

Documenting Digital Projects: Instituting Guidelines for Digital Dissertations and Theses in the Humanities

Roxanne Shirazi and Stephen Zweibel

Dissertations and theses with interactive digital components seldom fit neatly into the institutional format requirements designed for traditional humanities texts. This creates challenges for students, administrators, and librarians who are charged with preparing these works for library deposit. While disciplinary acceptance of digital dissertations in the humanities may be increasing across institutions, little attention is given to the mechanics of documenting and submitting such projects. Readers, also, are challenged to find and interpret digital projects that may not be entirely described in the accompanying paper. To address this, the authors examined a set of digital theses and dissertations at their institution to determine how these digital components might fit into traditional manuscript formatting guidelines. This article introduces the resulting set of local documentation guidelines for digital dissertations and theses aimed at improving access, preservation, and reproducibility.

Introduction

In 1937, Kate Turabian created a set of guidelines that adapted the *Chicago Manual of Style* for use by graduate students in the preparation of their dissertations and theses.¹ Her role as dissertation secretary at the University of Chicago is legendary, and her work has resulted in a uniform format for dissertations across the United States that includes prefatory pages mimicking those found in book publishing. The functions of these pages are many; taken together, they succinctly draw out relevant information about the manuscript to facilitate its long-term research use. These pages—or front matter—serve as a sort of shorthand for assessing and previewing the contents of a manuscript. Yet they are relics of the print age, having been designed to help navigate through large paper documents.

Graduate students in the humanities are engaging in novel methods of scholarship, with a growing emphasis on digital tools and multimodal presentation.² These developments are resulting in graduate works that are “born digital” but lack a clear path for submission to the library for preservation and access. The singular nature of these works often requires bespoke preservation solutions, but other aspects of the workflow could be standardized, particularly the documentation required. What information about digital projects should libraries be capturing and clearly

*Roxanne Shirazi is Dissertation Research Librarian and Stephen Zweibel is Digital Scholarship Librarian at The Graduate Center, CUNY; email: rshirazi@gc.cuny.edu, szweibel@gc.cuny.edu. ©2020 Roxanne Shirazi and Stephen Zweibel, Attribution-NonCommercial (<https://creativecommons.org/licenses/by-nc/4.0/>) CC BY-NC.

identifying when students deposit digital dissertations and theses? What would constitute a *digital* front matter, designed to efficiently describe the “aboutness” of a digital project?

This article describes the development of documentation guidelines that extend the functions of a publication’s print front matter into the digital sphere as part of an ongoing transition to accommodate digital dissertations, theses, and capstone projects. Faced with a local need to standardize submission information, the authors set out to establish a framework for documenting digital projects in the humanities within the confines of Turabian style.³ Working with a local collection of digital dissertations and theses, a mixed methods analysis identified broad themes that were then paired with existing manuscript guidelines to generate a set of additional preliminary pages. The analysis confirmed that each digital deposit—though complex in terms of file formats, research techniques, and goals—contained shared characteristics that could lend themselves to a single framework of documentation. The resulting manuscript guidelines were presented to institutional stakeholders for feedback, went through several rounds of revision, and were ultimately adopted as a requirement for digital project narratives accompanying digital dissertations and theses.⁴

Faculty and graduate students want to experiment with digital publishing methods in their degree programs, but the degree requirements are often enforced by administrative offices operating outside the realm of research and scholarship. The authors propose a set of required information for digital dissertations in the humanities that can help clear the administrative hurdle to the adoption of more experimental methods and modes of publication as part of graduate work. If universities are to embrace nontraditional electronic theses and dissertations (ETDs) in the humanities—multimedia projects, software applications, digital editions, and the like—it is essential that institutions adapt routine administrative processes to perform familiar tasks in unfamiliar territory.

Background

The humanities PhD job market in the United States has long been untenable; in recent years, efforts to modernize graduate training—and to expand possibilities for the dissertation requirement—have surfaced. In 2015, as part of a project on graduate training in the twenty-first century, MediaCommons published a series of articles grouped under the heading “Beyond the Dissertation as Proto-Monograph,” in which students discussed the decision to infuse multimedia and digital formats into their dissertations.⁵ The following year, the Council of Graduate Schools held a two-day workshop on the “Future of the Dissertation,” which brought together graduate school administrators, librarians, publishers, and scholars to imagine the possibilities of new formats and genres, along with the place of the dissertation in the scholarly publishing ecosystem.⁶ Later that year, the National Endowment for the Humanities would announce the “Next Generation PhD,” a multi-institution grant program to “broaden the career preparation of a PhD student beyond a career in the academy,” an effort that would undoubtedly tackle the seemingly immutable dissertation requirement.⁷

While the existence of these initiatives suggests that digital dissertations have “arrived,” they are primarily directed toward participants in the creation and evaluation of such works: students and faculty. Such initiatives fail to address issues surrounding administrative frameworks of acceptance and institutional deposit—those procedures managed by a graduate school’s Thesis Office, Registrar, and (often) the library that must be brought in line with any proposed changes to the format of a dissertation. The typical dissertation goes through

an academic review, guided by the input of a faculty committee and culminating in a formal “defense” of the work, usually an oral presentation by the student. After a successful defense, the student must then deposit the dissertation with the library. The manuscript is inspected for compliance with institutional document format guidelines and may be subjected to multiple rounds of correction before the deposit is approved and accepted by the university. This inspection may be managed directly by the library, or through a campus Thesis Office that then delivers approved theses to the library.

Library deposit is a separate degree requirement that must be completed in order to graduate. It establishes the dissertation as an original contribution to a field of study by making it accessible to researchers; to be made accessible, it must comply with a standard format for monograph publications (such as that developed by Turabian), determined by the institution. As the content of the dissertation expands to include multimedia elements, as students and departments aim to move beyond “the proto-monograph,” depositing students are faced with a conundrum: the office responsible for approving the format may enforce requirements that were designed with only print materials in mind.

