STARS

Electronic Theses and Dissertations, 2004-2019

2015

The Practice and Benefit of Applying Digital Markup in Preserving Texts and Creating Digital Editions: A Poetical Analysis of a Blank-Verse Translation of Virgil's Aeneid

William Dorner
University of Central Florida

Showcase of Text, Archives, Research & Scholarship

THE PRACTICE AND BENEFIT OF APPLYING DIGITAL MARKUP IN PRESERVING TEXTS
AND CREATING DIGITAL EDITIONS: A POETICAL ANALYSIS OF A BLANK-VERSE
TRANSLATION OF VIRGIL'S *AENEID*

by

WILLIAM DORNER
M.A., Rhetoric and Composition, University of Central Florida, 2010
B.A., Creative Writing, University of Central Florida, 2007

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Texts and Technology
in the College of Arts and Humanities
at the University of Central Florida
Orlando, Florida

Spring Term
2015

Major Professor: Mark L. Kamrath

ABSTRACT

Numerous examples of the "digital scholarly edition" exist online, and the genre is thriving in terms of interdisciplinary interest as well as support granted by funding agencies. Some editions are dedicated to the collection and representation of the life's work of a single author, others to mass digitization and preservation of centuries' worth of texts. Very few of these examples, however, approach the task of in-text interpretation through visualization.

This project describes an approach to digital representation and investigates its potential benefit to scholars of various disciplines. It presents both a digital edition as well as a framework of justification surrounding said edition. In addition to composing this document as an XML file, I have digitized a 1794 English translation of Virgil's *Aeneid* and used a customized digital markup schema based on the guidelines set forth by the Text Encoding Initiative to indicate a set of poetic figures—such as simile and alliteration—within that text for analysis. While neither a translation project nor strictly a poetical analysis, this project and its unique approach to interpretive representation could prove of interest to scholars in several disciplines, including classics, digital scholarship, information management, and literary theory. The practice serves both as a case-in-point as well as an example method to replicate with future texts and projects.

ACKNOWLEDGMENTS

This document and project would not be possible without the guidance and support of my dissertation committee:

- Dr. Mark Kamrath, General Editor, Charles Brockden Brown Electronic Archive and Scholarly Edition; Professor, Texts & Technology Ph.D. program

- Dr. Rudy McDaniel, Assistant Dean of Research and Technology, College of Arts & Humanities; Director, Texts & Technology Ph.D. program

- Dr. J. D. Applen, Associate Professor, Texts & Technology Ph.D. Program

- Syd Bauman, Senior Programmer/Analyst, Women Writers Project, Northeastern University

Thanks to the encouragement of the above individuals, my project has finally reached a presentable stage. In particular, thanks to Dr. Kamrath for taking me on in 2010 as an assistant and giving me TEI coding duties with the Charles Brockden Brown project, as well as agreeing to serve as my chair and mentor. Thanks to Dr. McDaniel for his coding insights and assistance with the Texts & Technology Ph.D. process. Thanks to the rhetorical expertise of Dr. Applen, especially for his and Dr. McDaniel's 2009 book, *The Rhetorical Nature of XML*, that greatly influenced the direction of the project—before I even knew what my project would be. Finally, thanks to Syd Bauman for his technical expertise and for years of teaching and spreading the word of XML, XSLT, and custom TEI schemata. His influence on the field as a whole, including the ongoing development of the TEI, cannot be overstated. I am fortunate to have all of these great minds as foundational and ongoing support for this endeavor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

CHDR   **C**enter for **H**umanities and **D**igital **R**esearch

CMS   **c**ontent **m**anagement **s**ystem

CSS   **C**ascading **S**tyle **S**heets

DTD   **d**ocument **t**ype **d**efinition

ECCO   **E**ighteenth-**C**entury **C**ollections **O**nline

FRBR   **F**unctional **R**equirements for **B**ibliographic **R**ecords

GML   **G**eneralized **M**arkup **L**anguage

HTML   **H**yper**T**ext **M**arkup **L**anguage

ISO   **I**nternational **O**rganization for **S**tandardization

MITH   **M**aryland **I**nstitute for **T**echnology in the **H**umanities

MLA-CSE   **M**odern **L**anguage **A**ssociation—**C**ommittee on **S**cholarly **E**ditions

