Project-Based Learning for Graduate Students in Digital Humanities

Thomas Augst¹, Deena Engel¹

¹ New York University

Funding: The author(s) received no specific funding for this work.

Potential competing interests: The author(s) declared that no potential competing interests exist.

Abstract

This essay reports on a five-year summer internship sponsored by the Graduate School of Arts and Sciences at New York University that sought to apply computer science pedagogy in project-based learning [PBL] to digital humanities training of graduate students from diverse humanities disciplines and programs. While much student training in the field tends to occur in academic courses or workshops devoted to particular tools and methods, this program used the iterative process of project development to design an inclusive, efficient context for graduate students with limited experience with technology to learn digital humanities skills appropriate to their professional and scholarly objectives. Describing the framework of PBL computer science pedagogy, the essay considers the technology learning objectives of a broad variety of projects undertaken by 50 MA and PhD students from disciplines ranging from English and History to Fine Arts and Linguistics. Emphasizing the role of peer-learning and cultural differences between STEM and humanities learning contexts, the essay draws on the program coordinators' teaching experience and student commentary to assess learning outcomes of a PBL approach for the professional and scholarly development of humanities graduate students.

Introduction

With the growth of research across disciplines applying computational tools and methods and employing digital forms of communication, universities have developed a variety of resources and opportunities for training graduate students in the digital humanities. Many of these opportunities have tended to be course-based, offering training with particular tools and methods, or take the form of one-time or multi-day workshops hosted by libraries, professional organizations, student groups or other campus entities such as humanities centers. This essay describes a different model: a summer internship program in digital humanities created by the Graduate School of Arts and Science at NYU, which began in 2015 and concluded in the summer of 2020. The program represented a five-year experiment in the application of computer science pedagogy to project-based training of graduate students across diverse humanities disciplines and programs. By reviewing this program's pedagogy and assessing its outcomes, we seek to recognize broader lessons this program may have for training humanities graduate students in computational tools and methods across different learning contexts.

From the outset, the chief goal of this program was to offer inclusive, project-based training in digital humanities tools
and methods for graduate students, in ways that would advance their professional horizons both within and beyond the academy. The program was made possible by the generosity of the Polonsky Foundation, which contributed funds to support seven paid internships for graduate students in the humanities per year, as well as funding supervision by faculty. Faculty were drawn from both the English and the Computer Science departments in an effort to role-model collaboration between a Humanities and a STEM department. For each of the program's five summers, a group of seven students was joined by an additional three graduate students from the Department of English, whose internships were supported by NYU alumnus and former trustee, Kevin R. Brine. Given the dedicated resources for the English Department, interns from the English Department represented 35% of the students participating in the summer workshops. (See Figure 1.) The next largest cohorts of students came from the Institute of Fine Arts, the Museum Studies program, the History Department, and the Archives and Public History Programs, with the remainder of the interns representing an additional ten departments or programs ranging from Near Eastern Studies and Classics to French and Linguistics. (A full list of all of the students and their bio's along with a brief description of each project can be found on the Graduate School of Arts and Science website at New York University (“Internships”, 2020). The program was open to graduate students in MA as well as PhD programs in the Humanities: 44% of the interns were MA students and 54% were from PhD programs. A total of fifty students participated in this internship program over the course of five years.

![Figure 1: Projects by Academic Department](https://doi.org/10.32388/4W5TXN)

The projects interns worked each summer demonstrated a variety remarkable for the chronological breadth of...
cultural heritage materials being explored, as well as for the diversity of methodologies that define scholarly inquiry in the humanities in our contemporary moment. Following a formal application process, selected students and their faculty mentors or project supervisors met with the program's faculty coordinators to define project goals and learning objectives for each internship. Students received a summer stipend of $5,000 each and were expected to devote 300 hours in research and training for a digital humanities project. Students were also required to participate in a project development workshop that met over lunch every two weeks over the summer with the program's faculty coordinators; these meetings included occasional guest instructors from NYU Libraries and other departments in the humanities. As we shall see with the technology-based learning goals and challenges shared by interns, project development opens opportunities for intellectual discovery and peer-exchange across disciplines otherwise limited by departments, programs, and subfields of specialization.

A major goal of this program was to leverage graduate students' dedicated training and compensated time to advance digital humanities research by faculty at NYU. A third to one-half of accepted proposals each year were for internship placements of students across a range of faculty-led, grant-funded projects, including: Saving Data Journalism (Boss et al., 2019) to capture and preserve data journalism projects in a scholarly archive; Shadowlines (Kisin, 2018), a database and mapping of the circulation of Indigenous made objects into collecting institutions; Digital Aponte (Rodriguez et al., 2017), dedicated to the life and work of an artist and antislavery activist in colonial Cuba; digital enhancements for fieldwork documentation for an archeological excavation of Sanam Temple (“Sanam Temple Project”, 2018); David Wojnarowicz Knowledge Base (Wharton et al., 2017) and the Joan Jonas Knowledge Base (Wharton et al., 2020) of the Artist Archives Initiative (Wharton and Engel, 2019) to build online open resources for research on contemporary art among others. External institutional collaborations also hosted students working on projects at art museums including the Solomon R. Guggenheim Museum, the Whitney Museum of American Art and SFMOMA (San Francisco Museum of Modern Art); the Weeksville Heritage Center, an historic site and cultural center in Central Brooklyn; the Chinatown Tenants Union, a pan-Asian community-based organization in New York City; and The New York Society Library, whose City Readers (Bieck, 2016) offers on-line exploration of digitized early circulating records. A substantial number of doctoral student interns pursued independent projects related to their scholarly research and publication.