For example, a digital dissertation developed in Scalar—a nonlinear electronic publishing platform in which readers follow various “paths” rather than read from start to finish—might be acceptable as a digital medium by a forward-thinking group of dissertation committee members.⁸ But when it comes time to deposit the dissertation with the library, or submit for approval to the campus Thesis Office, the author of such a work may very well be required to assign an arbitrary order to the nonlinear text for the sake of outputting to a PDF.⁹

Likewise, dissertators who include images as core arguments in their papers have found that the university formatting regime (which may insist that images be separated from the

FIGURE 1

Scalar publications allow readers to choose among several “paths” created by the author to accommodate nonlinear narratives. Note that a single page can belong to multiple paths, as indicated by the “Citations and context” box in the upper right corner.

CONDITIONS OF (IN)VISIBILITY: CULTIVATING A DOCUMENTARY IMPULSE IN THE DIGITAL HUMAN...

THE EYES AND HANDS OF LEWIS HINE

Lewis Hine and Invisible Work

I am going to guess that most of you are familiar with the work of the photographer, Lewis Hine. And probably most of you know his work with the National Child Labor Committee in the 1910s and the instrumental role that his photographs had in exposing the horrid conditions that these children were working in, whether it was industrial factories or newsboys out on the street.

The standard telling of the Hine trajectory places him as a master of social documentary photography during the reform movement at the turn of the century; witnessing Ellis Island, child labor, and immigrant communities during the 1910s; then sort of detouring through formalism in the 1920s; and peaking with the construction of the Empire State Building in 1931 before entering relative obscurity, losing his home and going on relief until his death in 1940.

Citations and context

- Step 1 of the "The Eyes and Hands of Lewis Hine" path
- Step 2 of the "Conditions of (In)Visibility: Cultivating a Documentary Impulse in the Digital Humanities" path

main text) effectively destroys their intent.¹⁰ The imposition of strict formatting requirements for dissertation deposit is part and parcel of the doctoral degree-granting process, but there is little discussion of how it might better serve the research community.

At the same time, key scholarly societies in the humanities, such as the American Historical Association (AHA) and the Modern Language Association (MLA), have introduced guidelines for the evaluation of digital scholarship, though primarily for the purposes of promotion and tenure.¹¹ The emergence of these guidelines is evidence of a changing mode of scholarship: humanities academics are conducting research and generating data that comes in vastly different formats. Such guidelines are useful for those evaluating digital work and could be helpful for creators as they prepare their work for evaluation, whether it is through a digital publication's peer-review process or examination by a dissertation committee. At present, however, the guidelines from scholarly societies do not address practical considerations, such as how to prepare the final product for transfer to a library, nor do they include guidance on collating the technical information about a project—what is commonly referred to as a project's "documentation."

Both the AHA and MLA statements emphasize the concept of medium specificity—that materials should be evaluated in their native format. Graduate students and their evaluation committees may be able to follow this recommendation in practice; but, to file the finished product with the campus library for future research use, students are frequently made to anchor the project with a PDF, with all other file formats considered "supplemental" to the work.¹² An institution may specify that digital projects at the master's level be accompanied by a narrative component discussing the background, scope, and utility of the digital work; doctoral dissertation requirements may be less prescriptive, though an extended narrative discussion of methodology is assumed.¹³ While these required narratives may function as a digital project's documentation, there are currently no universally accepted format specifications for identifying and collating key elements of a work with multiple digital components.

To complicate matters, a digital dissertation is a unique, hybrid form of digital publication that may not readily translate into the more general sphere of scholarly publishing.¹⁴ Publishers that produce such enhanced publications typically contain the work within a single online space or platform; there may be a variety of file formats and relationships that complicate digital preservation, but there is likely to be a single access point for all materials. In contrast, digital dissertations are complex digital objects with disparate elements existing across platforms and in various formats.¹⁵ For example, Amanda Visconti's 2015 digital dissertation at the University of Maryland consisted of an interactive project website, an additional dissertation website to document the project, and a GitHub repository containing project code.¹⁶ Jesse Merandy's 2019 digital dissertation at The Graduate Center, CUNY had nearly the same components, as well as a video.¹⁷ Multiple sites and source code might be linked together and visible in the dissertation website, but the narrative description (the PDF) deposited with the library might not include such basic information as project URLs and the location of source code repositories.¹⁸ In university institutional repositories and ProQuest's Digital Dissertations database, the PDF becomes the discovery anchor of digital dissertations and their digital packages, regardless of whether it is the main component of the work. For this reason, the present study is limited to documentation appearing in the PDF component of a digital dissertation or thesis.

Traditional dissertations, as monographs, follow the conventions of book formatting; the Turabian style that is required across universities is a derivation of the *Chicago Manual of Style*

and is intended to help students prepare their work for future research use.¹⁹ The insertion of a formal title page, copyright statement, table of contents, list of figures, and so on provides a structural layer that helps transform a student “paper” into a work of scholarship with enduring research value. The front matter works in tandem with the back matter—appendices, references, glossaries, and the like—to comprise a scholarly apparatus.

Familiar sections, such as the table of contents and list of illustrations, add a navigational structure to the work but also, importantly, preview its contents and pull out key information to help researchers assess whether the work will be of use. According to the *Chicago Manual of Style*, “the criterion for when to include a list of illustrations is whether the illustrations are of intrinsic interest apart from the text they illustrate.”²⁰ Tables and figures are crafted with labels and notes to communicate research results even in the absence of the surrounding text; they are then listed out in the front matter of a book (or dissertation) because of their intrinsic value as research objects. What could such a system of documentation look like in the hybrid publication sphere of digital dissertations, which references not just the narrative text but the digital components too? How does one quickly identify digital elements that bear intrinsic research value?

To approach the problem of documenting digital projects in an established academic style, it is first necessary to review what constitutes a digital dissertation project and how libraries and archives have approached the preservation of such materials.

The Digital Environment

Like all scholarship, a digital academic project needs to provide answers to basic questions of methodology: *How was this done, and why? Is it replicable?* A digital application depends on its execution, its ability to run on a computer; a digital project that cannot run for an interested scholar is like an unreadable text. Beyond execution, digital projects need documentation—the context and instructions necessary to be fully understood and evaluated by researchers. Too often, digital projects appear as black boxes, with results that are visible and interesting, but underlying processes that are hidden from view.²¹ For scholars who wish to critically evaluate, work with, or build upon a digital project, such a presentation is incomplete.