NINES   **N**etworked **I**nfrastructure for **N**ineteenth-Century **E**lectronic **S**cholarship

OCR   **o**ptical **c**haracter **r**ecognition

ODD   **o**ne **d**ocument **d**oes it all

OHCO   Ordered Hierarchy of Content Objects

RNG   **R**ELAX **NG** syntax

SGML   **S**tandard **G**eneralized **M**arkup **L**anguage

SQL   **S**tructured **Q**uery **L**anguage

TEI   **T**ext **E**ncoding **I**nitiative

TLG   **T**hesaurus **L**inguae **G**raecae

UTF   **U**nicode **T**ransformation **F**ormat

XML   e**X**tensible **M**arkup **L**anguage

XSLT   eXtensible Stylesheet Language for transformations

XTF   eXtensible Text Framework

A NOTE REGARDING SPELLING

Virgil's Latin name, Publius Vergilius Maro, would suggest an Anglicized name of "Vergil," as several of my sources also refer to him. However, the alternate spelling "Virgil" is equally employed, and to maintain consistency with Beresford's translation, I will spell it likewise when possible.

CHAPTER ONE: INTRODUCTION AND OVERVIEW

The current project is the result of the confluence of several interests that—on the surface—seem to possess very little in common. I sing of arms and a man, of verse and meter, of bytes and code, of networks and databases—the markup and transformation of a text with the weighty history of the millennia propelling it into the digital age.

The *Aeneid* of Virgil is a text with an illustrious and distinguished history, serving at one point as defining nationalistic propaganda of the Roman Empire, and in subsequent centuries as a rigorous translational exercise for the enterprising Latin student. In 1794, James Beresford, a Merton College fellow at Oxford University, published his contribution to the epic poem's history: a full translation into English, structured entirely in blank verse. In 2012, I came upon an online auction listing of this book, in decent condition for its age, and, owing to my love of the original as well as old books in general, promptly purchased it.

It was not long before I realized I could continue the text's legacy. For being over 215 years old, the book was certainly in a tolerable condition—it was readable, but its front and back covers were detached, and the pages were beginning to feel brittle. From my associated work on the Charles Brockden Brown Electronic Archive and Scholarly Edition, working with original texts from the same time period, I decided that this book would be a good artifact to preserve, and would serve as a case study in modern digital preservation. Consequently, one of my first goals was the acquisition of high-quality facsimile images so that I could limit my handling of the artifact itself as much as possible. To this end, the University of Central Florida library's Digital Initiatives division was eminently helpful.

My first exposure to the digital humanities occurred in 2005, in an undergraduate corpus linguistics class. For the first time, I realized there was a place where I could marry my love of

language and my enthusiasm for digital technologies. It was here, under the instruction of Dr. David Bowie, that I witnessed and practiced the possibilities of digital markup in a setting decidedly *not* limited to the realm of HTML website design: the Text Encoding Initiative. I attribute the experience of this class as one of the single most-influential to shaping my interests over the next several years. Indeed, because of it, I have been involved for several years in the development of the Charles Brockden Brown Electronic Archive and Scholarly Edition—a project built on a complex system of digital markup files—itself a major influence on my choice of the content-management systems involved herein.

Until my acquisition of Beresford's translation, other interests of mine remained segregated from my concurrent studies and work: Classical Latin and Greek, as well as Early Modern English poetry and its imitation, all occupied various parts of my time, but almost never intersected with my more common undertakings. At this one juncture, finally all of my various interests found a common ground. Combining the classics, poetry, digital text preservation, and a desire to show that digital technology is not restricted to the computer scientists and programmers of the world, the current endeavor was born.

In an attempt to answer the questions of why scholars would and should devote time to the endeavor described herein, Chapters Two and Three provide the theoretical foundation upon which the project is based, offering definitions of key terms and concepts in addition to a synopsis of scholars' contributions to defining text, representation, and digital markup. From these chapters, a theoretical framework is established to demonstrate that whether oral, physical, or digital, texts possess key features that define them as texts, and that markup can describe any of them. Using scholars' broad definition of "ekphrasis" as "the verbal representation of visual representation" (Heffernan 3; Mitchell 152), the third chapter also develops an argument to

demonstrate that ekphrasis is a useful representational concept both in terms of its traditional roots as well as how it applies in the digital age—including how this project and its influences implicitly demonstrate and develop the concept.