In describing the diversity of digital humanities projects advanced by the summer internship program over a five-year period, this essay seeks to offer a critical reflection on needs and opportunities for graduate training in the humanities at this time. By assessing patterns in the experiences and outcomes of these internships, we also seek to outline tactics of project-based learning that can efficiently and effectively adapt training in the use of computational methods and tools to diverse scholarly and professional goals in the humanities. In the pages that follow, we describe specific areas of technology learning that interns engaged in during the summer internships, and we seek to assess major outcomes of the internship program for graduate training. We also report on what students say that they found most beneficial in this program, and the many ways that their projects benefitted from the learning goals they initially brought to the internship. We seek to illustrate the spectrum along which learning digital tools and methods can advance a student’s professional development in the humanities. Our experience demonstrates that a pedagogical focus on project development, combining individualized tutoring in technology with peer-sharing, can facilitate efficient and effective paraprofessional training that widens access to data literacy while also advancing STEM-Humanities collaborations in building 21st century
forms of scholarship.

Project-based Learning (PBL) and Computer Science Pedagogy

A search on the keywords Project-based learning in the ACM Digital Library (ACM Digital Library) reveals a range of resources on using a Project-based Learning (PBL) approach throughout a variety of topics in Computer Science and for a number of student audiences. As early as 1999, instructors wrote that “The computer science (CS) educational community has recently realized the potential of project-based learning (PBL) in CS education...” (Scherz and Polak, 1999). Other instructors have credited PBL with students’ improvement in critical thinking for software development courses (Güney et al., 2019), with student success in a compilers course Lara and Quesada, 2019 and the flexibility of PBL to support diverse student groups (Thevathayan, 2018), among others.

Many of the facets of PBL as developed in the Computer Science pedagogy field within the framework of the classical education model apply to the diverse digital humanities student body (Veselov et al., 2019). PBL encourages course development that supports “different types of classes, such as traditional lectures, time in the lab, group meetings and design discussions to enable student voice, reflection, critical thinking and critique, essential elements of PBL that are commonly not sufficiently addressed in traditional course organizations” (Lara and Quesada, 2019). Charles Thevathayan at Royal Melbourne Institute of Technology (RMIT) in Australia, described PBL for software engineering students in this way (Thevathayan, 2018):

PBL allows abstract software engineering principles and practices to be learnt experientially making the course more appealing to diverse student cohorts. Employers favor PBL as students get exposed to current industry practices, processes and tools, thus narrowing the industry-academia gap. PBL however, poses a number of challenges. Academic staff need to find efficient ways of managing many additional tasks related to PBL such as getting licenses, developing technical materials, managing projects and handling team issues. Students involved in PBL need to be exposed to teamwork and project management skills concurrently.

PBL can be effective in a number of fields (What is PBL?, 2020). In this section of the essay we explore how to leverage Computer Science pedagogy around PBL to provide a framework for graduate students in the humanities to take on projects that require them to learn and implement new technological skills. PBL can fill an important role in supporting students in bridging the gap between the training offered by academic disciplines and their roles and responsibilities outside of the academy. Through practical challenges of designing and executing projects in digital humanities, students gain opportunities to see their work through life-cycles of scholarship within a rapidly changing digital media ecology. Through the process of project iteration, students learn to appreciate content-area expertise typically acquired through an often-solitary process of research discovery.

The internship program asked students to undertake sustained, independent work that enhanced their skill-sets in the use of digital tools and methods, and facilitated their exploration of scholarly and professional interests outside their familiar course structures and scholarly disciplines. Through informal peer discussion that occurred in the first hour of
each bi-weekly meeting of the summer workshop, students acquired experience explaining the particular significance of the humanities content of their respective projects to a cross-disciplinary audience, while sharing practical challenges in the application of technology to the design, implementation, preservation, and publication of digital scholarship. The major technology-driven learning needs over the five years of the program were seen in the areas of web publication and preservation (26%), Geographic Information Systems (GIS) tools and mapping (24%); working with text (20%), working with datasets (12%), and programming (8%), among others. (See Figure 2). A project-based learning model offers flexibility in training across diverse computational tools and methods, tailored to specific professional and scholarly learning objectives of individual students.

![Students' Primary Learning Goals Upon Entering the Program](image)

Figure 2: Projects grouped by the student's primary learning goal upon entering the program.

Project-based Learning (PBL) and Digital Humanities Pedagogy

There are “seven essential project design elements” of the “Gold Standard” for PBL (What is PBL?, 2020). These elements, along with one more (Lara and Quesada, 2019) constituted the framework for the organization of the program's pedagogy, and offer structure as well as flexibility in supporting students’ independent learning and assessing their progress across a diversity of humanities projects. Following is a list of those elements and how they are addressed in the Polonsky-Brine internship program:

1. **Project-based learning begins with a challenging problem.** As part of their application statements, students accepted
into this program were asked to describe their background and interests in digital humanities, and to describe a significant question or problem that they wish to research and/or implement during the course of their summer internship. Prior to the student application deadline, faculty from the Faculty of Arts and Sciences were invited to submit descriptions of digital humanities projects that could be advanced by an internship placement, and encouraged to identify and recommend a graduate student for the project. Most MA students accepted into the program applied to work on faculty projects, having coordinated their applications with program mentors in advance. The remainder of successful applications to the program were from doctoral students seeking to advance digital humanities skill-sets related to their doctoral research interests. The scholarly rigor was evaluated first by the selection committee and later through a preparatory meeting in April or May with the student, his or her faculty or project advisor and the Internship program faculty coordinators.