Here it might be useful to establish some definitions and explanations of formats to situate some of the challenges of digital preservation. A digital dissertation may take many forms—a website, a software application, a collection of datasets, a collection of text- or image-based files, to name a few. To ensure access, libraries encourage the use of open-standard (nonproprietary) file formats, which are more likely to remain accessible in the future; these include PDF for documents, .txt for text, .csv for tables and numerical data, FLAC for audio, TIFF or JPEG2000 for images, and MPEG-4 for video. Indeed, for the purposes of access and preservation, a website is itself just a collection of text and image files. Its ability to run depends on its preserved format; programs and applications become lost to us when they have been created in a proprietary file format that is no longer operable.

A “platform” is an environment in which a program can run. A platform could be a piece of software or a website; what makes it a platform is its ability to run another piece of software. A software “application” is a kind of unit most of us are used to working with, such as a web browser, word processor, and the like. A software “library” is a collection of scripts used by programmers to build various applications. Software libraries are created such that they can be incorporated by subsequent programmers without their having to understand all

its details. For example, a software library that makes it easier to show images on a website could be used by another programmer in the development of a digital book.

Peculiarities of the applications used to collect or analyze data may call into question the validity of a data-driven digital project, necessitating attention to issues of reproducibility in digital projects.²² Yet file types can become inaccessible over time: for example, the company that makes a closed-source application used to develop a digital project could go out of business, leaving users who wish to run the project at a loss. In the end, libraries may not be able to ensure the continued operability of student digital projects, but we are charged, nonetheless, with maintaining the work as a scholarly contribution and record of accomplishment.

Literature Review

The question of how to preserve multimodal digital projects and make them accessible from a digital preservation standpoint is beyond the scope of this article, which approaches digital dissertations primarily as *research objects*, not simply digital objects. Educopia Institute's *ETD Plus* project (2014–2017) and the earlier *Lifecycle Management of ETDs* address the management of complex digital objects but are designed to train students, administrators, and librarians in best practices of digital curation. The projects contain a toolkit and "guidance briefs" with information on copyright, file formats, version control, and so on, but neither attends to how this information gets represented in the project narrative or PDF.

The library literature around ETDs broadly coalesces around two areas: collections management and scholarly communication. With regard to the former, topics such as metadata application and interoperability between repositories,²³ digital preservation,²⁴ and repository infrastructure are commonplace.²⁵ The latter area comprises publication issues such as the copyright status of legacy works,²⁶ access embargoes,²⁷ and, more recently, integrating scholarly communication instruction into graduate education.²⁸ Broader examinations of ETDs may look at the roles of institutional stakeholders,²⁹ or survey the prevalence of ETD programs more generally.³⁰ Few publications address digital dissertations specifically; as a call for proposals for a forthcoming edited volume notes, "While digital dissertations have been around for twenty years or more, the precise processes by which they are defined, created and defended remain something of a mystery."³¹ Rather than formal publication, it seems that this information has been confined to blog posts, conference panels, and webinars.³²

This article is primarily concerned with documentation of the digital components of a multimodal work. Conceiving of digital dissertations in the humanities as hybrid publications composed of disparate print and digital components, the authors looked to the digital humanities for inspiration. Somaya Langley suggests drawing on documentation conventions from the performing arts, such as technical riders, stage plots, and input lists, which outline instructions and requirements for staging music and theatre performances. Additionally, practices in the arts such as "documenting around the work" through video recordings of interactive installations may be appropriate for digital humanities projects.³³

Rockwell et al. describe in detail the considerations necessary to deposit an archival version of a digital project in an institutional repository, noting the four main aspects of their project that were captured for preservation: content, code, process, and the user experience.³⁴ Interestingly, they did not attempt to deposit a working version of the project, preferring instead to provide source materials outside their native format, alongside documentation that would convey the interactive experience without strictly recreating it:

Access is not recreation, and over the long term the chances that someone can recreate the hardware and software platform on which an installation could work will approach zero. We are better off giving them something they can understand and re-implement, if needed, than something they can't install.³⁵

To document the user experience, Rockwell et al. chose to write a narrative description of the project that included relevant screenshots and deposited this as a PDF with the project data files. This begins to resemble the unique publication contexts of digital dissertations and suggests a broader utility of efforts to formalize digital project documentation for student works.

Methods

To identify what information within a digital dissertation should be extracted and collated into a narrative's front matter, the authors performed a qualitative content analysis on 29 ETDs deposited at our library during the period 2014–2019. The goal of this study was to obtain and analyze data on the types of digital projects submitted to the library to inform the design of an additional set of preliminary pages for digital dissertations and theses.

Given that digital dissertations are complex digital objects that may represent a range of disciplines and scholarly conventions, a mixed-methods approach was devised to examine the ETD sample and synthesize a range of data to address an issue of institutional policy. It was apparent that the frequency of project types (a quantitative approach) would be useful to prioritize our documentation efforts, but a qualitative approach to identifying categories or themes across projects would allow latent information to emerge from the deposits. Rather than count and analyze the individual file types present in each digital project using a deductive analysis, as a digital preservationist might, this study attends to ETDs as research contributions; as such, an inductive coding scheme was applied to a more expansive unit of analysis.

The abstract, submission PDF, supplemental files, and corresponding metadata for dissertations and theses containing multiple digital components formed the unit of analysis for the study, a unit identified here as "digital deposit." Unfortunately, the library's data on digital deposits is incomplete and spread across multiple library systems. The Digital Commons repository provides a content inventory report that identifies submissions containing supplemental files. Deposits with digital components (a total of 129) were identified, but several of them were simply copyright permission letters included with traditional dissertations. After removing these records, this data was then cross-referenced with a spreadsheet of student consultations on digital preservation maintained by the digital services librarian, which rendered a total of 29 digital deposits for our study.