In order to explore in any depth the electronic edition of Beresford's work, Chapter Four foregrounds several scholars' definitions and examples of the "electronic scholarly edition." Next, expanding on the definition and examination of markup in the third chapter, this chapter explores the application of markup in several of the projects beyond the Charles Brockden Brown Archive that have influenced my approach with the Beresford translation. While the existence of my digital edition would not be possible without the influence of such established projects in the digital humanities, my own offers a visualization layer of interpretation that is not currently common practice to employ in scholarly editions, but that I believe has its place in such academic products. While user interfaces typically allow the showing and hiding of editorial notes and other supplemental content, the overlaying of an interpretive layer over the facsimile text is yet to become a typical practice. It is my hope that the project has room to expand, or even that other texts can benefit from a similar approach. Further, my edition enables the simultaneous side-by-side comparison of the facsimile and interpreted text, both of which scroll synchronously with each other. This chapter also details the decisions I made and the technical procedure I used to prepare the book for representation in a digital environment. Finally, it puts forth the argument that more scholars outside of the digital fields should recognize and consider a similar approach in their own studies.

Chapter Five examines the textual history of the *Aeneid* as a text continually undergoing transformation (exemplifying Jerome McGann's "deformance" and Jeff Rice's "detournement," concepts explored in more detail in the second chapter), and the situation of the text at the time

when Beresford undertook his translation. It also details several influential and enduring translations of the text into English, from the sixteenth century through the present. From the beginning, I decided that I would showcase the strengths of digital markup by digitizing Beresford's text and applying custom interpretive tags to it. While mine is not the first marked-up edition of either the *Aeneid* or even James Beresford's English translation of it, it is the first to apply a custom set of tags to the text for the purpose of poetic analysis that can be visualized. Since I have an affinity for prosody, it did not take long for me to decide on a set of poetic devices on which to focus for this edition; Chapter Six includes an explication and analysis of those poetic figures, including definitions of each and discussions of their usage in and importance to the text.

Chapter Seven reflects on the experience of the project and investigates where it fits within the conversations of the various fields with which it engages. Additionally, this chapter examines what have I learned based on the successes and failures I encountered, and how the project might contribute to current and future dialog in the Classics, digitization, and representation based on that self-awareness.

CHAPTER TWO: WHAT IS A TEXT?

Representation, Mediation, Interpretation

Egbert Bakker in "Mimesis as Performance" equates the Ancient Greek's conception of the Muses to quintessential, unmediated thought. These nine divinities each represent a humanistic or scientific field, and are the source upon whom so many ancient artists and scholars called for direction and inspiration. If the muses are the original thought, then humans are the mediators, the representers. As Bakker says, "The poet's 'seeing' things is a remembrance of earlier poets'—and ultimately the Muses'—seeing of that same reality. The vision of the Muses was detailed and specific, which in less mystic terms amounts to emphasizing the importance of imagery for the Homeric performance as a cognitive act of re-creation" (Bakker 24). That is, the earliest enduring writings of Classical Greece, from historians, philosophers (Plato and Aristotle), playwrights (including Aeschylus, Sophocles, Euripides, and Aristophanes), and the lyric poets (among whom Sappho, Anacreon, and Pindar number) on the cusp of widespread literacy relied on exquisite detail to mediate their thoughts through performances and inscriptions, assisting in committing them to personal as well as cultural memory.

Language, Ferdinand de Saussure argues in his *Course in General Linguistics*, "is the social product deposited in the brain of each individual" (23). It is the means humans use to express and share ideas, and in its spoken form, offers the shortest path to the origin of an idea—the brain, wherein "mental facts (concepts) are associated with representations of the linguistic sounds (sound-images) that are used for their expression" (10). With the birth of language, then, so too were born representation and mediation. W. J. T. Mitchell, in *Picture Theory*, defines representation quintessentially as "a practical activity of using things to refer to other things" (355). It might be accurate to argue that the only unmediated form of information is that which

lacks representation: thought itself. Bakker makes this argument with the Muses, and Saussure does with mental facts within the brain: as soon as thought is represented by words, by pictures, or by sounds, it is inherently mediated, separated by at least one step from the original thought that bore the work. This object of mediation is what I consider the essential definition of a "text": a representation of thought that is capable of being interpreted. Given this definition, any textual object serves as an intentionally crafted window providing a rhetorically unique perspective of the original thought.