2. **Sustained inquiry:** Students work on the topic of their choice throughout the program which allows for an iterative process of research and development. Students learn to assess and make choices in managing their project work, weighing timely pursuit of project objectives against the exploration of tools and methods. If traditional forms of scholarship teach students to use the research process to efficiently narrow questions and conclusions regarding its significance, the iterative process of learning computational tools and methods learning typically requires revisions of questions and objectives that define the project's mission.

3. **Authenticity:** The project is not simply "an assignment" but is rather a scholarly study of value to the student's discipline, either through his or her dissertation or through digital humanities projects. By seeding or extending work on multi-year, faculty-led digital humanities projects, the internship program facilitated institutional collaborations in the training of next-generation scholars and arts professionals while supporting faculty to further their own research and professional goals.

4. **Student voice and choice:** Students select and design the project as well as their approaches, selecting and using computational and/or digital tools to complement their scholarly research and professional interests. Empowering humanities students to be critical and creative users of technology in their research and teaching requires building confidence and authority to make informed choices about tools and software. The bi-weekly cohort meetings played a significant role in supporting students as they made decisions and voiced concerns about their projects.

5. **Reflection:** The students reflect on their projects and their progress independently; with their project or faculty mentors; with the other students in their cohort; and with the program faculty advisors. Peer-sharing about the iterative process of project-development encouraged students to reflect on and discuss the broader variety of digital humanities scholarship tools, methods and goals. Near the end of the internship program, students were asked to write 3-5 page reflections on their learning experience; although not intended for publication, these essays created a space for students to assess how their projects had evolved, to articulate the learning objectives they had accomplished, and to discern the broader lessons of their experiences for their research trajectories and professional development.

6. **Critique and revision:** Students met regularly with their faculty advisors, in most cases weekly, to assess the project status, evaluate the student’s work and to plan the next steps. Students also gave thoughtful feedback to others and listen to feedback in turn during the workshop meetings with the cohort throughout the summer. Many students met independently in small groups when they had technical or project development issues in common.
7. **Public product:** The result of each student’s work was made public in ways consistent with project goals at the outset such as through his or her dissertation; through their contribution to a larger project they have participated in; to a digital publication or web presence and/or a conference presentation. All of the students present their findings publicly in the early fall at a campus event dedicated to this purpose. In the later weeks of the summer workshop, interns rehearsed presenting their projects in brief 7-10 minute presentations prior to the “project showcase” that occurred in September following completion of the internships.

8. **Success Skills:** Students participated in “group work, self-management, [and] collaboration” (Lara and Quesada, 2019). Real-world projects in digital scholarship bring responsibilities, choices and consequences about uses of time and investments of energy, especially for collaborative projects in which a graduate student intern has partnered with a faculty member or may be the latest addition to a team, charged with project duties for a limited period of time.

### Technology learning outcomes for graduate students in the humanities

Each summer of the program the independent work of interns on their project was supported by bi-weekly meetings of the cohort in a project development workshop. The purpose of these one- or two-hour workshops was to introduce major technology topics and relevant terminology; realistic expectations were discussed so that students did not expect to master a skill in this short time-frame but rather were encouraged to explore the topics. For example, each summer all students were asked to participate in a workshop on GIS software which was conducted through the NYU library. Students also participated in a workshop on working with data taught by one of the Computer Science faculty mentors covering topics such as examples and definitions of data formats (XML, CSV, JSON, etc.); data curation; terms and definitions when working with a relational database and how to prepare data for use in an SQL environment along with an explanation of the role of relational databases in software that students are already familiar with such as wikis, WordPress and other content management systems.

The Computer Science faculty also provided one-on-one meetings or in small groups to help students with programming needs and review technical issues. These meetings were designed to serve both pedagogical and research goals, teaching programming, web development and SQL through a specific problem that a student wished to address. We did not undertake a specific programming class in this program but rather worked with students individually based on the level they were at and their project needs. In several cases, the Computer Science faculty built simple Python scripts with students which the students then worked on further to extend, modify and develop as their projects progressed; this approach was especially productive for students working with unstructured data such as doing textual analysis and learning how to manage files in Python as well as for students doing statistical work and/or cleaning texts programmatically. The Computer Science faculty offered an SQL workshop for students working with structured data using **DB Browser for SQLite** (Piacentini, 2020) described as “for users and developers who want to create, search, and edit databases”. The workshop examples were tailored to meet the students’ needs; for example, a student working with European Medieval manuscripts was encouraged to download a dataset from the Metropolitan Museum of Art’s github site (“The Metropolitan Museum of Art Open Access CSV”, 2020) so that he could apply the concepts of SQL queries against a single table working with data that he understood, and he could therefore identify how well his queries were working by the results that he retrieved.
The computer science faculty also provided training on CSS as well as support in learning tools such as WordPress to assist students working in web development to customize their projects. We found that teaching these topics to humanities students worked well using the principles above of PBL: the students were eager to learn for the purpose of their projects; they shared their learning experiences and encouraged each other; and were able to set realistic goals. This approach shielded humanities students from some of the frustrations and disappointments that can otherwise accompany the iterative process of project development. For example, one student was working with 18th century texts in Ottoman Turkish but struggled to find a way to use Python to obtain the results that he sought. A computer science faculty mentor encouraged the student to look into literature on this topic in technology journals; it soon became clear that textual analysis of 18th century texts in Ottoman Turkish is a known computational problem that has not yet been satisfactorily solved for this student’s purposes. The student shifted his project focus accordingly to accomplish more feasible goals for summer work, having acquired a fuller understanding of the computational challenges and opportunities in his field of historical language research.

