

USING DIGITAL TECHNOLOGIES TO COLLECT AND STUDY STUDENT WRITING

André C. Buchenot

Despite its widely cited utility for professional and program development,¹ instructors rarely research students' writing outside of the context of a particular course. Put differently, we are keen scholars of student writing during the semester—we investigate composing processes across multiple drafts; we observe collaboration in classroom activities; we reflect on student learning through our written comments—but we seldom continue our study after assigning a final grade and committing the remaining papers to a filing cabinet or recycling bin.

The limited presence of extracurricular research on student writing can be attributed in no small part to the material demands of working with student documents. In their 1988 study of the frequency of written “errors,” Robert Connors and Andrea Lunsford collected over 21,500 documents from over 300 instructors and ended up with an “imposing mass” occupying “approximately 30 feet of hastily-installed shelving” (398). Twenty years later, Andrea Lunsford repeated the study with Karen Lunsford and collected only 1,826 documents, attributing the lower participation to the “tedious, the time-consuming, the mindnumbing task of filling out dozens upon dozens of (Institutional Review Board) forms” (787-88). Taking a longitudinal approach, Nancy Sommers and Laura Saltz collected “more than 600 pounds of student writing, 520 hours of transcribed interviews, and countless megabytes of survey data” over the course of four years (126).

Fortunately, developments in academic technologies and composing practices offer possibilities for shifting the labor of

studying student writing away from imposing masses of paper and toward curated collections of files. In terms of academic technologies, the key development is the widespread adoption of *networked computing* at secondary and post-secondary institutions. Teachers engage with networked computing when they check school email, share files, and use learning management systems (LMS) such as Blackboard or Canvas. For composing practices, the key development is the near ubiquity of *digital writing* or writing that “exists as pixels and bits on a computer at some point in the composing process” (McKee and DeVoss np). Unlike its chirographic antecedent, the material character of digital writing allows it to be collected, organized, tagged, and indexed through electronic means. If teacher-researchers are able to marshal networked computing and digital writing in this way, it is possible to create a powerful tool for research.

I created such a system using readily available digital technologies (our campus LMS, a laptop, office software) and some minimal assistance from colleagues (roughly ten minutes, once a semester). The result was a digital archive of over 2,000 student documents that have been used to conduct assessments, design teaching interventions, and establish a clearer sense of student learning. In this article, I outline methods that helped me create this system focusing on those principles that might transfer to other institutional settings. I begin by discussing the labor of using student writing in composition pedagogy and scholarship. Then, I review key developments in academic technologies and composing practices that enabled the creation of this digital archive and that might be used to create similar resources at other institutions. The remainder of the article discusses efforts to use my campus’ LMS as an entry point for collecting, storing, organizing, and analyzing a substantial corpus of student writing. Through this discussion, I present concrete details for using digital technologies to support research that can then be used to improve teaching. Overall, I contend that changes in networked computing and digital writing have opened up engaging possibilities that make systematic research on student writing distinctly possible.

At first blush, instructors might be hesitant to take on additional work, but it is worth the modest investment of time and effort to study student writing with a precision that is unavailable in the anecdotal and ad hoc studies that characterize much of the research in the field. When we draw on systematic studies of student writing, we move away from what Steve E. Graham and Karen R. Harris call “teaching lore”—informally collected knowledge about teaching—and toward actionable and persuasive evidence (92). Such evidence is useful for the practical work of advocating for our students and programs. Rather than arguing passionately for the value of revision, instructors might use data gleaned from a digital archive of student writing to show the ways student writing improves when it develops over multiple drafts. In addition to contributing to evidence-based practices, this kind of data can contribute much needed systematic research to the larger field of composition and rhetoric.

The Labor of Using Student Writing in Composition Pedagogy and Research

Placing student writing at the center of instruction is a defining move of composition pedagogy. This centrality of student writing defines the everyday activities of our teaching: We photocopy drafts for class discussions; we write feedback to encourage revision; we plan activities for peer review; we read closely for evidence of learning. In sum, we expend significant energy attending to the concrete labor of treating students’ writing as “real” writing—that is, writing that deserves sustained and careful attention. Decades of scholarship support these choices including Donald Murray’s assertion, “the text of the writing course is the student’s own writing” (5) and Bruce Horner’s argument that “much that has been accomplished in composition has come from the practice of paying close attention to student writing” (523). Placing student writing at the center of research is also a defining move of composition research. This centrality is seen in research publications drawing on large collections of student writing such

as Mina Shaughnessy's *Errors and Expectations*, Deborah Brandt's *Literacy in American Lives*, or Lee Ann Carroll's *Rehearsing New Roles* as well as more theoretical pieces employing a close reading of smaller sets of student writing such as Min-Zhan Lu's "Professing Multiculturalism," Richard Miller's "Fault Lines in the Contact Zone," and Lad Tobin's *Reading Student Writing: Confessions, Meditations, and Rants*.² Joseph Harris articulates an ideal relationship among student writing, pedagogy, and scholarship arguing that:

