson explains the intense discussions by his students over how to tag a photograph of Dawn Wilson and Marisa Richmond taken sometime in the 1990s. While online searches revealed that the two were well-known African American transgender activists, there was “nothing inherent in this photograph” that allowed Rawson and the students to discern how Wilson and Richmond “identified when the image was taken” (Rawson, “Rhetorical” 16). Rawson explains:

Even if we are relatively certain that we know how they identify now, at what point in each of their lives did they begin identifying in that way? What are each of their preferred terms for describing their racial identity and their gender identity, both now and when the image was taken? What other identities would they claim? By even asking these questions, we are queering the archive by carefully approaching the power involved in archival description and by recognizing the imbalance that exists as an unnamed and invisible force is tasked with classifying named and visible people. (17–18)

Rawson’s discussion helps to demonstrate how complicated the process of tagging archival material can be, and the powerful role that researchers have in the process. Discussing Rawson’s work together with my research for the Pointe-au-Chien and BCCM helped students think deeply about the political stakes of metadata construction. They spent significant time coming up with a standard of tagging racial categories that was mindful of the historical inequities that have been enacted through classification systems. Rather than implementing the outdated and offensive terms for racial categories used in the 1930s life histories, especially for communities of color, students instead applied modern terms used by community members of the respective ethnic or racial group in their tagging standards. Students also attended to gender inequities, especially the invisibilizing of women’s names, through the standards they set in their metadata schema. For example, many of the life histories defined a married woman’s name by inserting “Mrs.” in front of the husband’s name (Mrs. Jim Smith). However, students noticed that by reading the life history itself for context clues, it was easy to determine the woman’s first name. Therefore, the students
used these context clues when available to at least give the women some marker of their own identities.

While students saw many important affordances to their metadata standards, many archivists would frown at their practices. Traditional archivists usually use the terms within the original documents to index material, because, they argue, it works to maintain the authenticity of the documents. Additionally, archivists are keenly aware that the classification systems are necessary to actually find relevant information for users, and those users change over time. This desire to maintain uniformity is another reason that archivists usually argue for using language from the original documents so as to avoid constantly changing classification systems. On the other hand, many of those original documents were created within an extremely unequal system whose purpose was to undermine the histories of Native communities and communities of color.

Students recognized this complicated balance that required attention to power structures and retrieval efficiency, but nonetheless they made a rather strong rhetorical argument for their changes. During a class discussion regarding the standards for tagging race categories, one student asserted, “nobody would use those offensive terms when they typed something into the search box.” Another student responded, “and we wouldn’t want them to either.” Through this important exchange, students understood metadata as composition—they were crafting an ontology of the archive for their intended audience. Moreover, by directly addressing inequities in the archival infrastructure, students participated in feminist and justice-oriented historiographic methods, what Rawson calls “queering the archive.” They were recovering and reconceiving traces that might not have otherwise been seen. Such practices constitute the type of “strategic contemplation” (Royster and Kirsch) that Jessica Enoch and Jean Bessette argue enhances feminist historiographic priorities, which can help address “hesitations feminist rhetoricians may have regarding digital tools” (636).

Moreover, in many ways this is exactly the kind of work that Graban calls for when she argues for feminist metadata standards that provide “a new form of recovery by disrupting several assumptions; that historical relationships are drawn according to traditional taxonomies of storage, ordering, and use” (“Ripple” 9). Therefore, when students were invited into digital archival work in which they were in charge of key decisions in metadata composition, they were given the space to enact their own disruptions of
traditional data practices. Moreover, as part of their metadata schema, students worked collectively via a Google Doc to document their rationale for each metadata standard they created. Not only does such documentation constitute rigorous archival work and underscore the rhetorical structure of the archive, but it also helps future classes and data users (both in the Photogrammar project and at the UNC archives) understand the rationale used to construct the schema. Users will then be able to decide if and how they want to use the data based on the rationale behind its creation.