A project-based approach that emphasizes peer-learning among a small seminar-size group of students lowers cultural and disciplinary barriers that can inhibit technology learning in the humanities. By adopting a model of a project development workshop, the internship program supported individual students through limited training in technologies, scaled to specific research and professional goals. The project-focused tutorials in technology were combined with more expansive exposure to tools and methods of broad relevance that occurred during workshop meetings, and by watching other interns develop their projects. As students learned from the progress and detours of their own respective work as well as those of others in their cohort, they learned to trust their own authority outside of the typically specialized disciplines of humanities graduate training, and to articulate their shared methodological interests and project goals of digital scholarship. As student reflective essays about the internships suggest, this approach especially serves students in the humanities and others facing cultural and disciplinary barriers to learning about technology. Hands-on practice with appropriate support in executing a project can build confidence in one or more specific skill-sets in digital humanities, while exposing students to a diversity of methods, tools, and projects through their cohort that they would be less likely to encounter in a particular skills-based course or workshop.

There are a number of pedagogy principles of project-based learning in the computer science field that apply well to the groups of motivated and curious students who participated in the internship program. At the same time, there are a number of cultural differences in pedagogy and implementation in STEM fields vs. the humanities which we found helpful to address directly in our conversations and group discussions with the students. Humanities students often work alone, consistent with conventions of single-author dissertations and monographs that have long been prevalent in humanities disciplines. In STEM fields, by comparison, collaboration and sharing (the role of “open-source software” for example) are more often the norm. Given the opportunity, it is important for humanities students to learn to collaborate in their efforts and to share credit. When participating in collaborative projects for their internships, explicit parameters and goals for student work were essential for establishing in advance clear expectations. As projects evolve in the course of an intern’s work, as projects and selected methods typically do, students will be more resourceful and flexible in adapting to new goals when treated as collaborators from the outset, empowered to make decisions about uses of time and supported in acquiring training appropriate to project needs.
Humanities students were encouraged to think through answers on their own but they need to be encouraged to ask questions about technology as they often otherwise hesitate for fear of appearing "dumb". It can be useful to set a classroom “rule” e.g. “There is no such thing as a dumb question”, and gently remind students during class discussion that they may not preface their questions or comments with “This is probably a stupid question, but ...” Humanities students often need encouragement to see that learning how to shape programming or technology questions enables them to find answers, whether from colleagues or teachers or from external resources such as software and hardware documentation, both physical (printed) and online. It is important to provide reasonable tools with appropriate support and appropriate expectations for humanities students taking on tasks and challenges for which they have not been formally prepared. Digital humanities tools that are framed as “easy to use” can do more harm than good as humanities students may lack the skills to know how to learn new tools and therefore become discouraged and lose confidence in their abilities to handle computational and digital-based work. As Morgan observes, “Despite its considerable dazzle, easiness is an abstract and intangible quality. The promise of easiness, or an easy-to-use tool, is that some process (whether display, formatting, organization, or analysis) can be accomplished with minimal difficulty, confusion, or extra labor” (Morgan, 2018). It is important to provide reasonable software options such as those with a graphical user interface for computational work e.g. DB-Browser for SQLite to learn SQL or Voyant Tools (Sinclair and Rockwell, 2016), “designed to facilitate [text] reading and interpretive practices for digital humanities students and scholars as well as for the general public, “for an introduction to textual analysis. Appropriate training options and open discussion of feasible goals and timeframes will help to alleviate some of the discouragement that can otherwise accompany students' work in project development.

Project-based learning introduces students to processes of iteration, which requires open-ended experimentation with multiple tools and methods for accomplishing project goals. With an iterative process, one makes mistakes and learns from them. A master's student in English learning to customize a WordPress site found herself "afraid to touch the code or the CSS for fear of ruining a part of the site." Project iteration, however, requires that students learn to make mistakes and to learn from them. "I still fear making mistakes and being wrong, but I have come to understand that making mistakes and being wrong does not always lead to negative consequences; they can, in fact, lead us to discover the very solution we had been looking for all along.” Effective mentoring in the context of project iteration builds broader familiarity with a fuller spectrum of digital tools; it enables an appreciation of the project’s technology goals within a fuller context of information architecture, and encourages confidence in navigating a constellation of concepts and methods in the course of their work.

Learning Outcomes: Expanding Research and Professional Skills

For doctoral students early in their programs or just beginning to write dissertation proposals, the internship offered the opportunity to broaden the horizons of their doctoral research through computational methods. Graduate students in the humanities typically receive extensive training in qualitative methods of analysis, using case-studies, illustrations, and examples to develop interpretations and arguments. For MA students with a strong interest in doctoral study, the internship offered intellectual inspiration for pursuing literary scholarship with digital methods and tools. A science student who returned to graduate school in English appreciated “retaining the methodological frameworks I inherited in my
scientific training”; this student would go on to a doctoral program in English, convinced that “better understanding what
digital tools are out there and how to collect, organize, and present a wide range of data types appears to me a necessary
step in my graduate education.”

On-going collaborative projects brought interns awareness of the specialized and diverse skill sets involved in long-
term project development. An MA student in the Institute of Fine Arts on the Sunam Temple project gained experience in
curating and visualizing field data collected during archeological field work, while assisting the project director in creating
a public-facing website for the project, as well as collaborating with another project member on the design of an app for
collecting anthropological field data in the specific conditions of the Sudan where the Temple of Sunam is located. While
working on the Joan Jonas Knowledge Base, another MA student discovered that “It is beneficial to learn and maintain
tools that may not be required in your particular field because you may never know when they will be needed.”