Each digital deposit was then manually coded using an inductive coding scheme (see table 1) to identify the most common categories and content types of digital project submissions. The coding scheme used in this analysis deploys the controlled vocabulary for research objects contained in the Taxonomy of Digital Research Activities in the Humanities (TaDiRAH).³⁶ This taxonomy includes a broad variety of research activities, objects, and techniques used by humanities researchers and was therefore deemed appropriate for the purposes of this study. Notably, however, there are no use notes or definitions for each term, which allowed the present study to inductively develop such applications; only the terms themselves were used. Due to the manageable size of the sample, the authors coded the deposits together simultaneously over the course of two sessions, in conversation with each other until terms were in

agreement. This method was sufficient as the research aim was not generalizable knowledge but local application. During each coding session, the application of the coding scheme led to further refinement of the scheme in an iterative process. For example, a discussion between coders regarding WordPress websites revealed a deficiency in the scheme. WordPress is a content management system and can be considered itself a software application. The tendency of students to make use of WordPress to render a simple website, however, led to the distinction between websites coded with object type “code” versus more complex websites that were also coded with “interactive.”

Deposits were coded using as many terms as applicable. A simple website presenting a visual memoir was coded with two object types: code and images. A more complex website, presenting analytical essays and data visualizations of a textual corpus, was coded with five object types: code, data, images, interaction, and text-bearing objects.

TABLE 1
Coding Schema Using TaDiRAH Vocabulary

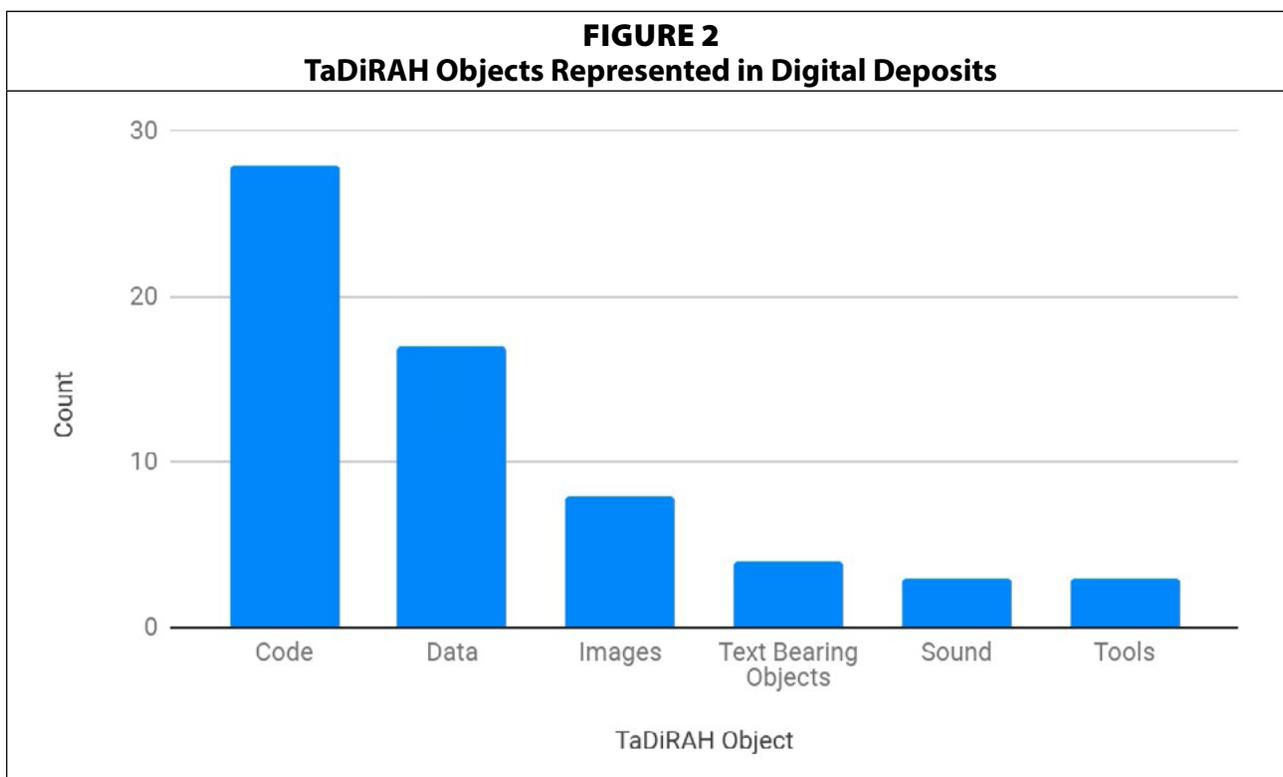
TaDiRAH Research Object	Description
Code	Any use of a programming language or mark-up language
Data	A collection of facts or points of research to be analyzed or interpreted
Sound	Audio file containing sound emissions
Tools	Complete applications or programs that users can manipulate
Images	Visual artifacts
Text Bearing Objects	Digital objects such as a digital edition or textual corpus, where the textual content is primary
Interaction	Contains a programmed response to user input, such as a visual representation that can be manipulated by users
Infrastructure	Part of the structure that allows for other works or services to exist
Video	Moving visual images
<i>Note.</i> TaDiRAH = Taxonomy of Digital Research Activities in the Humanities. http://tadirah.dariah.eu/vocab/index.php	

The data were then analyzed alongside existing prefatory pages to determine what key information about a digital project could be efficiently documented in the narrative. The authors consulted the 8th edition of Turabian and the *Chicago Manual of Style* (17th edition).³⁷ Following this, the authors developed a prototype set of additional front matter sections and explored its application in three representative digital deposits to inform the final product.

Findings

A frequency analysis of the coded deposits indicated that almost all deposits had software code as part of the project. Data were also widely used, while images and text-bearing objects were somewhat less common. Figure 2 displays the distribution of object types represented in the 29 digital deposits examined in the study.