Taking students seriously as writers defines the intellectual work of composition. And thinking in public about the work students have done in your courses helps you become a more reflective and self-critical teacher. I thus think we need not only to allow but also to encourage teachers to cite and use student writings much as they might draw on critical essays or novels or poems—that is, as part of the repertoire of texts they've read and that have informed their thinking. (23-24)

When teacher-researchers "think in public" about research on student writing, they often include an accounting of the concrete labor involved in collecting and studying these texts. That is, they describe their research process in terms of feet and pounds of student documents. I return to these material demands to explore the differences in valuation between the labor of teaching with student writing and the labor of researching student writing. With regards to teaching, the myriad tasks required to teach writing using student texts tend to be invisible because they have very little academic exchange value. Promotions are not awarded because instructors stay up late responding to papers; they are awarded for positive evaluations, high test scores, and, at universities, publications. In contrast, the myriad tasks completed to *research* student writing are highlighted because they do have academic exchange value as evidence of rigorous research methodology. Detailing a research project's methods suggests the

teacher-researcher has done her or his due diligence and offers crucial details for other academics who might seek to replicate/validate/extend the study being described.³

Given this disparity in visibility and value, it is easy to understand why instructors are hesitant to take on the additional labor: We have more than enough to do now. Why take on more work that might not be valued? In response to the (quite reasonable) hesitation, I argue that by taking on even a small amount of research labor, it might be possible to 1) enjoy the benefits associated with the academic exchange value of research, 2) draw attention to the labor of teaching in order to “re-value” it, and 3) contribute to the field of composition and rhetoric. Put differently, when it is connected to research on student writing, it is possible to associate everyday teaching activities with the academic exchange value usually reserved for overtly scholarly activities. I am not suggesting that the everyday labor of teaching student writing does not have educational or other forms of value, nor am I suggesting that academic exchange value is somehow superior to the use value of teaching. I am suggesting that putting the labor of teaching in dialogue with the labor of research presents some promising possibilities for rethinking the meaning of our work.

Developments in Academic Technology and Composing Practices

Contributing to this promising dialogue are recent developments in academic technology and composing practices that allow teacher-researchers to blend research practices in with their teaching. In terms of academic technology, the key development is the wide adoption of *networked computing*. Broadly, networked computing includes all the technologies used to share information across computers including tools ranging from servers and modems to electrical wiring. These technologies have been a fixture of university campuses for some time, but networked computing is not a college-only phenomenon; the National Center

for Education Statistics estimates that 93% of public school classrooms with computers have access to the Internet and a majority have access to email and file sharing (Gray et al. 3). As digital files in a network, digital information can move smoothly between computers along paths created by networked computing.

These paths created by networked computers are only useful for research because of the shift in student composing practices from ink-and-paper writing to *digital writing*.⁴ The Writing in Digital Environments (WIDE)⁵ research collective defines digital writing as:

the art and practice of preparing documents primarily by computer and often for online delivery. Digital writing often requires attention to the theories and practices of designing, planning, constructing, and maintaining dynamic and interactive texts—texts that may wind up fragmented and published within and across databases. Texts that may, and often do, include multiple media elements, such as images, video, and audio. (np)

This definition usefully highlights how the processes and products of digital writing are materially different than earlier forms of writing. As the WIDE collective points out, the process of digital writing takes place (primarily) on a computer, smartphone, tablet, or other device connected to the Internet. Because of this connectivity, a digital writer never composes alone. She or he can consult colleagues, review published texts, look up usage guidelines, and access a host of other resources. This radical connectivity benefits research on student writing because working within networks has become a routinized process for students. Asking them to share their digital writing over networks for the purpose of collecting data is a major departure from their existing, digital writing processes.

The WIDE collective's definition describes the products of digital writing as "dynamic and interactive" and often including "multiple media elements." These characteristics are easy to see in

multimodal, web-based texts like a Tumblr blog that includes images, quizzes, and animations alongside prose, but they are also present in seemingly straightforward texts like Microsoft Word documents. In both modally-rich blogs and word-processed documents, the substance is made possible by encoded instructions that exist “beneath” what is readily seen when we compose or read on computers. Douglas Eyman and Cheryl E. Ball describe code as “the underlying structure that has to function properly in order for a digital text to achieve its design goals and support the rhetorical functions of usability and accessibility” (116). Regardless of a writer’s awareness, code is present when she or he composes on a computer. Put differently, when we write on paper, we write with ink or graphite or some other mark-making medium. When we write on screen we write with code, even if that code is hidden from us by a user-friendly interface. This begs the question, if it supports digital writing regardless of our awareness of it, why is considering the role of code in writing important to writing instructors? There are several answers to this question (see Kristin Arola’s “The Design of Web 2.0,” Lisa Dush’s “When Writing Becomes Content,” and Annette Vee’s “Is Coding the New Literacy Everyone Should Learn?”) but for the purposes of this article, code is important because it is the feature that makes digital writing amenable to electronic collection and organization. In other words, code is important because it makes it possible to move writing swiftly through networks and store it purposefully in digital archives.