**Working with data**

The tools of digital scholarship transform students' understanding of the humanities content they are working with, as
data. Many projects focused on gathering, curating, analyzing and building visualizations of datasets related to each
student’s specific project and goals. The concept of data is new for many humanities students and the process of
preparing data, analyzing data and rendering data visualization provides powerful tools for research with wide-ranging
applications in diverse fields of professional employment. Humanities students can be encouraged to consider nuances in
the data and to bring attention and skills to evaluating qualitative aspects of the data and to think through non-obvious
questions e.g. which data are missing? Who (or which communities) are not represented in this dataset? What is the
humanistic impact of these numbers? At the scale of corpus building, data visualization, or archive preservation, students
learn to conceptualize texts, historical artifacts, art objects and images as digital data that can be analyzed, mapped,
preserved, and formatted for research discovery and scholarly re-use. By learning to curate the data produced by their
research process, and structuring it according to commonly used standards, interns gained access to a range of
computational tools and methods permitting quantitative approaches to humanities research. Software such as MySql for
database management, Tableau for data analysis, and GIS platforms (Qgis, Arcgis online, and Arcgis Story Maps)
enabled doctoral students to incorporate quantitative data into qualitative research, opening novel connections and
insights across humanities content areas. Given the rapid obsolescence in computational tools and approaches, the
problem-driven process of learning to format and visualize humanities research as data allows scholars to experiment
with methods of analysis and communication. “Since much of our research outputs are expected to remain qualitative,” as
a doctoral student noted about his internship experience, “project-based work affords an opportunity to engage in a more
experimental, comparative approach to methods of analysis, at a moment when the tools for such analysis are
themselves undergoing rapid change”.

A large number of summer projects focused on creating and visualizing geo-spatial data, reflecting the growing
interest in spatiality in humanities scholarship, the so-called ‘spatial turn’, as well as the general prevalence of GIS
technology in our daily lives. A Master's student in history mapped locations in mid-19th century New York City that
participated in and supported the by-then-outlawed American slave trade. This process involved extracting geographic
information from a wide range of existing scholarship on the slave trade, structuring this information in consistent format,
and visualizing it on both nineteenth-century and contemporary maps of the city. A Master’s student in English used the internship to gather geographic data as part of her on-going research into a 19th century manuscript diary kept by an evangelical minister in 1820s Manhattan. After transcribing a volume of the diary, the student created a dataset of addresses the minister visited during his ministry among sailors and their families. Learning how to map these addresses using ArcGIS, the student created a StoryMaps Project (“Classic StoryMaps”) that explores the social and spatial implications of building local religious communities (Fenster, 2019); the practice she acquired with structuring and mapping spatial data gave the student a new way of reading texts. Gaining experience using cartographic software while examining databases related to the city of Baltimore, a PhD candidate in English linked these digital tools and skills to public humanities and activism, working with local artists and community organizations to produce maps of “infrastructure inequities” around Baltimore. As this student later observed, “These skills have opened up a host of research, teaching, and professional opportunities, as the fellowship gave me the confidence to pursue more digital humanities projects and gain literacy in new softwares and programs. Currently, as I apply for positions at various public humanities institutions, I constantly tout my experience working in digital humanities.”

Tools for data curation and mapping allow students with little previous experience with computational tools to discover new questions about research materials with which they were otherwise familiar. For many interns in early phases of their doctoral research, the iterative process of working with data opened entirely new avenues for inquiry. A student of medieval literature observed that when he "ran queries on a large body of data in a CSV file regarding the material component of an artifact, or displaying provenance data from a CSV file in a mapping program, the materiality of artifacts, and the rootedness of this materiality in particular times and spatial context could never be elided." Internship work with mapping software and data organization revealed for a doctoral student in History new findings from her research on the Bronx Zoo’s trade in animals in the early twentieth century, while StoryMaps software facilitated publication for a general audience (McLeod, 2020) and inspired a “new outlook on what my dissertation can be — it will almost certainly include some interactive digital components and be more public-facing.” Another doctoral student working on environmental history came to the internship with the intention of learning GIS tools and software for advancing her research in urban and environmental history. During her internship, she produced a digital archive in Omeka (“Omeka”) of georectified maps highlighting the historical course of urban rivers and waterways in São Paulo while building analysis of political ecology into interactive features of a bi-lingual Neatline exhibit. While experimentation with GIS tools developed the student’s understanding of digital techniques for spatial visualization, it also deepened her commitment to public history and expanded her horizons in approaching her undergraduate teaching assignments in the following academic year as she integrated GIS tools and software into her students’ coursework. Tools for spatial analysis and visualization were one of the major approaches by which students learned to work with data. Other students invested their time during the internship in transforming texts into datasets, learning to apply computational methods of reading to diverse literary and historical research materials. While subscription databases have in recent decades made huge quantities of periodicals and newspapers accessible for research, automated processes of ‘machine reading’ enable researchers to analyze textual corpora at a scale that facilitates discovery of new patterns. Word frequency study using software such as Voyant Tools for textual analysis permits graduate students without programming experience to frame questions or interrogate concepts with more fine-grained and far-reaching attention to historical phenomena of language change.
doctoral student in history used text analysis tools to research corpora in nineteenth-century United States newspapers, widening his interests beyond editorial commentary about religion to the fiction and advertisements that appeared around them. In contrast to the reliance of literary analysis on close reading, computational tools for linguistic analysis offer many approaches to discovery of textual patterns in historical texts.