Data Visualization

Once students finished indexing their life histories, they then moved into the next phase of the project: data visualization. While there are many data visualization programs, I chose Tableau because it is one of the prominent programs used within digital humanities, it was recommended by our university’s data visualization librarian, Lorin Bruckner, and the company offered free desktop licenses for educational purposes. Teaching students how to use Tableau required significant classroom time, because it was a new and more complex program that again required students to think critically about the role of rhetoric in data construction. However, as Daniel Anderson explains, such time is often well spent as “experimenting with unfamiliar technologies can facilitate a sense of creativity that can lead to motivation” (44). Even more pressing than the newness of the program was the fact that data visualization highlighted issues in students’ metadata construction and required them to think differently about the Life Histories Collection.

Students’ initial visualizations were, to put it simply, a mess. The reason for this “messy data” was that in the parlance of data scientists, it had not been properly “cleaned”; in other words, no one had meticulously gone through each line in the spreadsheet to ensure that all terms were uniformly written. Having clean data is essential because of the way in which the computer “reads” the data. For example, if one student wrote someone’s education as “elementary” and another wrote “grade school” and yet another wrote “Elementary,” the computer would read these as three different tags even though students intended them to be the same. However, this problem actually turned into a fantastic collaborative learning experience about composing data. First, students began to see how search queries work, because they understood what the computer actually reads. In her weekly writing insight, Madison Seals explained, “We also have to
keep in mind the technology we’re working with. In class, we had to ensure that we were using exactly the same terms across the board (within each category) to classify the life histories. That way, when we move on to data visualization, there won’t be any extra/unnecessary categorical inputs that are essentially the same as others (e.g. ‘widower’ and ‘widowed’ or ‘Black’ and ‘black’).” Students also quickly saw similarities between these problems and issues they had encountered when searching scholarly databases for a desired journal article and even online searches when they were looking for a new restaurant. Such insights are key to helping students develop information and digital literacies necessary for navigating the technologies of the twenty-first century.

Moreover, because students had seen the result of their messy data in their first attempts at visualization, they were eager to clean it. During the resulting data-cleaning session, the class erupted into a cacophony of questions and answers: “Miriam, why did you capitalize the racial category? Do you consider that a proper noun?” “Are we going to allow plural words?” “How do you know the race of the writer Cora Bennett—wow, you found her on a census report from ancestry.com!” This process of collaborative data construction blurred the lines of authorship, while simultaneously creating a communal sense of responsibility. As Johndan Johnson-Eilola explains, “collaboration helps symbolic analysts work together to solve problems while crossing complex disciplinary domains…team members brainstorm ideas and solutions, critique each other’s work, and provide support and feedback to the teammates” (259–60).

Once the data was cleaned, students were able to focus on the new type of research that was enabled through visualization. Richard White explains that data visualization “generates questions that might otherwise go unasked, it reveals historical relations that might otherwise go unnoticed, and it undermines, or substantiates, stories upon which we build our own versions of the past” (6). Echoing this sentiment, Krista Kennedy and Seth Long argue that “visualization of data fosters interpretation and allows patterns to be detected—and patterns, as Franco Moretti bluntly puts it, tell us that ‘something needs to be explained’ (2005, 39)” (Kennedy and Long 147). While data visualization enables a new way of seeing an archival collection by analyzing the entire corpus at once, students struggled with how to begin this process as it represents yet another way to think about important rhetorical dimensions of data. In her weekly writing insight,
Madison Seal wrote, “with so many possibilities, it’s hard to decide just what kind of visualizations I want to create.” Another student wrote that using Tableau was “a very interesting and helpful process in gaining insight into the patterns and details of the life histories during the Great Depression,” but that it was a challenge to figure “out how to best use Tableau to make it as advantageous to the research as possible.” While it took some time to convince students that all scholars using data visualization methods struggle with this part in the research process, it also offered the opportunity to demonstrate to them that “the critical problem of what and how to visualize reveals that there is a rhetoricity to historical data work” (Graban, “Ripple” 2). Therefore, as students made choices about which data to visualize and how to present that data, they saw that data visualizations, too, are rhetorical constructs. They came to understand that data visualizations are only able to demonstrate one small aspect of the data in a very particular way determined by the desires of the author to convince the reader of the significance of seeing the data in such a way. Moreover, they were using this process of data visualization as a methodology for analyzing a collection that had never been analyzed through data visualization before. As a result, a number of students made innovative insights, especially regarding the use of dialect to represent speech patterns of certain interviewees that will truly contribute to the scholarly community.