To reiterate, networked computing and digital writing can support research on student writing by blending the existing labor of teaching with the labor of researching student texts. Students and teacher-researchers are already producing digital writing using networked computing technologies. In the overwhelming majority of educational settings, asking students and teacher-researchers to modify their use of these technologies for the purposes of research does not add undue burden. Further, the fact that digital writing is made of code rather than paper means it can be collected, organized, stored, and analyzed using basic software.

In what follows, I move away from these abstract descriptions and into examples of how these practices took place in my research.

Using Recent Developments to Conduct Research on Student Writing

In this section, I offer concrete recommendations that teacher-researchers might use in their own research and professional development projects. The section describes my methods for collecting and organizing student writing as well as some suggestions for analysis and future innovation. Underlining these descriptions and suggestions are a set of principles designed to be applicable across institutions.

Goal-Driven, Technology-Centric Research on Student Writing

The overarching goal of my research was to develop a model for assessing student learning using metrics taken from campus- and department-level outcomes statements. The outcomes stipulated that students should demonstrate mastery of a range of competencies such as “apply, analyze, evaluate, and create knowledge” (“Principles”). The complexities of these competencies motivated my decision to base my research on end-of-semester student writing rather than test results or surveys or any other text that might evidence student learning. Similarly, the complexities of studying a large, diverse English department motivated my decision to collect an expansive corpus of student writing in an effort to represent the range of teaching and learning happening.

These seemingly academic research decisions informed all of the technological decisions that I detail below. All teacher-researchers looking to conduct research on student writing should begin by articulating similar research goals before considering technology options. This is not to imply that research goals will completely dictate technology use; there will always be a give-and-take between goals and technologies.⁶ Still, beginning with a

clear sense of an ending will help focus technology use and avoid systemic problems with technical implementation.

The LMS as Data Collection Tool

My data collection method was designed to increase faculty involvement by limiting the impact on the everyday work of instructors and students. I accomplished this goal by identifying the key functions of routinely-used technologies and, then, developed a protocol to collect student writing using these familiar technical features. The result was a data collection that did not require instructors or students to engage in activities that significantly departed from the regular labor of the course. Based on these principles, I decided to collect student writing using our campus' LMS, Oncourse Collaboration and Learning or, more commonly, Oncourse. This software features many of the usual functions of contemporary LMS—grade tracking, test administering, email messaging—but what drew me to Oncourse was the way it was integrated into the everyday labor of teaching in the English department. At the start of each semester, new Oncourse sites are created for every section of every course and instructors are expected to populate these sites with syllabi and other course documents. Instructors are not required to ask students to submit writing via Oncourse, but many do because 1) it helps to manage the paperwork of collecting student writing and, 2) it allows them access to the Turnitin plagiarism detecting software. In addition to being familiar and accessible, Oncourse was attractive because of its collaborative administrative functions. The LMS allows a course's instructor of record to "enroll" other instructors into the course's Oncourse site with administrative privileges, giving them the ability to assign grades, post content, and download student writing. These collaborative options allow for a researcher to access and collect student writing from multiple courses with minimal involvement from the instructor of record.

The data collection protocol that grew from these technologies had three steps. First, instructors were asked to inform their

students of the research project and distribute a one-page study information sheet that included my contact information and instructions on how to “opt-out” of participating. Because the project did not require instructors or students to engage in activities that significantly departed from the regular labor of the course, my university’s Institutional Review Board deemed that the study posed little to no risk to students or instructors and did not require an informed consent document.⁷ Second, instructors invited their students to upload a document written in the later part of the semester—a decision motivated by the assumption that many courses assign a lengthy writing assignment due at the end of the course. Students were not compelled to participate in the study. If they did not want to be involved they could choose not to upload a document or, if uploading was already required by the course, they could ask for their documents to be omitted from the archive. Finally, instructors were asked to “enroll” me into their Oncourse site with administrative privileges which allowed me to download student writing into a digital archive located on a secure university server.

To test the protocol’s viability, I ran a pilot study that collected data from introductory- and senior-level courses taught in one semester. Only one of the five courses included required students to submit their end-of-semester writing through Oncourse while the others invited students to submit documents electronically for the sake of the study. At the start of the semester, 89 total students were enrolled in these courses and roughly half (n=43) submitted documents for the study.⁸ Collectively, these students submitted 164 documents including essays, short stories, poems, and reflections on the semester. Not surprisingly the course that required submissions had the highest student participation (n=17 of 22 or 77%) and the second highest number of documents submitted (n=66). The results of the pilot suggest that the protocol was successful from a technical standpoint; instructors were able to allow me access to their courses, and I was able to download student writing. Logging into a course and downloading all of the student files took less than five minutes per course. The

results also suggest that the protocol has limitations when it comes to student participation. Simply inviting students to participate did not yield a fully representative sample. To collect such a sample, more direct collaboration with instructors is required to create a teaching/learning situation that highlights the role of the LMS as described below. Since this pilot, electronic submission of student writing via Oncourse has increased, in part, due to the surge in online courses where every assignment is submitted electronically and, in part, due to what seems to be an increasing familiarity with the campus LMS.