Texts that humanists manipulate with such ingenuity and invention on a small scale through methods of slow reading or close reading, can be analyzed through computational methods and visualized within an ever-changing arsenal of tools. A doctoral student in linguistics wrote Python scripts to run against large-scale Twitter data in order to study a regional New York City English dialect, demonstrating how scholars can use large-scale corpora to analyze naturally occurring activity online to understand the expression of identity and attitudes to language use. A doctoral student in English participated in the development of City Readers, using metadata from digitized borrowing records to generate statistics, visualizations, and datasets analyzing the types of books women in New York City read in the 1790s. A doctoral student at the Institute of Fine Arts translated Coptic colophons in the manuscript corpus, while learning to mark up manuscript texts with TEI ("Text Encoding Initiative") and use statistical analysis tools in R to understand monastic book production in Egypt in the ninth and tenth centuries. A Master’s student in English used text encoding to represent and analyze a manuscript text from the Victorian period, contributing to a faculty member’s on-going project to digitize a collaborative diary written by two women writers under the composite name Michael Field (Thain, 2020). As that student observed, “Digital code may never capture the thrill of holding a book, but the encoding process generates a surprising intimacy with the text, one that is comparable to the close-reading practices prized by literary scholars.”

Several doctoral students from various disciplines used their internships to develop their research interests and training in their respective disciplines by learning to use computer programming languages to transform text corpora into unstructured data formats available for tools of ‘machine reading’ or ‘distant reading’. Scholars with both limited and advanced experience in programming learned to write Python scripts to analyze textual phenomena at a much larger scale than would be possible with human eyes alone. One student wrote Python scripts to wrangle a corpus of over a thousand eighteenth-century essays on economic subjects for further exploration. This student had only had a brief introduction to Python before the workshop; the Computer Science instructors developed a “bare-bones” script for the student, walked him through it slowly and in great detail and made suggestions for additional features and efficiencies which the student carried out himself with the Computer Science faculty available to answer questions. An English doctoral student with professional experience in digital text modeling, analytics and visualization developed a user front-end for “The New Fascicles”, his custom application that 'reads' the poetry left by Emily Dickinson in forty fascicle books using Mallet topic modeling software (McCallum, 2002) to discern the potential topics of these poems and their shared themes, creating an interactive, series of data visualizations for these topic models. A PhD student in linguistics created Phonological corpus tools (Szabó, 2018) for studying two under-researched Indigenous languages, one each spoken in Latin America and in Eastern Africa. This project benefits research on these two languages in particular as well as phonological research in general through the software tools that the student wrote. A third-year doctoral student in Classics used natural language processing tools in Python to create a comprehensive single corpus of digitized ancient Greek texts. Adapting emergent tools for text acquisition and data visualization available in linguistics, computer science, and data science, along with traditional methods of philology and textual criticism, the student tested methods for
analyzing and visualizing the presence and development of concepts and themes in ancient Greek dramatic texts. In doing so, she developed a methodology for her doctoral research that she described as a “robust, transparent, and reusable research process” to inform the writing of her dissertation prospectus in the subsequent semester. A doctoral student in French seeking to analyze the complex rhyme structures and textual references in a fifteenth-century work ended up not using Python, as she originally intended, and instead turned to linguistics methodology -- underscoring the importance of pursuing multiple possible solutions to one's computational goals.

Another group of students focused on reading, rather than writing, source code across several different programming languages. The conservation of time-based media art is a new specialization within the field of art conservation dedicated to conserving works of art that are digital-born such as artworks in digital video, audio or other digital media; and computer-based artworks that embody one or more computer hardware and/or software components. Practitioners and theorists in this field bring expertise and knowledge in the humanities as they approach these works in scholarly collaboration with art historians and curators. So too they cultivate technical skills required to manage the care and re-exhibition of artworks that are based on technology. NYU’s Institute of Fine Arts houses one of three full-time art conservation programs in the United States and it is the only conservation program in the United States that offers a concentration in time-based media art. Four students from this program have participated in Polonsky internships across three major U.S. art museums (the Solomon R. Guggenheim Museum in New York City; SFMOMA, the San Francisco Museum of Modern Art; and the Whitney Museum of American Art in New York City). All four students were tasked with diving deeply into the source code of the respective complex software-based artworks that they studied to produce extensive technical documentation and formal risk assessment reports along with recommendations for conservation practices and preparation for re-exhibition of these fragile works of art which are seminal works in the museums’ collections. The internships allowed the students to do in-depth studies of the source code as the foundation to the artworks and to their projects. The conservation of time-based media art is a field that inherently embodies both sides of digital humanities - the digital and the humanities - and as a result, the students in this program contributed a great deal to the class discussions as practitioners in a field that bridges science and art.

**Web Publication and Preservation**

A large number of projects each summer consisted of web development projects, advancing archival preservation in library and museum contexts, as well as website development for public engagement with scholarship. Summer internships offered students an opportunity to imagine different audiences for their research, and to model interactive, media-rich means of storytelling so readily available through web development tools. WordPress, Omeka, and StoryMaps among others offer dynamic publication environments where scholarship can find new audiences outside of traditional modes of publication, and engage those audiences more actively than with traditional forms of scholarship. During her work as researcher and editor for a WordPress research portal at the Institute fine arts, an MA student also developed a WordPress website for a survey course in Chinese Art for a course she went on to teach at Parsons. A master's student in Museum Studies created a prototype for online educational resources and an interactive map related to the living cultural heritage of the Inka Road in Peru. From customizing appearance to modeling user experience of visual and textual content, practical design challenges of website development opened new ways of conceiving the significance and
potential audiences for humanities research.

How should we measure the value or impact of scholarly activity, when it is designed for public engagement outside of academic conferences, journal articles and monographs that for so many decades have furnished the infrastructure for scholarly publication and communication? For students completing dissertations, the opportunity to publish archival sources and to situate them within critical or historical contexts on websites can produce shifts in intended audiences, expanding potential publics for scholarship. Where traditional methods of scholarly training in the humanities inevitably incline students to identify with sub-fields within larger disciplines defined by content specialization, several interns discovered through their web development projects that imagining audiences beyond content-specialists can reframe intellectual questions and professional ambitions for their work. While working on a digital companion for his dissertation about music criticism and opera in 18th century England, narrowly conceived at the outset as “an online resource for literary scholars,” one student used the design and prototyping of a WordPress site to redefine the project scope and presentation for a general audience. Preparing to undertake his first entry into the academic job market, this student observed that “This, in turn, has had important implications for my dissertation, enabling me to identify ways in which my research can more directly engage issues of contemporary concern and readers outside the scholarly community.”