Examining the functions of Turabian’s preliminary pages revealed four broad categories of documentation that would serve digital deposits in an expanded front matter (see table 2). The first is a simple inventory, like a packing list, of the project components since a large



Note: A total of 29 digital deposits were coded with as many terms as applicable.

number of digital deposits contained multiple object types. Second, a note discussing technical requirements would supplement the inventory list and provide more context to readers. Requiring that students include a freeform space addressing what they feel to be important high-level information about the project mitigates any restrictions that might be introduced by conforming to a strict Turabian-inspired style. Third, for works reliant on software code or data, a list of important variables or terms used in the work would aid legibility. And fourth,

TABLE 2
Mapping New Guidelines to Turabian Style

New Guidelines	Turabian	Characteristics
Digital Manifest	Table of Contents	An inventory of items included in the deposit. Gives a high-level preview of the content by tying together scattered elements.
Note on Technical Specifications	Editorial Method	A prose explanation to readers. A direct meta-conversation with the reader about the content. A flexible space for extra information that might not fit anywhere else.
List of Variables	List of Abbreviations	A list of key terms with description. This requirement forces writers to document their work according to best practices. Serves readability and accessibility.
Glossary of Functions	Glossary	Terms and definitions of key functions. A quick reference to key points allowing for understanding of main ideas. Helps orient readers who may be unfamiliar with the text's subject.

Note. Turabian front matter sections are referenced from Turabian, Kate L. *A Manual for Writers of Research Papers, Theses, and Dissertations: Chicago Style for Students and Researchers*. 8th edition. Chicago: University of Chicago Press, 2013.

for applications and interactive websites and the like, a glossary of functions could help explain the inner workings of the project.

Discussion

The aim of this research was to aid in the design of analogues to traditional dissertation front matter sections. The present study led the authors to create a “Digital Manifest” and “Note on Technical Specifications” as required elements of the standard front matter for manuscripts accompanying digital projects. Additional, optional sections to be considered include a “List of Variables” and a “Glossary of Functions.” Students are also explicitly instructed to remember to cite any software, code, and corpora that informed their work and are referred to Alan Liu’s crowdsourced guide.³⁸ The final guidelines are reproduced in the appendix. Future work will include mapping the information solicited by new front matter sections to appropriate metadata fields, to facilitate faceting by a variety of new access points based on technologies and software used in digital dissertations.

Between 2016 and 2018, the prototype was presented to institutional stakeholders including colleagues in the library, the campus Digital Initiatives group, and graduating students to gather informal feedback, and a broader professional audience was targeted through conference presentations in the ETD and digital preservation communities. Comments from all groups led to multiple rounds of revision before formal guidelines were implemented in May 2018. Since then, the application of these guidelines on new digital deposits and continued professional conversation has generated further data that will inform future iterations.

The venture described in this article demonstrates the need for further research in this area. The most recent cross-institutional analysis surveying the landscape of digital dissertations (within the broader context of ETD programs) and the administrative requirements placed on them was released in 2010.³⁹ How prevalent are these works among library ETD collections? Little information is available on the number of students creating digital dissertations in the humanities, the platforms and technologies they use, and their experiences of depositing these works with their institutions.

Conclusion

Throughout the United States, dissertations and theses adhere to near-identical document format guidelines—a format that was designed to allow graduate works to achieve visibility alongside published scholarly works. The dissertation, particularly, has been constructed as a proto-publication experience for graduate students, and graduate schools and libraries across the country still enforce manuscript submission requirements that are themselves relics of the print age. Dissertations that incorporate digital components present distinct challenges for administrators and librarians, both in terms of access and preservation. This article has argued that attending to the formatting and administrative minutiae that accompany such works is an overlooked but necessary step toward adopting digital methods and multimodal publications in graduate programs. The guidelines developed here can serve as a bridge between the traditional dissertation format and all of its attendant routines and conventions, and newer, more experimental work.

As efforts to modernize the PhD experience abound, it is incumbent on libraries and graduate offices to establish administrative procedures that provide the requisite room for a variety of publication media without abandoning the systems that have been in place for more

than 80 years. There may well be a future in which fully digital dissertations are the norm, and schools no longer impose requirements that were established primarily for print materials. Until then, however, libraries must consider how to facilitate the incorporation of digital project work into legacy processes. The guidelines put forth here are a step in that direction.

Acknowledgments

The authors wish to thank Lisa M. Rhody and Matthew K. Gold for their generous and thoughtful feedback on early versions of the guidelines. We also thank the anonymous reviewers whose helpful comments greatly improved the manuscript's structure and clarity, and Jen Zweibel for her excellent eye and copyediting skills.

APPENDIX.

Format Requirements for Project Narrative or White Paper

Most digital projects designed to fulfill the requirements of a degree include a project narrative, white paper, or accompanying essay(s) in addition to the digital component(s). If your department has such requirements in place, follow the library's format guidelines for textual dissertations, theses, and capstone projects when preparing your document. Additionally, include sections specific to digital projects as discussed below.

Preliminary Pages, or Front Matter

In addition to the requirements in place for all dissertations, theses, and capstone projects, the following sections are specific to works including digital or online components. Required pages must appear in the page sequence outlined in the next section.

Abstract

All dissertations, theses, and capstone projects must include an abstract; there is no length limit. The abstract will appear online with the descriptive information (title, author, program, and any other descriptive information) associated with the work and may be reproduced and indexed in a variety of academic research databases and Google Scholar. The abstract must also be included within the deposited manuscript as page iv of the preliminary pages.

For digital projects, abstracts should describe the project scope and include relevant URLs for associated elements such as videos, websites, or code repositories (such as GitHub link); if applicable, describe what data has been collected.

Digital Manifest

All dissertations, theses, and capstone projects that contain digital projects must include a "Digital Manifest" in the preliminary pages. Formatted like a Table of Contents, this section provides a master list of all the components—print and digital—that, together, constitute the dissertation, thesis, or capstone project *as it was submitted to your program for approval*. List and briefly describe the project components to form an inventory or "packing list" for the deposit. For each component included in the deposit, indicate its file type, a brief description, and URL, if applicable (see figure 1). For example, a capstone project that is a website containing geospatial visualizations created using CARTO might deposit: an archived version of the website (submitted as WARC files), data files exported from CARTO (as CSV, Shapefile, KML, GeoJSON, or SVG files), and an accompanying white paper. All of these items would be listed in the Digital Manifest.