I want to stress that using the LMS as a data collection tool was an appropriate choice because this networked computing software 1) is integrated into the everyday work of teaching, 2) includes functions that support collecting student writing, and 3) serves my larger research goal of assessing student learning. Given different parameters, an LMS might not be the ideal networked computing option for data collection. Teacher-researchers must assess their local contexts before committing to a particular technology for conducting research on student writing. That said, I argue that the principles outlined here might be applied to a variety of circumstances and networked computing software. For example, an institution seeking to study student writing but lacking an LMS might decide to use email for data collection. Email data collection can be as simple as asking students to submit an assignment via email to a teacher-researcher or as sophisticated as asking students to email an assignment directly into a folder located on a cloud storage platform. A discussion with local IT support will likely uncover more varied and powerful options for using networked computing than I am able to outline here. Such a discussion will be useful so long as it attends to the teacher-researcher's goals and local contexts.

Creating a Digital Archive: File Structuring

When I use the term “digital archive,” I refer to any secure, deliberately organized collection of electronic files compiled for the purposes of documentation or research. Digital archives might

be physically located on the hard drive of a laptop or the disk array of a file server or in the memory of a USB flash drive or any other device that can hold electronic media. The challenge of creating a digital archive is not in obtaining the technology to store files, but structuring it in such a way that it promotes future examination. In practical terms, this means organizing folders and naming files based on research goals. I admit that the topic of data structuring is not the most exciting one, but purposeful organizing and sorting can make the difference between a useful archive and a frustrating, digital mess.

For the pilot and the data collection that followed it, my file structuring scheme was informed by my research goal of assessing student learning using campus and department outcomes statements. These statements identify two sets of competencies: the competencies students should master by graduation and the competencies students should practice on their way to graduation. The second set of competencies are distributed over the four *levels* of courses (100 level, 200 level, etc.) that roughly correspond to the four years students are taking classes. I created a system of folders that echoed the importance of development over time by storing student writing according to the year, semester, course, and section in which it was produced. Figure 1 shows an iteration of this organization. Given a different research objective, a different file structure might be more appropriate. For instance, if a teacher-researcher is following a cohort of students enrolled in the same class over the course of a year, she or he might use each student's ID as the foundation of the structure and use assignment and draft numbers as subfolders as illustrated in Figure 2.

Name	Kind
▼ Research	Folder
▼ 2016	Folder
▼ Semester1_2016	Folder
▼ ENG101_5678	Folder
ST01_critical_5678.docx	Word
ST02_critical_5678.docx	Word
ST03_critical_5678.docx	Word

Figure 1: Proposed File Structure Emphasizing Time

Name	Kind
▼ Research	Folder
▼ Student01	Folder
▼ Assignment01	Folder
S01_A01_Draft_01.docx	Word
S01_A01_Draft_02.docx	Word

Figure 2: Proposed File Structure Emphasizing Assignment Drafts

Creating a Digital Archive: File Naming

In this section, I discuss ways of increasing data granularity—the concentration of identifiable details or *grains* of information in a system—using a file renaming scheme. The file structure I used in my digital archive offers a coarse granularity by introducing four grains of information into the system: Year, semester, course, and section number. Knowing that I was going to collect a large corpus of files, I wanted to develop a finer data granularity in order to facilitate research on student learning over time. To do this, I developed a file renaming scheme to incorporate key data into the identifier of each piece of student writing. The files I downloaded from Oncourse were named according to their writer’s preference and featured names like “EnglishPaper4,” “poemCrystal,” and “AB_Resume” that offered little in the way of systematically identifiable information. I renamed them using readily available metadata. Briefly, *metadata* can be thought of as

“data about data” or information that describes other information. The information affixed (often literally) to library books is a useful example of metadata used for organization; each book in a library is tagged with encoded information about its subject, author, and year of publication to aid sorting and searching. For many of the same reasons, I renamed each file entered into the archive based on its author, its content, and the section in which it was produced. Put differently, my formula for file renaming was Student ID + File Contents + Course Section Number. Using this formula, a rhetorical analysis written by Andy Buchenot in English 101 section 5678 would become “ST01_critical_5678.docx” in the archive. To protect privacy,⁹ student ID’s were used in place of names. To assign content, I skimmed each piece to determine if it was “critical” for expository and analytic essays, “creative” for fiction and poetry, “reflective” for pieces that examine a student’s own experiences, or “other” for outliers such as résumés or genre collages.