Admittedly I am not the first to offer a definition of the word "text." More traditional definitions exist and are far more established, not the least of which is that of post-structuralist Roland Barthes. Of the word "text," he says, it "is not to be thought of as an object that can be computed" (156), but rather "is held in language, only exists in the movement of a discourse" (157); further, "the Text is experienced only in an activity of production" (157). More distinctly, a text is not necessarily limited to a discrete physical object, or what Barthes distinguishes as a "work": "a fragment of substance" (156) or "the object of a consumption" (161). Instead, "the metaphor of the text is that of a *network*" (161)—a text is produced as a reader reads or a listener hears. Unlike an isolated work, then, a text possesses a legacy; that is, a text is inextricably bound to the interpretations belonging to its audience. While I might argue against Barthes' claim that a text cannot be computed, his definition is a seminal one that bears contextual investigation.

Barthes is not the only one to consider a text as a relationship of interpretations. According to literary historian Elizabeth Clark, "[Jacques] Derrida argues that there is no pure originary text that has not been 'touched' by other texts" (132). Instead, texts are "'substitutive chains' of allusions and references to other texts with which they interact" (132). Derrida's

argument—and Clark's own interpretation of it—shares some common ground with Barthes';
namely, when the text transcends the limitations of a physical medium and becomes something
without any distinct, tangible bounds, it is defined instead by and for the audience that receives
and discusses it.

In their investigation of the development of reading in the Western tradition, Guglielmo
Cavallo and Roger Chartier follow up on the distinction between the text and the work earlier
made by Barthes and echoed by Derrida. No text exists in itself, they claim. Rather, it relies on
its representation to be consumed and understood: "Authors do not write books," they say;
instead, "they write texts that become written objects" (5). Authors' ideas are made manifest in a
product such as a book that can then be interpreted by some audience. Chartier in his own
chapter expands on this idea: without a reader or other form of audience, he argues, a text does
not exist (275). That is, a work can only be regarded as a text if it has been engaged by other
individuals apart from its creator, supporting both the assertions of Barthes—that a text develops
as a networked relationship of interpretations—as well as my above definition of a text as an act
of representation that enables further interpretation. This description of text as an associative
process will be taken up again shortly in its relation to critical interpretation and permanent
transformation.

At this point, a long-standing model of textual representation is worth a discussion. A
concept that might assist in the understanding of what makes a text, especially in relation to the
use of markup for digital representation, comes from DeRose, Durand, Mylonas, and Renear in a
seminal 1990 article, "What Is Text, Really?" What a text *is*—that is, the essential components
that define a text as unique—is determined by an "ordered hierarchy of content objects (OHCO)"
(6). Such "content objects" are individual components that together, in a certain order, define a

text; these objects include quotations, paragraphs, chapters, and other blocks of information that are essentially distinct from how they appear—that is, their content, rather than their form or appearance on a page or screen, is the information that defines a text (5). Clearly, in this case, a text *can* be limited more narrowly to a single object, unlike the text-as-a-network that Barthes describes.

Despite this forthright logical definition of a text, it took less than a decade for several of the same authors who defined the OHCO to question its applicability, especially in regards to the actual experience and practice of the digital encoding of literary texts. Though they fail to outright reject the OHCO they put forth, Renear, Mylonas, and Durand express discontentment with the structural certainty it seemed to imply. They attempt refinement by stating that "the spirit of the OHCO hypotheses is borne out to the extent that texts qua intellectual objects still seem to be composed of structures of meaning-related features and that, moreover, these structures are often hierarchical." That is, they attempt to retain some of what the OHCO model originally presented while attempting to free the "text" from the strict limits of a nesting hierarchy.