By publishing their scholarship through interactive web-based projects, graduate students experimented with ways of engaging publics with research artifacts and materials, while acquiring experience in writing for larger and more diverse audiences outside of the academic processes of peer-review. Another doctoral student in English designed a prototype for a digital edition of ephemeral 17th century English pamphlets that, she argues, are significant to the history of free speech. Online critical editions of texts offer scholars the opportunity to make the content as well as the form of humanities scholarship more public. Learning about text encoding, and the design of an on-line platform “fundamentally changed my understanding of academic practice,” as she later commented: “[this was an] opportunity to engage with an interdisciplinary group of scholars and graduate students; acquire a new set of technical skills; observe an alternative model of collaborative pedagogy; and create an online resource that will be significant contribution to my scholarly credentials as I launch my academic career.” In the course of web development internships, doctoral students discovered motives and ambitions for humanities research that challenge forms of publication and validation that traditionally define and reward scholarly activity in the academy.

Occurring in the midst of the 2020 Covid-19 pandemic, the final summer of the program underscored specific needs and opportunities posed by digital environments for scholarship, unfolding in real-time. The internship gave a doctoral student in history ways to think about “how I use archives and communicate my findings from them,” and especially in the context of 2020’s Covid pandemic, helped her navigate “the immediate hurdles of an academic world gone digital,” and offered practical means “to think through what the future of my work can be in the as-yet unknown terrain of post-Covid academia.” As it became clear that the museum facility would remain closed in months ahead, an English MA student working with NYU’s Grey Art Gallery became focused on creating an engaging, interactive online exhibit featuring digital images from its collection. In the process of building website prototypes using WordPress, Omeka with Neatline, and Storymaps, her academic background in narrative had practical impact in opening the gallery’s collection to on-line discovery while in-person visits were not possible.
Professional Development

Especially for students whose graduate training entails specialization in humanities materials and discipline, the iterative exploration of computational tools and methods builds confidence in their capacity for technology learning more generally, empowering them to be more critical and informed users of software products. The process of exploring solutions for project goals inevitably takes one down paths that turn out not to be fruitful towards the project goals. As one intern observed during a discussion, “I feel like my life right now is going down wormholes.” In contrast to traditional forms of humanities inquiry, defined by academic conventions of the scholarly essay or the dissertation, project-based learning assumes that this open-ended process of exploration, encompassing multiple computational methods or sampling of multiple softwares, for example, yields valuable learning outcomes regardless of the particular product they may eventually select. While some projects did not advance beyond prototype phases, or remained on-going after summer internships ended, the iterative process of project-based learning offered significant rewards of their own. An MA student earning a dual degree in Library and Information Science and History found that the internship played a significant role in her professional employment. “My Polonsky project directly ushered me into a job in the field where I could share the skills I learned with others.” In her new role as Reference and Digital Humanities Librarian at Fordham University, this student found that her experience with “data collection, mapping, data visualization, and web development enabled me to hit the ground running giving project advice to students and faculty.”

Digital scholarship projects in areas such as archive preservation and scholarly communication provide training in the kinds of professional activity that, if less typical of scholarly activities of research and writing familiar to them, offer graduate students practical experience in team-based processes and institutional standards of humanities work across cultural heritage, museum, and library and information science contexts. An internship with a faculty member’s project A View from the View (McCormick, 2020) showcasing an archive of historical postcards of Beirut, exposed a Master’s student in Middle Eastern studies to a more comprehensive set of skills than she would have encountered in courses focused on particular tools or methods: content management systems, website design, metadata schemas, database building and cleaning, visual sources, and modeling interfaces that contextualize data for different audiences. This internship experience led immediately to other opportunities to work on digital humanities projects, while inspiring new interest in incorporating mapping into her thesis research. One internship entailed creating a digital archive of an organization’s digital-born publication, the VoCA Journal, by Voices of Contemporary Art, which is published by a “non-profit organization that generates critical dialogue and collaborative programming to advance a vital model for the stewardship of contemporary art” (“VoCA Journal”). A Master’s student in Museum Studies learned about a wide range of collection management tools including Omeka, Collective Access (“Collective Access”), MuseumPlus (“MuseumPlus”) and PastPerfect (“PastPerfect Museum Software”) along with web archiving tools including WayBack Machine (“Wayback Machine”) and Archive-It (“Archive-It”), both from the Internet Archive, along with ArchiveSpace by Lyrasis (“ArchiveSpace”), and Conifer (a new version of Webrecorder by Rhizome) (“Conifer”, 2020). So too, this project entailed extensive consultation regarding long-term preservation issues with field experts, that significantly expanded the student’s professional network in the field in which she intends to work.

Internships with a range of institutional partners allowed students to cultivate their agency outside the classroom, as they collaborated with teams of professional archivists, curators, and technologists. Summer work with on-going projects
offered students meaningful recognition of their capacity to undertake independent work, and to be trusted as a working professional in museums and archives, cultural, and library and information science, and cultural heritage organizations. The process of problem-solving and need for accountability to other stakeholders entailed by collaborative projects make them especially conducive to building professional competencies for careers outside of academia. Through on-going discussions with project teams, students acquire flexibility in communicating across academic and professional contexts, focused on advancing the public mission of diverse arts and humanities organizations and institutions.