List of Variables (if applicable)

If the digital project is a database or includes code, a List of Variables must be included in the preliminary pages or as an appendix (back matter).

If brief, the List of Variables should be located in the front matter, or preliminary pages; lengthier sections are more appropriate as an appendix. Format according to Turabian instructions for lists of abbreviations, presenting information in two columns and in alphabetical order (see figure 2).

Glossary of Functions (if applicable)

If the digital project consists largely of code, a Glossary of Functions must be included in the preliminary pages or as an appendix (back matter) to document the uses of functions and code. Formatted like a traditional glossary of terms, the Glossary of Functions should list important functions included in the software code alongside a brief explanation of what each function does. Format according to Turabian's instructions for Glossary (see figure 3).

A Note on Technical Specifications

All dissertations, theses, and capstone projects that contain digital projects must include a "Note on Technical Specifications" in the preliminary pages. Use this section to provide an introductory note to readers that will serve as a high-level overview of the project's components and technical specifications (see figure 4). Include here any information about components housed outside of the library deposit, such as GitHub repositories, and where to find the latest version of materials.

References

In your references section, include the platforms, software libraries, and code used in your project. These can be separated from other bibliographic citations included in your manuscript if desired. For guidance, consult the Software Sustainability Institute's recommendations (<https://www.software.ac.uk/how-cite-software>). Additional examples are collected in Alan Liu's "Citing Bits: Sources and Suggestions for Citing Software, Platforms, Code, Corpora" (2017).

Required Page Sequence for Preliminary Pages

Title Page
 Copyright Page
 Approval Page
 Abstract
 Preface, Foreword, and/or Acknowledgements (optional)
 Dedication and/or Epigraph (optional)
 Table of Contents
 List of Tables (if applicable)
 Lists of Figures, Illustrations, Charts, Diagrams, and the like (if applicable)
 Digital Manifest (required for all digital projects)
 Lists of Variables (if applicable, unless submitted as an appendix)
 Glossary of Functions (if applicable, unless submitted as an appendix)
 A Note on Technical Specifications (required for all digital projects)
 Body of Text (pagination switches to Arabic number 1)
 Appendix or Appendices
 References
 Autobiographical Statement (optional)

Sample Pages

Figure 1: Digital Manifest. This sample page shows what a Digital Manifest might look like for a digital dissertation containing a white paper, a project website, blog posts, a digital edition, and software code.

Figure 2: List of Variables. This sample page shows a List of Variables for a capstone project that consists of a software application.

Figure 3: Glossary of Functions. This page shows a Glossary of Functions for a capstone project that consists of a software application; note that only key functions are included.

Figure 4: A Note on Technical Specifications. This section may be a prose paragraph or a list of specifications, depending on the nature of the project. Note how the sample walks potential users through the process of installing the project and putting it to use; this section is analogous to a “readme” file..

Digital Manifest

- I. **Dissertation Whitepaper (PDF)**
- II. **WARC files**
 - a. **Project Website**
Archived version of <https://github.com/amandavisconti/ulysses>
 - b. **Blog posts**
Archived version of <http://literaturegeek.com> - a research blog with 31 posts (average length: 2,000-3,000 words) plus cross-postings to HASTAC.org plus 2 unique blog posts at mith.umd.edu.
 - c. **Digital Edition**
Archived version of <http://www.infiniteulysses.com>
- III. **Code and other deliverables**
Zip file containing the contents of the GitHub repository at time of deposit.
(<http://github.com/amandavisconti/infinite-ulysses-dissertation>)

Figure 1: Digital Manifest. This sample page shows what a Digital Manifest might look like for a digital dissertation containing a white paper, a project website, blog posts, a digital edition, and software code. List all components included in your library deposit. *Source credit:* Adapted, with permission, from the online documentation submitted by Amanda Visconti (University of Virginia) for “Infinite Ulysses,” a digital dissertation at the University of Maryland.

List of Variables	
ADMIN_EMAIL	The administrator's email address in the DH Box database
ADMIN_PASS	The administrator's password
DEFAULT_HOSTNAME	IP address of the host server
DEMO_ENABLED	If True, will enable the demo page of DH Box
INSTITUTION	Will display this string on the front page
LOCALHOST	If true, will point URLs towards 'localhost' rather than '0.0.0.0'
SECRET_KEY	Allows for protected sessions
SQLALCHEMY_DATABASE_URI	Location and name of the DH Box database
TESTING	If true, enables Flask's testing tools

Figure 2: List of Variables. List variables on the left in alphabetical order and the corresponding description on the right. This sample page shows a List of Variables for a capstone project that consists of a software application. *Source credit:* Adapted, with permission, from the documentation accompanying the DH Box project, submitted by Stephen Zweibel as a master's capstone project at the Graduate Center, The City University of New York.

Glossary of Functions

add_user(container, username, password): Add a user to a running DH Box.

build_dhbox(username='test'): Builds a new DH Box seed, renaming the old one if it exists. Username variable defaults to "test".

configure_dhbox(user, the_pass, email, demo=False): Use Docker exec to SSH into a new container, customizing it for the user.

create_user_and_role(): Create an admin user to test with.

delete_untagged(): Find the untagged images and remove them.

demo_dhbox(username): Make a demonstration DH Box with a random name. Expires in an hour.

download_dhbox(username='test'): Downloads a DH Box seed, renaming the old one if it exists. Username variable defaults to "test".

execute(container, args): Execute a list of arbitrary Bash commands inside a container.

get_hostname(): Determine the IP address of the host server.

how_long_up(container): Find out how long a container has been running, in seconds.

kill_and_remove_user(name, user=True): Kill a running DH Box container and remove the user if there is one.

replace_admin_dhbox_image(): Updates admin's DH Box to the new seed image.

setup_new_dhbox(username, password, email, demo=False): Create a new DH Box container, customize it.