As with file structuring, file naming should be designed to serve a teacher-researcher’s goals. The scheme presented above was designed to help me collect student writing in an effort to find evidence of student learning at various course levels. A research project with a different goal would necessarily use a different file naming scheme. A teacher-researcher following student writing produced in one class over multiple drafts might use a file naming scheme based on Student ID + Assignment Number + Draft Number. Files named in this way would help a researcher quickly find multiple iterations of the same assignment completed by multiple students.

Automated Processes to Support Research

Simply renaming files and organizing them into a digital archive opens up possibilities for broad, automated analysis. For instance, using a file manager like Finder or File Explorer, I can search through the metadata contained in the file name to get a broad sense of the student writing assembled in the archive. A search for “critical” shows how many analytic/expository essays I have

collected and gives a partial indication of how many are being assigned in the department. Adding Boolean operators, I can create more complex searches that will tabulate how many critical essays I have collected in a particular class or at a particular course level. The same principles can be used for other research projects (How many files do I have from Student 01? For how many assignments have I collected at least three drafts?) so long as the appropriate metadata has been included in the file names.

The digital archive also opens up possibilities for fine grained automated analyses of the contents of student writing. A familiar example of this kind of analysis is word processing software like Microsoft Word that can produce quantitative data about word count, average sentence length, and assign a readability score for single documents. A more sophisticated example is corpus analysis software like WordSmith Tools which can return data on word frequency and concordance across multiple documents. It is beyond the scope of this article to fully survey the automated options for researching digital writing, but digital archives might be extremely useful for supporting the kinds of analyses provided by these software. While quantitative analysis can produce intriguing representations of student writing, I contend that they are most useful when used in tandem with human produced assessments. Knowing that, on average, essays in a 300-level contain more words than essays in a 100-level course is interesting, but it is not actionable information unless it is examined alongside an analysis of the essays' content.

The Database as Research Tool

Incorporating additional metadata further increases data granularity but requires a more complex system than deliberately naming files and folders. To handle this additional metadata, teacher-researchers might explore using database management software to handle the computational heavy lifting of tracking information and connecting it to the relevant files. When I use the term “database” in this article, I really mean “relational databases,” a common type of database comprised of a series of tables

containing related but different information. Database software builds relationships across the information contained in these tables allowing a user to find patterns among diverse data.

The database I have created for my research on student writing is comprised of several tables of metadata about the documents, courses, and students described in the digital archive. I have one table (for clarity's sake, Table A) comprised entirely of information about the documents contained in the digital archive. Each row of the table starts with a unique document code and each column that follows contains metadata about that document including the ID of the course for which it was written and the ID of the student who wrote it among several other pieces of metadata. I have another table (Table B) comprised of information about the courses referenced in the digital archive. Here, each row starts with a course ID and each column follows with metadata about the course (when the course was held, whether the course was held in-person or online, etc.). Both tables contain a shared piece of information, the course ID, but use this information in different ways. In Table A, it describes a document and, in Table B, it is a unique identifier. The database management software I use, Microsoft Access, creates relationships between the information in these two tables and allows me to conduct complex searches through the reams of metadata. Using the database management software, I can search out all of the documents created for English 101 in the fall terms of 2014-2015, for example. I could run the same search in the digital archive itself using a file manager, but the database management software conducts the search much more quickly. The database management software also allows me to run searches that would not be possible to conduct simply using the archive. As an example, I recently conducted a search for critical essays written by English majors during their senior year. To find this information, the database connected several tables and hundreds of pieces of metadata to create an orderly list of 174 documents that met my criteria. I took that list to the archive (to help protect privacy, I only connect the database to metadata about student

writing and not the student writing itself) and assembled my list of desired files.

The decision to use a database was informed by my interests in assessing student learning using institutional outcomes statements. To assist in this process, I wanted to design resources that would allow me to gather collections of student writing with shared characteristics that could be assessed by a team of trained readers. For example, I could create a collection of student writing produced for 300-level courses and, then, assess these documents based on the 300-level student learning outcomes. With the database, I am able to isolate a greater number of variables in the collections I create. Continuing the previous example, I can learn something about the influence of online teaching on student learning in 300-level courses by creating two collections of student writing—one from in-person sections and one from online sections—assessing them, and then comparing any differences in the outcomes. This kind of systematic study of student writing is what Graham and Harris call for when they urge us to move away from “teaching lore” and toward “high quality intervention studies” that rely on a systematic analysis of teaching and learning (93).

The decision to use a database was also informed by the material demands of my local context. I was fortunate to have use of Microsoft Access through my university, but there are a variety of free options for constructing relational databases including MySQL Workbench and LibreOffice Base, both of which have readily available tutorials and support communities online. There is not a universal database management program; I used Access because it could be used to serve my research goals and was already integrated into my university’s server. Teacher-researchers should evaluate their institutionally available technologies before making any software decisions. Part of evaluating institutionally available technologies is making connections with IT specialists. I benefitted immensely from the support of a knowledgeable, patient IT staff whose help informed the shape of my database as well as my digital archive. Whether

you are creating an expansive database or just a modest archive, developing a good rapport with institutional technology staff is a crucial step.