During the process of indexing the life histories, students puzzled over how to deal with many of the life histories that were written with a type of dialect, or what others might call phonetically spelled vernacular speech, which contained purposefully misspelled words and excessive use of apostrophes. Take for example, the life history of Eliza Hall, an African American washwoman, written by Louise L. Abbitt. In this passage, Hall is discussing her limited schooling:

When it got awful cold, the older boys tuk turns about bringin’ in wood an’ buildin’ the fire in the ole pot-bellied stove in the middle o’ the room. Some of us’d bring sweet ’taters to school an’ roast ’em in the ashes under the fire, so they’d be done in time fer dinner. I don’t reckin I learned so much in the time I tended school, ’ceptin to print my name an’ figger a little. (3642)

They came to understand that data visualizations are only able to demonstrate one small aspect of the data in a very particular way determined by the desires of the author to convince the reader of the significance of seeing the data in such a way.
Because this type of written dialect was glaringly present in some life histories and not others, students decided to index the presence of dialect, a particularly perceptive move as this issue of written speech was significant within the Federal Writers’ Project. While there is little scholarship on the Life Histories Collection, many scholars have discussed the use of written dialect in the Slave Narratives, a FWP project that exclusively focused on writing life histories from former slaves. In this project, as with the life histories, writers were told to write the stories in the words of the interviewee. For the ex-slave narratives, this directive “made the question of how to describe the black vernacular of primary importance” (Stewart 78). Because the presentation of black vernacular was “often a reflection of white interviewers’ idiosyncratic (and frequently racist) interpretations,” directors within the Federal Writers’ Project “composed guidelines for the transcription of dialect” (79). Catherine Stewart argues that while the goals were to preserve “the flavor of speech,” “the use of phonetic speech resulted in stereotypical depictions of African Americans as provincial folk people, whose racial difference was embodied in their speech” (79). Toni Morrison has explained that this “eye dialect,” which “relies on phonetic spellings and apostrophes to make visible to the eye sounds that cannot be heard in the medium of print,” “renders the speech of black characters ‘as an alien, estranging dialect made deliberately unintelligible by spellings contrived to disfamiliarize it’” (qtd. in Stewart 79).

Reading and discussing this issue of written dialect, some students began to wonder if there was a correlation between the race of the interviewee and the use of written dialect, a query that was possible because students made the choice to tag the life histories with either using substantial dialect or not. One student, Zoe Leatherwood, then executed this comparison of interviewee race and presence of written dialect in her data visualization. Her results were stunning, as they demonstrated a strong correlation between an interviewee’s race and the decision to use dialect in the writing of the life history (see Figure 6). Because the Life History Collection contains interviews from both African American and white interviewees, her results mark a significant scholarly finding and warrant further research as they demonstrate how written dialect was used to mark race more than region. Therefore, such a finding demonstrated the way in which “data visualization is central to data interpretation . . . . Without this transformative step of
rendering text visual, certain trends and patterns may go unnoticed, hidden within the textual or numerical aggregate” (Kennedy and Long 147).

As Zoe’s work demonstrates, assignments that require students to create and visualize metadata for archival collections automatically set students up to conduct original research, often leading them to new insights that directly contribute to scholarly conversations. In so doing, “they transform from thinking of themselves as students to seeing that their insights and their work have value to an academic field. They can see themselves not as ‘outsiders’ but as ‘novices’ to academic work” (Grobman 188, cited in Hayden 418). Moreover, when centering data construction and archival infrastructure, this transformation also equips students with digital literacy skills necessary to navigate the infrastructure that now governs our access to information because they understand its inherent rhetoricity.

Conclusion

In discussing the affordances of a pedagogical approach that uses archival rhetorics to center information infrastructure and data construction in the development of digital literacies, I demonstrate that collaborative work with archives and existing digital humanities projects can be mutually helpful to all parties. Archivists are just beginning to tackle the endless work of tagging archival documents and therefore will always need help with this labor. Moreover, it is exactly this work of the transfer of text to data that proves so generative for the development of digital literacies. This act of transfer requires students to directly take on issues of historical authenticity, the importance of attending to multiple audiences in composition practices, the assessment of what gets counted as evidence and how to best communicate the results, and the fundamental power of archival rhetorics.
This act of transfer requires students to directly take on issues of historical authenticity, the importance of attending to multiple audiences in composition practices, the assessment of what gets counted as evidence and how to best communicate the results, and the fundamental power of archival rhetorics and information infrastructure to shape knowledge production. In so doing, students build these critical digital literacies by making their own “historiographic interventions” (Glenn and Enoch 18) in the historical record. They learn contemporary composition and digital literacies by doing history—making history more assessable to a larger audience. As Jack Turner, the student who initially grappled with how best to create tags for the Elsie’s heartbreaking life history, explained in one of his final weekly writing insights in the class, “Data has always been a part of our lives, but modern computing and data collection techniques have multiplied its potential implications . . . An understanding of how data is used and created is therefore greatly important for all of us.”