Figure 3: Glossary of Functions. Formatted like a traditional glossary of terms, the Glossary of Functions should list important functions included in the software code alongside a brief explanation of what each function does. This page shows a Glossary of Functions for a capstone project that consists of a software application; note that only key functions are included. *Source credit:* Adapted, with permission, from the documentation accompanying the DH Box project, submitted by Stephen Zweibel as a master's capstone project at the Graduate Center, The City University of New York.

A Note on Technical Specifications

Currently DH Box requires a version of Linux that has apt-get. DH Box is a Python 2.7 application that uses a tool called Docker to create “containers,” akin to virtual computers, for users on the web. Users access DH Box tools directly from a web browser, performing statistical analysis with R Studio; creating digital exhibits with Omeka; or using MALLET for topic modeling, all without having to configure or install anything new on their computers.

To work with DH Box as an administrator, users need to install Docker and pip (Python Package Manager) before cloning the project’s git repository (<https://github.com/DH-Box/dhbox.git>) to their local computer. After following the install instructions, admin users navigate to the site on <http://localhost:80> and create relevant user names and passwords before launching their new DH Box.

Locations for the tools in DH Box are based on the IP address of an admin’s DH Box, for example xx.xxx.xxx.xxx. RStudio is located at xx.xxx.xxx.xxx:8787, that is, port 8787. Omeka is located at xx.xxx.xxx.xxx:8080, that is, port 8080. Its Super User account has the same username and password as the admin user’s DH Box. The Bash shell is accessed at xx.xxx.xxx.xxx:4200. Once there, users input their DH Box username and password to get access to IPython, MALLET, and the Natural Language Toolkit.

Figure 4: A Note on Technical Specifications. This section may be a prose paragraph or a list of specifications, depending on the nature of the project. Note how the sample walks potential users through the process of installing the project and putting it to use; this section is analogous to a “readme” file. *Source credit:* Adapted, with permission, from the documentation accompanying the DH Box project, submitted by Stephen Zweibel as a master’s capstone project at the Graduate Center, The City University of New York.

Notes

1. Penelope Kaiserlian, "Kate Turabian's Manual: A Best-Seller for Fifty Years," *Scholarly Publishing* 19 (April 1988): 137.
2. Sharpe, Celeste Truong Vy. "Digital Dissertations and the Changing Nature of Doctoral Work," *Perspectives Daily*, April 23, 2019, <https://www.historians.org/publications-and-directories/perspectives-on-history/april-2019/digital-dissertations-and-the-changing-nature-of-doctoral-work>.
3. In this article, reference to Turabian refers to the 8th edition; Kate L. Turabian, *A Manual for Writers of Research Papers, Theses, and Dissertations: Chicago Style for Students and Researchers*, 8th ed., Chicago Guides to Writing, Editing, and Publishing (Chicago, IL: University of Chicago Press, 2013).
4. Roxanne Shirazi and Stephen Zweibel, "Format Requirements for Project Narrative or White Paper" (2018), https://libguides.gc.cuny.edu/ld.php?content_id=44359564.
5. "Beyond the Dissertation as Proto-Monograph: Examples and Reflections," *#alt-Academy: A Mediacommons Project* (April 26, 2015), <http://mediacommons.org/alt-ac/cluster/beyond-proto-monograph-new-models-dissertation>.
6. Council of Graduate Schools, "Proceedings of the CGS Future of the Dissertation Workshop" (Council of Graduate Schools, 2016), <http://cgsnet.org/future-dissertation-workshop>.
7. "NEH Announces \$1.7 Million for 'Next Generation PhD,'" *National Endowment for the Humanities* (August 9, 2016), <https://www.neh.gov/news/press-release/2016-08-10>.
8. "About Scalar," *Alliance for Networking Visual Culture*, <https://scalar.me/anvc/scalar/features/> [accessed 8 March 2019].
9. This hypothetical scenario is based on a personal communication between the authors and a scholar regarding their experience depositing a multimodal dissertation.
10. Melanie Lee, "The Melancholy Odyssey of a Dissertation with Pictures," *Pedagogy* 15, no. 1 (December 10, 2014): 93–101.
11. American Historical Association, "Guidelines for the Professional Evaluation of Digital Scholarship by Historians" (June 2015), <https://www.historians.org/teaching-and-learning/digital-history-resources/evaluation-of-digital-scholarship-in-history/guidelines-for-the-professional-evaluation-of-digital-scholarship-by-historians>; Modern Language Association, "Guidelines for Evaluating Work in Digital Humanities and Digital Media" (January 2012), <https://www.mla.org/About-Us/Governance/Committees/Committee-Listings/Professional-Issues/Committee-on-Information-Technology/Guidelines-for-Evaluating-Work-in-Digital-Humanities-and-Digital-Media>.
12. See, for example: "ETD Administrator FAQ—I Want to Submit a Dissertation in a Non-PDF Format (e.g. Audio File, Video File, Etc.). How Can I Do That?—ProQuest Support Center," <https://support.proquest.com/articleDetail?id=kA01W000000k9fKSAQ> [accessed 8 March 2019].
13. "Thesis/Capstone Project," <https://www.gc.cuny.edu/Page-Elements/Academics-Research-Centers-Initiatives/Masters-Programs/Liberal-Studies/Courses/Thesis-Capstone-Project> [accessed 8 March 2019]; "Digital Dissertation Guidelines," *History and Art History*, <https://historyarthistory.gmu.edu/graduate/phd-history/digital-dissertation-guidelines> [accessed 28 August 2019].
14. Celeste Sharpe, "They Need You! Disability, Visual Culture, and the Poster Child, 1945–1980" (2016), <http://mars.gmu.edu/handle/1920/10555>.
15. Bill Donovan, "Managing the Lifecycle of ETDs: Curatorial Decisions and Practices," in *Guidance Documents for Lifecycle Management of ETDs*, eds. Matt Schultz and Katherine Skinner (Atlanta, GA: Educopia Institute, 2014), 5.1-5.20, <https://educopia.org/publications/gdlmetd>.
16. "Amanda Visconti's Doctoral Dissertation," <http://dr.amandavisconti.com/> [accessed 17 April 2018].
17. Jesse Merandy, "Vanishing Leaves: A Study of Walt Whitman Through Location-Based Mobile Technologies," *Vanishing Leaves*, <http://vanishingleaves.com/> [accessed 20 September 2019].
18. Amanda Visconti, "'How Can You Love a Work, If You Don't Know It?': Critical Code and Design towards Participatory Digital Editions" (PhD, University of Maryland, 2015), <http://drum.lib.umd.edu/handle/1903/16580>.
19. "History of Kate Turabian's 'A Manual for Writers,'" *Chicago Manual of Style Online*, <https://www.chicagomanualofstyle.org/turabian.html> [accessed 8 March 2019].
20. "3.38: When to Include a List of Illustrations," *Chicago Manual of Style Online*, 17th ed., <https://www.chicagomanualofstyle.org/book/ed17/part1/ch03/psec038.html>.
21. See, for instance, the section on interfaces in Fiona Barnett et al., "Queer OS: A User's Manual," in *Debates in the Digital Humanities*, 2016 ed., <http://dhdebates.gc.cuny.edu/debates/text/56> [accessed 8 March 2019].