Using a database to index a digital archive is a fairly advanced version of conducting research on student writing. This level of complexity is not necessary to conduct high quality research. As described above, simply structuring and naming files with an eye toward research opens up possibilities for research. Much can also be accomplished by entering metadata into a spreadsheet or even just a table. My point is not to champion the database as the only way to conduct research, but to argue that it is one particularly fruitful tool.

Continuity and Flexibility

As it continues to define communication in the twenty-first century, digital writing will likely change in form and content as technologies develop. Consistent organization within a digital archive can account for some of these changes by providing a foundation for continuing research. A hierarchical file structure based on research goals, for example, will stretch to accommodate changes in the preferred file format while still affording insight for research. In practical terms, a .docx file might be tagged with metadata and entered into a digital archive in much the same way that a .doc file might. So long as the file structuring scheme is followed, these new file types can be folded into existing research. However, some new forms of student writing pose more challenges for established archives. For instance, student-produced websites, blogs, and other networked texts do not exist as individual files stored on a single computer. Instead, these texts are comprised of multiple files stored across many computers. These multiple, distributed texts are far harder to corral into an archive than a discrete .docx file uploaded to an LMS. There are options for capturing these network texts in a digital archive including uploading direct links to the web-based content, capturing static images of the page, and copying the HTML mark up that defines the form of the text. There is not a single best

practice for this kind of data collection, but a teacher-researcher should be prepared to develop new protocols for incorporating texts in ways that serve her or his overall research goals.

It is worth brief mention that, even accounting for normal data degradation, digital files benefit from a longer lifespan than their paper counterparts. Routine back-ups further increase any archive's longevity creating options for digital archives to become useful sources of institutional memory. The digital archive is a resource that, potentially, stretches beyond an individual research project.

Recommendations and Next Steps

Given the mercurial nature of digital technology, it is risky to make recommendations tied to specific software or hardware for fear that it might go the way of MOOs, MUDs, and Myspace. In the recommendations that follow, I focus more on principles that might be applied to a variety of technologies and institutions. My use of LMS, digital archives, and database management software is a strong model, but teacher-researchers should be ready to adapt that model to fit their local conditions using the principles below.

Articulate Research Goals

As explained above, my research goal is to use student writing in service of assessing student learning as defined by the outcomes in departmental and institutional statements. The influence of this research goal can be seen in every step of my research methods, from data collection to metadata indexing. Admittedly, operating under my particular research goal closed down as many opportunities as it opened up. My focus on end-of-semester documents meant I did not develop a method for examining the drafting process, for example. Conducting any kind of research means making such choices in order to create a cohesive project. In projects that involve digital technologies, these choices also shape the ostensibly neutral tools that we use to work with student writing. Of course, no tool is truly "neutral" as it is a product of a particular set of assumptions and values, a point

Andrew Feenberg makes eloquently in his discussion of the philosophy of technology (5). My larger point is that teacher-researchers should not lose sight of research goals as they navigate the complexity of digital technologies. There will be give-and-take between research goals and technological possibilities, but the overarching research goals should be a foundational part of any project to use digital technology to study student writing.

Identify Commonly Used Functions that Support Research Goals

As mentioned above, the time and effort involved in conducting research on student writing is a significant barrier preventing teacher-researchers from taking on such projects. Thankfully, we find ourselves in a moment where the incorporation of networked computing and digital writing into the work of teaching have made research on student writing more possible. To make the most of these possibilities, teacher-researchers should assess their local conditions in order to marshal the appropriate technology. This means investigating what technologies are routinely used at an institution and which of the technology's functionalities might be used to support research. At my institution, for instance, our LMS is used frequently to communicate with students and, increasingly, to collect student writing. At other institutions, an LMS' primary use might be to continue in-class conversations on an online bulletin board or to host student blogs. At yet another institution, the only networked computing option might be the Google for Education suite of applications. In each of these situations, teacher-researchers should identify overlaps between their research goals and the available functionalities in order to develop protocols for using technology to study student writing.

Encourage Faculty “Buy In”

Research on student writing often starts in a single teacher-researcher's classroom, but, to conduct the kind of research that Graham and Harris and others recommend, it is necessary to gain

a wider frame of reference by securing support from colleagues. One way to engender support from colleagues is to involve teachers and administrators in the process of designing a research goal. This might mean everything from sending out email to holding informal meetings to preparing a formal proposal. What is important is to make your colleagues into stakeholders, into individuals who have something to gain from spending time adding another instructor into their course homepage or explaining to their students the importance of research. It won't be possible to involve everyone in this way, but it is an admirable goal to reach out to as many as possible. Regardless of individual instructors' involvement in the planning stage, it is crucial to limit the time and energy they are asked to participate in a research project. Data collection, for instance, should be as integrating into the everyday labor of teaching writing as possible. This goal can be achieved by using technologies that other instructors are already using as much as possible. The less labor instructors must take on to participate, the more likely they are to contribute their time.