By building shared competencies with technology among students with diverse humanities interests and training, the internship program fostered critical reflection about methods and outcomes of humanities research more generally. As interns found common intellectual ground around learning methods and tools of digital humanities, they came to appreciate the challenges of communicating across disciplinary boundaries. An intern and project directors may have initial misunderstandings, about project duties and goals that require brokering. Ongoing peer discussion and support from faculty and project advisors helped students to learn to find these gaps and address them, while helping them to navigate the variety of terminologies employed in different disciplines of professional practice. "To a humanities graduate," as one intern observed, "a “theme” is clearly referring to subject matter, while a theme to someone with a computer science background, recalls a package of graphical details." The familiarity that interns acquired with language and goals of even modest computational training bolstered the initiative for cross-disciplinary exchange outside sometimes hermetic confines of specialized humanities training. A doctoral student found that dialogues with faculty in STEM fields was conducive to conversation about forms of inquiry that was “more inclusive and less alienating than many discourses in the humanities.”

This student became more confident in dialogues with public institutions of cultural preservation that he initiated as part of exploratory work on mapping medieval manuscript collections. Over the five years of the program, the bi-weekly sharing of project experiences opened an especially significant space for cross disciplinary conversation, and one where the intimidation of undertaking computational skills could be diffused through challenges and discoveries made by peers. As a doctoral student observed, "Even though our projects were diverse, it was surprising that others were encountering similar kinds of issues and difficult decisions as I was. Seeing a problem from their perspectives helped me to question assumptions I was making [about my own.]” As one student wrote, "As humanists, we are often intimidated and shy away from computing tools because there is this sense that you have to be a complete tech person in order to understand and use them. ...by talking to one another, I lost my “intimidation” for the digital part and realized that I had the tools and resources I needed." This student went on to use digital maps in her doctoral work on ancient coin hoards, and in the successful interview for a faculty position at the University of Basel, in which she was asked about her experience with digital humanities and the tools she would use in her teaching. Through the community they shared in workshop discussions as well as in cohorts that emerged around shared interests, interns learned to collaborate and support one another in problem-solving. For a doctoral student in English building a WordPress site for a public humanities website, The Liberty Hyde Bailey Project (Linstrom, 2020), a meeting with an MA student in Museum Studies about their shared interests in web permanence for multimedia materials produced a significant shift in the purpose of the site, from hosting digital surrogates of archival materials to creating thematic exhibitions of material already available elsewhere online. By the end of his internship, this shift in project mission had helped to position him for a career outside the academy, utilizing scholarly expertise in literature and history.
to reach wide audiences with engaging content.

For many doctoral students facing an ever-shrinking market for tenure-track academic employment, project-based learning in the context of summer internships offered sustained exposure to alternative career pathways, for which most doctoral programs offer no meaningful preparation. As Katina Rogers observes, “meaningful and sustainable academic employment is an increasingly distant prospect for many doctoral recipients, with a dwindling proportion of tenure-track jobs available to an ever-growing pool of graduates” (Rogers Putting 6). In this challenging climate for traditional tenure-track academic employment in the humanities, project-based learning supports the efficient acquisition of para-professional skills that empower graduate students to recognize the value of their scholarly training to employment outside the academy.

Conclusions

This program benefitted from grant support that allowed graduate students to invest substantial time in project-based learning, without delaying progress-to-degree; so, too, it provided resources to support faculty coordination and technology mentoring. Even without such resources, however, the project-development model employed by this internship program could be adapted to a range of other learning contexts, both formal and informal. Many graduate programs offer course credits for independent study, and so too libraries on many campuses offer technology training. With advance planning across departments, coordination of existing resources, and engagement of faculty in needs and opportunities for professional training in emergent tools and methods of digital scholarship, the outcomes achieved at NYU through the Polonsky program could be realized in other learning contexts through the application of PBL pedagogical principles: customized tutoring tailored to a particular project; peer support for cross-disciplinary learning, and teaching that models STEM-Humanities collaboration in teaching of computational methods. Our experience also demonstrates the benefits for faculty that come with practice-based learning with real-world research projects. Project-based learning facilitates collaboration between faculty and students, and can form the basis for partnerships with both on-campus and off-campus organizations and institutions.

We found that the application of project-based learning in selected and relevant computational and digital skills to graduate students in the humanities, was successful for both students and faculty and the institutions with whom they worked. Both M.A. and Ph.D. students benefit from “real-life” projects that prepare them for the job market after graduation and Ph.D. students further benefit from the opportunity to gain skills that directly support their research and professional training. Students were exposed to computational applications outside of their disciplines but devoted their time to their own work and honing skills relevant to their respective disciplines and goals. Faculty mentors benefitted from this program through the interns’ endeavors as well as from the infrastructure and discussion with the humanities and computer science faculty, introducing them to formats and methods of digital humanities scholarship with which they may have been otherwise unfamiliar. Graduate programs and university research benefit from the successful completion of visible projects and publication of work in the field of digital humanities. Ongoing collaboration between humanities and computer science faculty will continue through co-teaching and co-mentoring projects in digital humanities and will continue to lead to advances in pedagogy for graduate as well as undergraduate students in the humanities.
Acknowledgements

The authors would like to thank the NYU librarians who provided DH workshops: Zach Coble and Andrew Battista; Joe Versoza in Computer Science for joining us and mentoring students; and Grace Afsari-Mamagani, Agathe Jarczyk, Craig Kapp and Jared McCormick for reading this manuscript and providing thoughtful feedback.

References

- Morgan, Paige C. “The Consequences of Framing Digital Humanities Tools as Easy to Use.” College & Undergraduate