Other definitions of what constitutes a text are discussed by Christian Wittern in "The World of XML Markup," in which four separate contextually-specific concepts of "text" are delineated (2). All of these are separate again from the OHCO thesis and my own definition above:

1. From a broad perspective, text is essentially something created to express ideas;
2. From a linguistic perspective, text is a communicative act, fulfilling the principles of textuality;

3. From a literary-theory perspective, text is the object studied; and

4. From an information-theory perspective, text is equivalent to character data.

Clearly, then, "text" possesses no universal definition suitable for all situations. For such a reason I assert that my own definition of text—*a representation of thought that is capable of being interpreted*—as one that encompasses as many conceivable situations as possible in which a text might exist—whether inscribed on a stone tablet, inked on parchment, or stored as electronic data on a computer hard drive. "Reading," then, is one of the most basic means of interpreting a textual representation. Chartier asks us to consider the act of reading "as the act by means of which a text takes on meaning and acquires efficacy" (275). We can continue applying this definition to various texts: a sculpture can be read, a dramatic performance can be read, a book can be read, and a web page can be read. Each of these forms of text constitutes a different method for representing thought, but each also enables some form of reading and interpretation.

According to *The PREMIS Data Dictionary for Preservation Metadata*, available from the United States Library of Congress, a mediation or representation in the digital age comprises "A digital object instantiating or embodying an Intellectual Entity. A representation is the set of stored digital files and structural metadata needed to provide a complete and reasonable rendition of the Intellectual Entity" (32). Here, the word "text" has been eschewed in favor of the less-meaning-imbued terms "object" and "Intellectual Entity," but the meaning itself endures despite the generic terminology: a "reasonable rendition of [an] Intellectual Entity" should indeed encompass the concept of a text. Whether channeling the Muses or rendering "Intellectual Entities," this definition of a text is far broader than the idea of a "work" such as a single

manuscript or physical artifact; indeed, an entire digital database might (and should) be construed as a "text."

Further complicating the distinction between texts and other terms for intellectual creativity is the Functional Requirements for Bibliographic Records (FRBR) model, which distinguishes among several layers of textual access and reference—referred to as Groups 1, 2, and 3, each of which contains several individual "entities." Unfortunately for the sake of discussion and precision, terms do not remain consistently defined between different scholars and experts, and in the case of FRBR entities, there exists some overlap with terms used up to this point. In particular, while scholars like Barthes, Derrida, Cavallo, and Chartier often refer to a "work" in regards to a single instance of representation, FRBR Group 1 entity definitions subvert the relationship thus far established between text and work by defining "work" at the most general level, or as Allen Renear, Christopher Phillippe, Pat Lawton, and David Dubin elaborate, as "a distinct intellectual or artistic creation" as it exists in an abstract state. In the context of this project, the *Aeneid* is an FRBR "work"—such categorization relegates the FRBR "work" entity to a level equivalent to Barthes', Derrida's, and others' notion of "text."

A second FRBR Group 1 entity is the "expression," or "the intellectual or artistic realization of a work in the form of alphanumeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of such forms" (Renear et al., "XML Document"). Here, Beresford's translation is an expression of the *Aeneid* work. To narrow the focus further, the third Group 1 entity is the "manifestation": "the physical embodiment of an expression of a work." For the *Aeneid*, a manifestation of it is the 1794 edition of Beresford's translation.

Finally, "a single exemplar of a manifestation" is an "item," the final and most narrow of the FRBR Group 1 entities. The exact copy of Beresford's translation that I possess is an item, and more closely resembles the "work" referenced by Barthes and others. According to Renear et al., common equivalencies and correspondences can be established between the Group 1 entities and more widespread terminology. In particular, they pair "expression" with "version"; "manifestation" with "edition"; and "item" with "copy." "Work," they maintain, "is already in common use in more or less the sense suggested above," that is, a "distinct intellectual or artistic creation." This assertion does *not* agree with the definition of "work" provided by Barthes and his contemporaries, however. Though the FRBR model is useful especially in the context of library cataloguing and classification, for consistency and familiarity with literary sources I will prefer the terms "text," "edition," and "work" over the FRBR entities "work" or "expression," "manifestation," and "item."