Develop File Structuring and File Naming Schemes

A digital archive is only useful if it is organized in a way that lends itself to being searched and analyzed. When I undertake the slightly dull process of creating folders and renaming files, I am reminded of new media theorist and rhetorician Karl Stolley's advice that "file naming and organization is essential to keep yourself sane" (45). Stolley is writing about creating directories for websites, but the lesson translates nicely: An organized set of files is always easier to use and far less maddening than a disorganized set of files. This advice becomes especially true as the number of files in an archive stretches beyond 30, 50, 100, or 1,000. It also bears repeating that the utility of the organization depends on the goals of a research project. For my goal of creating targeted samples of student writing, a hierarchical structure starting with year and ending with course section was sufficient. For a project examining a small group of students over several years (i.e., the kind of longitudinal studies that Lindquist and

others are presently conducting), an organization starting with a student name and moving toward specific years and then documents might be more useful. In all cases, the organization and file naming should be internally consistent and supportive of a project's goals.

Consider the Future Applications

My final recommendation might be read as an extension of my first. Teacher-researchers should start the process of conducting research on student writing with a specific interest in mind—be it research questions or a curricular development or anything in between. That interest should shape the project design from the software used to store files to the protocols used to collect them. In the same way, a project that has started producing data should be improved and augmented based on the possibilities these data suggest. Research on student writing does not end in a conclusion; it ends in new questions, new teaching strategies, and new initiatives that invigorate teacher-researchers.

Acknowledgements

This project was supported by a grant from the Indiana University-Purdue University Indianapolis (IUPUI) English Department's Turner-Davis Boundary Crossing Award as well as a grant from IUPUI's Program Review and Assessment Committee (PRAC). I would also like to thank William E Stuckey and Thomas Upton for their support in the development of the archive and database.

Notes

¹ See the National Council of Teachers of English's *The Teaching-Research Connection*, the Conference on English Education's *Understanding the Relationship between Research and Teaching*, and the Two-Year College English Association's *Research and Scholarship in the Two-Year College* among many others.

² In his review of composition scholarship that uses student writing, Joseph Harris contends that Lu, Miller, and Tobin's works present "arguments in which the meanings of student texts matter—and are very much open to debate" (676).

³ Details about a study are necessary to produce what Richard Haswell calls RAD research—research that is replicable, aggregable, and data supported (198). Haswell rightly argues that RAD research is crucial to growing the field of writing studies (201).

⁴ Data on student computer use suggest that the majority of students in secondary schools have ample opportunities to produce digital writing. A 2015 Pew Research study reports that "87% of American teens ages 13 to 17 have or have access to a desktop or laptop computer" and 73% have access to smartphones (Lenhart n.p.). Students at colleges and universities report even higher rates of computer access. According to a report published by Educase in 2014, 90% of students own a laptop, 86% own a smartphone, and 47% own a tablet (Dahlstrom and Bichsel 14).

⁵ Since publishing this definition, the WIDE collective has renamed themselves as the Writing, Information, and Digital Experience research center. More about their current iteration can be found here: wide.cal.msu.edu.

⁶ See Bruno Latour's "Morality and Technology: The End of Means" for an engaging discussion of how our goals are adapted by the technical processes we undertake: "If we fail to recognize how much the use of a technique, however simple, has displaced, translated, modified, or inflected the initial intention, it is simply because we have *changed the end in changing the means* (252 original emphasis).

⁷ Policies on what kinds of research require review vary. In many cases, research on student writing only needs review if it is going to be disseminated outside of the immediate educational context (e.g., a conference presentation). Regardless of the circulation of the results, all studies involving human subjects (students, teachers, members of the community) should at minimum be discussed with local research offices.

⁸This initial enrollment figure might not reflect the final enrollment. When students drop a course after the start of the semester or simply fail to finish, they might still appear on the Oncourse roster.

⁹My collaborators and I are the only ones able to see these file names. When student writing from the archive is shared with others (usually in the form of paper copies), all identifying features are removed and a second coded name is assigned to further protect privacy.