22. See, for example, the discussion of retractions caused by errors in Microsoft Excel: Mark Ziemann, Yotam Eren, and Assam El-Osta, "Gene Name Errors Are Widespread in the Scientific Literature," *Genome Biology* 17 (August 23, 2016): 177, <https://doi.org/10.1186/s13059-016-1044-7>.
23. Marielle Veve, "Harvesting ETD Metadata from Institutional Repositories to OCLC: Approaches and Barriers to Implementation," *Journal of Library Metadata* 16, no. 2 (April 2016): 69–79, <https://doi.org/10.1080/1051712X.2016.1215730>.
24. Joy M. Perrin, Heidi M. Winkler, and Le Yang, "Digital Preservation Challenges with an ETD Collection—A Case Study at Texas Tech University," *Journal of Academic Librarianship* 41, no. 1 (2015): 98–104.
25. Le Yang et al., "A Complete System for Publishing Music-Related ETDs: Technology Development and Publishing Model," *Library Hi Tech* 34, no. 1 (March 9, 2016): 151–63, <https://doi.org/10.1108/LHT-10-2015-0096>.
26. Gail Clement and Melissa Levine, "Copyright and Publication Status of Pre-1978 Dissertations: A Content Analysis Approach" (July 2011), <http://oaktrust.library.tamu.edu/handle/1969.1/149190>.
27. Charles B. Lowry, "ETDs and Digital Repositories: A Disciplinary Challenge to Open Access?" *portal: Libraries and the Academy* 6, no. 4 (2006): 387–93.
28. Gail Clement and Jo Ann Bianchi, "Copyright and Publishing Literacy for ETD Authors: Applying the Theft of the Mind Model" (February 28, 2013), <http://oaktrust.library.tamu.edu/handle/1969.1/147616>; Amy Buckland, "More than Consumers: Students as Content Creators," in *Getting the Word Out: Academic Libraries as Scholarly Publishers*, eds. Maria Bonn and Mike Furlough (Chicago, IL: Association of College and Research Libraries, 2015), 193–202, https://atrium.lib.uoguelph.ca:8443/xmlui/bitstream/handle/10214/12087/Buckland_ACRLchapter.pdf?sequence=1&isAllowed=y.
29. Richard Fyffe and William C. Welburn, "ETDs, Scholarly Communication, and Campus Collaboration: Opportunities for Libraries," *College & Research Libraries News* 69, no. 3 (2008): 152–55.
30. Joan K. Lippincott and Clifford A. Lynch, "ETDs and Graduate Education: Programs and Prospects," *Research Library Issues: A Bimonthly Report from ARL, CNL, and SPARC*, no. 270 (June 2010): 6–15, <http://publications.arl.org/rli270/>.
31. *The Digital Dissertation: History, Theory, Practice* (forthcoming), eds. Virginia Kuhn and Anke Finger, <https://sites.google.com/view/cfp-the-digital-dissertation/home>
32. See, for example: "Getting from Here to There," *Scholars' Lab*, <https://scholarslab.lib.virginia.edu/blog/getting-from-here-to-there/> [accessed 21 September 2019]; "Digital Dissertations and the Changing Nature of Doctoral Work | Perspectives on History | AHA"; admin, "TxETDA Webinar Series 2019: Beyond the PDF: Evolving ETD Practices and Policies to Support the Next Generation of Student Work," *Texas ETD Association* (March 8, 2019), <http://txetda.org/index.php/2019/03/08/txetda-webinar-series-2019-beyond-the-pdf-evolving-etc-practices-and-policies-to-support-the-next-generation-of-student-work/>.
33. Somaya Langley, "Planning Your (Digital) Funeral: For Projects," *Digital Preservation at Oxford and Cambridge* (blog, October 30, 2017), www.dpoc.ac.uk/2017/10/30/planning-your-digital-funeral-for-projects/.
34. Geoffrey Rockwell et al., "Burying Dead Projects: Depositing the Globalization Compendium," *Digital Humanities Quarterly* 008, no. 2 (May 29, 2014): paras. 36–40.
35. Rockwell et al., "Burying Dead Projects," para. 43.
36. TaDiRAH, "TaDiRAH: Taxonomy of Digital Research Activities in the Humanities" (July 18, 2014), <http://tadirah.dariah.eu/vocab/index.php>.
37. Turabian, *A Manual for Writers of Research Papers, Theses, and Dissertations*; The University of Chicago Press Editorial Staff, *The Chicago Manual of Style, 17th Edition* (University of Chicago Press, 2017).
38. Alan Liu, "'Citing Bits': Sources and Suggestions for Citing Software, Platforms, Code, Corpora," editable spreadsheet, Google Docs (January 15, 2017), https://docs.google.com/spreadsheets/d/1Qc8VZDIn_N3J-HGEkk6QKPwRIGxM3DjB9hn1tIRsQA0/edit#gid=0.
39. Lippincott and Lynch, "ETDs and Graduate Education," 13.