Any work (in the sense that Barthes uses the word, equivalent to the FRBR Group 1 "item") becomes a text—that is, an interpretable representation of thought—when it can be witnessed by an audience. In order to be witnessed, it must be represented for consumption, be it through glyphs on a scroll or HTML on a web page. As W. J. T. Mitchell concisely states in *Picture Theory*, representation is "a practical activity of using things to refer to other things" (355). This act of representing a text not in its original form is "remediation" (note: the term "remediation" carries other connotations that do not relate to the present endeavor—all instances of the word in this text should be taken in the context of media studies and *not* in the context of educational or environment remedies). The concepts of (re)mediation and representation are older even than recorded language itself. There is no reason that the depictions of hunter-gatherers scratched onto a cave wall should not constitute texts: such marks represent a version

of events that occurred at some point in history and indicate a conscious, creative endeavor. According to Johanna Drucker and Emily McVarish in *Graphic Design History*, the signs of human language and graphic systems are "[c]apable of representing things, ideas, actions, and abstractions" (5). Cave paintings and other early "deliberate mark-making . . . not only responded to needs (or fulfilled a superstitious function) but also corresponded to ideas" that eventually led to stable, pictographic signs and systems developed through shared cultural meaning (6–7).

In the simplest sense, every representation of a text is inherently an act of interpretation: each is a new creation connected to the original thought, but is not an exact replica of the original. This uninterrupted connection of interpretations and representations is certainly not a new construct, though; for Jerome McGann, the connection is a series of transformative re-interpretations stretching between a text and all its forms back to the original thought that he calls "deformance": a combination of "deformation" and "performance." This relationship is the manifestation of interference by "later critical responses" (McGann 101) and clearly shares the critical underpinnings of Derrida's substitutive chains of allusions and references. No single text in the relationship chain is identical to the one before it, so every interpretation differs from and therefore transforms the "original." Each transformation is a realization of choices made by an editor, creating a lens through which the audience must view the text. In this regard about scholarly editions and how they are presented to an audience, Paul Eggert argues, "All editions of literary works interpose between the reader and earlier documents that present the text of the work" (97)—again, here, the "text" and the "work" are distinct.

Indeed, it is a relatively common opinion that "It is impossible for a transcription to reproduce the original object; it is always a selection of features from that object: the words but

not their size on the page or the depth of the chisel marks, major changes in type style but not variations in the ink's darkness from page to page or over time" (Lavagnino 338). Such features can of course be represented, but finer granularity equates to increased planning, increased time, and increased possibility of errors—both in human effort and data storage, in the case of electronic media. For such reasons, Lavagnino cautions against overcustomization of represented features, as Patrick Durusau (below) similarly cautions that poorly planned or executed markup can harm an edition: too much representation can distract an audience by drawing attention to unintended parts of the text—or even to the presentation medium or interface itself. Moreover, if the edition is a digital one, more data can harm its interoperability as well as its electronic longevity.

For Jeff Rice, a similar idea is manifested as "detournement," whereby remixing, re-appropriation, and the breaking of habitual patterns becomes "cool"—not unrelated to copyright concerns brought to the fore in recent years, examined academically by Lawrence Lessig. All media, whether new or traditional, "are always and at all times building upon the creativity that went before and that surrounds them now" (Lessig 29)—to see texts restricted behind the gates of copyright is a disheartening prospect, but is a very real and pressing concern for literary scholars who wish to remediate and represent texts that have come before. Speaking to the appropriation and even "theft" of culture, Rice contends that "all writing involves some degree of theft, particularly when writing is introduced into the digital, an area that relies to a great extent on the 'borrowing' logic associated with appropriation" (57). For Rice, then, any text—but especially a digital one—relies on ideas and efforts that have come before. Such an assertion is not out of step with the ideas of earlier scholars, including Barthes and the network; Derrida and substitutive chains; and McGann and deformance.

In his work on the nature of remediation, Jay David Bolter touches on many of the same core concepts as Jerome McGann in *Radiant Textuality*. For Bolter, the most simplistic definition of remediation is a "shift" in which "a newer medium takes the place of an older one, borrowing and reorganizing the characteristics of . . . the older medium and reforming its cultural space" (23). The shift from oral delivery to writing on a papyrus scroll is one such remediation, the shift to print another, and the shift to electronic literature yet another (Bolter 23–25). However, Bolter continues, "Remediation is not limited to technologies of writing"; as one example relevant to the digital humanities, "[N]ew visual media, such as computer graphics, virtual reality, and the World Wide Web, define themselves by borrowing from, paying homage to, critiquing, and refashioning their predecessors, principally television, film, photography, and painting" (24). There exists a chain of representation, then, linking new media to previous forms of media. In the examples Bolter mentions, computer graphics remediate visual arts such as sculpture, painting, and photography; virtual reality remediates film as well as perspective painting; digital photography remediates the analog photograph; and the World Wide Web absorbs and refashions almost every previous visual and textual medium, including television, film, radio, and print (25).