Works Cited

- Arola, Kristin L. "The Design of Web 2.0: The Rise of the Template, The Fall of Design." *Computers and Composition* 27.1 (2010): 4-14. Web.
- Brandt, Deborah. *Literacy in American Lives*. Cambridge University Press, 2001. Print.
- Carroll, Lee Ann. *Rehearsing New Roles: How College Students Develop as Writers*. Carbondale: Southern Illinois University Press, 2002. Print.
- Conference on English Education. "Understanding the Relationship between Research and Teaching." 2005. Web.
- Connors, Robert J., and Andrea A. Lunsford. "Frequency of Formal Errors in Current College Writing, or Ma and Pa Kettle Do Research." *College Composition and Communication* 39.4 (1988): 395-409. Web.
- Dahlstrom, Eden, and Jacqueline Bichsel. "ECAR Study of Undergraduate Students and Information Technolog." Oct. 2014. Web.
- Eyman, Douglas, and Cheryl E. Ball. "Composing for Digital Publications: Rhetoric, Design, Code." *Composition Studies* 42.1 (2014): 114-17. Web.
- Feenberg, Andrew. "What Is Philosophy of Technology?" University of Tokyo. 2003. Lecture. Web.
- Graham, Steve, and Karen Harris. "Conducting High Quality Writing Intervention Research: Twelve Recommendations." *Journal of Writing Research* 6.2 (2014): 89-123. Web.
- Gray, L., Thomas, N., and Lewis, L. "Teachers' Use of Educational Technology in U.S. Public Schools: 2009." *National Center for Education Statistics (NCES 2010-040) Institute of Education Sciences*, U.S. Department of Education: Washington DC, 2010. Web.
- Harris, Joseph. *A Teaching Subject Composition Since 1966*. Logan, UT: Utah State University Press, 2012. Web.
- . "Using Student Texts in Composition Scholarship." *JAC* 32.3-4 (2012): 667-94. Web.
- Haswell, Richard H. "NCTE/CCCC's Recent War on Scholarship." *Written Communication* 22.2 (2005): 198-223. Web.

- Horner, Bruce. "Students, Authorship, and the Work of Composition." *College English* 59.5 (1997): 505–29. Web.
- Lenhart, Amanda. "Teen, Social Media and Technology Overview 2015." Apr. 2015. Web. 24 Oct. 2015.
- Lindquist, Julie. "Time to Grow Them: Practicing Slow Research in a Fast Field." *JAC* 32.3-4 (2012): 645-66. Web.
- Lu, Min-Zhan. "Professing Multiculturalism: The Politics of Style in the Contact Zone." *College Composition and Communication* (1994): 442-58. Web.
- Lunsford, Andrea A., and Karen J. Lunsford. "'Mistakes Are a Fact of Life': A National Comparative Study." *College Composition and Communication* 59.4 (2008): 781-806. Web.
- McKee, Heidi A, and Dànielle Nicole DeVoss. *Digital Writing Assessment and Evaluation*. Logan, UT: Computers and Composition Digital Press: Utah State University Press, 2013. *Open WorldCat*. Web. 11 May 2016.
- . "Preface" in *Digital Writing Assessment & Evaluation*. Eds. Heidi A. McKee and Dànielle Nicole DeVoss. Logan, UT: Computers and Composition Digital Press/Utah State University Press, 2013. Web.
- Miller, Richard E. "Fault Lines in the Contact Zone." *College English* 56.4 (1994): 389-408. Web.
- Murray, Donald. "Teach Writing as a Process Not Product." *The Leaflet* 71.3 (1972): 11-14. Web.
- National Council of Teachers of English "The Teaching-Research Connection." 2005. Web. 23 Oct. 2015.
- Perelman, Les. "Construct Validity, Length, Score, and Time in Holistically Graded Writing Assessments: The Case against Automated Essay Scoring (AES)." *International Advances in Writing Research: Cultures, Places, Measures*. Fort Collins, Colorado: WAC Clearinghouse, 2012. 121-32. Web. 24 Oct. 2015.
- Porter, Jim. "Why Technology Matters to Writing: A Cyberwriter's Tale." *Computers and Composition* 20.4 (2003): 375-94. Web.
- "Principles of Undergraduate Learning." N.p., n.d. Web. 11 May 2016.
- Shaughnessy, Mina P. *Errors and Expectations: A Guide for the Teacher of Basic Writing*. New York: Oxford University Press, 1979. Print.
- Sommers, Nancy, and Laura Saltz. "The Novice as Expert: Writing the Freshman Year." *College Composition and Communication* 56.1 (2004): 124-49. Web.
- Sternglass, Marilyn S. *Time to Know Them: A Longitudinal Study of Writing and Learning at the College Level*. Mahwah, N.J.: Lawrence Erlbaum Associates, 1997. Print.
- Stolley, Karl. *How to Design and Write Web Pages Today*. Santa Barbara: Greenwood, 2011. Print.

- Tobin, Lad. *Reading Student Writing: Confessions, Meditations, and Rants*. Portsmouth, NH: Boynton/Cook/Heineman, 2004. Print.
- Two Year College English Association. "Research and Scholarship in the Two-Year College." 2010. Web. 23 Oct. 2015.
- Writing in Digital Environments (WIDE) Research Center Collective. "Why Teach Digital Writing?" (2005): N.p. Open Access. Web." *Kairos* 10.1 (2005): n.p. Print.