Acknowledgments
I would like to extend my deepest thanks to my collaborators on the Photo-grammar project: Lauren Tilton, Lauren Wexler, and Taylor Arnold. Working with this material together with such incredible people has been an amazingly rewarding experience. Additionally, I would like to thank panelists from Cultural Rhetorics (Pamela VanHaitsma and Jean Bessette) and CCCC (Daniel Anderson, Jim Ridolfo, and Ashley Hall), who gave me extremely helpful feedback when I presented various parts of this paper at the respective conferences. I am also indebted to the extensive help of Matt Turi (manuscript research and instructional librarian at the Southern Historical Collection) and Lorin Bruckner (data visualization librarian), who led workshops to help the students master archive and visualization skills, respectively.

Notes
1. All student work used in this article was done with expressed written consent in which students chose to allow their name to be used or to remain anonymous.
2. For a few examples of such scholarship, see Bessette; Buehl et al.; Cushman; Enoch and VanHaitsma; Gaillet; Graban, “From”; Graban et al.; Hayden; Kirsch and Rohan; Purdy; Ramsey et al.; Rawson, Archiving Transgender, and “The Rhetorical Power.”
3. For an excellent overview of such helpful advice for researching archives, see Kirsch and Rohan.

4. I developed this course as a result of my participation in university initiatives designed to bring data studies into humanities pedagogy. I am deeply grateful to Daniel Anderson, director of the Carolina Digital Humanities Initiative, for leading the Faculty Learning Group on Data Studies in which we developed many helpful pedagogical activities that directly informed my teaching. Additionally, I am extremely thankful to both the Carolina Digital Humanities Initiative and Data@Carolina for providing me with course development grants that gave me the time and resources to construct the course.

5. See Moretti for a detailed explanation of this method.


7. In order to teach the students important project management skills necessary for the completion of such a collaborative digital project, I used Trello, a free project management software. The platform is organized according to a series of lists, similar to sticky notes, that contain action items needed to be completed. Students are able to post their completed work, see what still needs to be finished, and send messages to each other.

8. The archive session was made possible by the organization and work of archivist Matt Turi at the Southern Historical Collection. His advice and instructional help was indispensable to the success of the entire class.

9. Students marked each life history with either yes or no for the presence of written dialect. They decided that if the life history contained substantive (defined as being used in more than one-third of the history) phonetic written speech with excessive apostrophes, the life history would be marked with yes. Of course, there is significant room for interpretation and variation between students, but the reality is that this is precisely the case for all metadata construction.

Works Cited

Abbitt, Louise, interviewer. “Reckin’ I’ll Be Washin’ an Ironin’ Till I Drop Daid.” Southern Historical Collection, Wilson Library, University of North Carolina at Chapel Hill, Folder 274, Federal Writers’ Project papers #3709.


——. “‘Flagged for Deletion’: Wikipe-dia, the Federal Writers’ Project, and First-Year Composition.” *Teaching Rhetoric and Composition through the Archives*, edited by Wendy Hayden and Tarez Graban, forthcoming.


Courtney Rivard
Courtney Rivard (crivard@email.unc.edu) is the director of the Digital Literacy and Communications Lab and teaching associate professor at the University of North Carolina, Chapel Hill. Her interdisciplinary work bridges rhetorical studies, digital humanities, and feminist studies to analyze archives as rhetorical constructs. She recently received an ACLS Digital Extension Grant for her collaboration with the Photogrammar Project. Her work can also be found in *Rhetoric Review, Digital Humanities Quarterly* (forthcoming), *Settler Colonial Studies*, and *Identity Technologies: Producing Online Selves*, edited by Rak and Poletti.