Bolter's ideas of remediation undoubtedly bear resemblance to other scholars', but Bolter provides an additional claim that separates himself; namely, he says, "[O]lder media can remediate newer ones within the same media economy" (25). The implementation of computer-generated graphics in film is a commonplace example in the modern media economy that includes the Internet, and Bolter gives the example that "TV screens often look like pages from the World Wide Web" (25). This assertion is certainly visible on many news outlets whose

television screen layouts very much resemble framed websites, with each frame dedicated to a section of news such as a headline ticker or a local weather report.

Like McGann's claim that deformance is a constant chain linking previous interpretations into one set of texts that reference and build upon one another, Bolter similarly argues, "Each medium seems to follow [a] pattern of borrowing and refashioning other media, and rivalry as well as homage seems always to be at work" (25). In both cases, "refashioning" (Bolter) or "reinterpretation" (McGann) are at play. That is, each theory relies on the existence of previous works—hence, of course, the *re* prefix. Without a previous model, representation is simply *presentation* and is another entity altogether. While Bolter's "remediation" and McGann's "deformance," however, share the act of transformation, there is a key difference. The former involves a change in the medium of representation; the latter, on the other hand, occurs when a critical analysis has been applied to a text. Either one results in a changed text, but deformance implies an act of interpretation that remediation on its own lacks. Of course, when a text undergoes remediation, its new representation becomes part of the chain of a text's deformance.

Speech, one of the original methods of representing thoughts, serves as a means of linking symbols with observed objects and phenomena. Speech is, essentially, the original medium. With speech, humans could communicate and share information, not only between two people, but from one person to many, as the rhetors of ancient civilizations seemed so adept at exploiting. It is perhaps ironic, then, that the primary source of our knowledge of these ancient speakers comes not from speech, but from another medium altogether: writing.

Writing functions similarly to speech, with symbols standing for sounds that in turn stand for thoughts, but on a different level of abstraction; that is, as Katherine Hayles states, "The content of media is other media" (Hales, *Electronic Literature* 90). Chirography, or the practice

of writing by hand, as Walter Ong defines in his discussion of oral and literate societies, is a means of preserving or memorializing thoughts. Whereas speech was ephemeral and only remembered by repetition and recitation, thoughts could be preserved in writing, even by a single writer. Nor was writing limited to the instant it was produced: as long as the medium exists in some form, whatever was written at one point in time can be viewed in another. Certainly the advent of writing became a reason for social change and upheaval; Plato famously lambasted the technology for its lack of ability to reason and argue, and writing also created a disparity through its requisite education. With education reserved throughout most of history for the most privileged, a knowledge of reading and writing became a sign of status. It was not until the relatively recent development of print writing (as opposed to chirographic) that the status quo shifted again. With mass production, writing and reading became available to a wider audience, and so too did the ability to understand, reinterpret, and create new knowledge by a larger community.

Though print literature widened the audience of knowledge production, it was not until radio and television that "new media" was born. No longer was information transferal limited to the arbitrary signs and symbols of writing. Now, humans had sounds and images with which they were familiar in speech and everyday life being presented to them. Even those without the education to read written or printed words could understand the images and sounds emitted from televisions and radios—given, of course, they could understand the language they were witnessing. Starting with these technologies, and continuing into the early age of the Internet through the present, Ong's age of "secondary orality" blossomed.

Now, in the digital age of the globally connected world, we can look at an electronic display and receive information—writing, speech, and images included—from anywhere in the

world nearly instantaneously. What would have taken weeks or months to travel just 200 years ago can now be transferred through fiber and across continents at the speed of light. So it is that any *new* medium that is created to communicate information is layered upon previous media: to take an example of the present day, a news web site contains print, which contains writing, which contains speech, which in turn contains the thoughts that are being represented. Each layer of this intricate chain of representation creates a new interpretation.

In fact, it is the very act of interpreting that lies at the heart of media studies and, indeed, much of humanity's endeavors. Ron Burnett, in *How Images Think*, establishes an important distinction between "data" and "information," wherein he argues that the former—data—is content without any sort of qualifications put upon it (201). However, once any sort of interpretation has been made upon data—once it has been mediated, for example—does it become "information." As Daniel Headrick similarly posits, this "data" is information that is akin to the metaphor of a tree falling in the forest but making no sound: only once such data has been meaningfully understood by humans does it become "human information" (3–4) that can be used, assimilated into knowledge, and passed on—that is, *interpreted*. This definition explicitly limits the content with which computers can be considered to function: computers process and produce "data," but never "information"; a human understanding is necessary for the conversion of data into information.

The Digital Humanities: A New Space for Old Practices

The intersection of traditional academic inquiry and digital technology is undoubtedly a new space, so new, in fact, that it still lacks a standard name. However, be it "new media studies," "emerging media," "digital humanities," or "texts and technology," its purpose is unified: how to harness the ever-increasing potential of digital technologies and apply it to the

academic pursuits of humanistic interpretation and investigation. Works of this century, including Jerome McGann's 2001 *Radiant Textuality*, are certainly not the first to explore how the computer and the Internet can be used to aid in traditional academic exploration, but they continue a long-running discourse in the humanities that ponders the very same topic.

It is likely that the late Father Roberto Busa became one of the first (if not *the* first) digital humanists when he initiated his *Index Thomisticus* in the 1940s with the assistance of IBM's founder. Not only was he an early digital humanist, however, he was arguably the first "digital classicist," considering the fact that his text was in Latin. It was Busa who desired to compile a database of Thomas Aquinas's work in a searchable database with the help of electronic technology, and he succeeded, with many iterations and changes in storage media. From thousands of punchcards, to magnetic tape, to CD-ROM, and most recently to an online archive, Busa's project in its current form is the product of more than six decades of work—and it has outlived its creator, as he predicted.

When Busa wrote the introduction to *A Companion to Digital Humanities* just a decade ago, he nonchalantly accepted the fact that future iterations of his project would be posthumous. He also acknowledged the changes in technology he had seen, and opined that there would be unforeseen ones. Perhaps only partially in jest, he remembered the day that "[i]n his mercy, God led men to invent magnetic tapes" as a storage medium, thus allowing his project to transition from what would have been 500 tons of punchcard weight to several hundreds of tapes, and eventually in the 1980s to a single compact disc (xvii). Busa, though, was not the only one to recognize the ever-changing nature of digital technology and what it could mean for humanistic inquiry.

In an essay that eventually became part of the book *Radiant Textuality*, Jerome McGann states, "We stand at the beginning of a great scholarly revolution" ("Radiant" 381). McGann refers to the potential offered to scholars by the Internet and the increasing availability and accessibility of digital technology. Since 1996—when McGann wrote the preceding—the Internet has become a hugely different entity from what it was then, with so many more connected to it at much higher transfer speeds. This fact, however, does not necessarily render McGann's statement obsolete; rather, it might perhaps be more relevant now than before.

Richard Finneran, echoed by McGann, believes that it is no longer relevant to focus on individual editions of any given work. Instead, everything produced on a topic influences the others, and nothing can be viewed in isolation—thanks to digital technology, our conceptions of "authorship, reading, the very nature of the text" have been "profoundly altered" (245). As McGann argues, every edition produced is both the product of and participates in deformance, defined as the point at which a text "is forced to take on meanings of which it was not originally possessed" (110). Deformance, he continues, is a result of interpretation. As soon as something is written about a book, for example, that book is forever changed. Similarly, translations of books into other languages are acts of deformance. A text takes on new meanings when there occurs a change to it; for McGann, interpretation of a text changes it, and the act of interpretation is also a performance. Any text that has been interpreted has also been changed, or deformed. In short, in our connected world, no textual object can exist in isolation.

Anticipating McGann's argument by several hundred years, James Beresford writes in the preface to his translation that he fears "that, in like manner as the most perfect Cast from a statue will, by certain abrupt joinings and raggedness of surface not common to it with its Original, betray itself as a copy, and will, moreover, discover to a scientific eye a difference and

19

inferiority furtively pervading the whole" (ix). Similar to McGann, Beresford believes that his translation, no matter how faithful to its source, will nevertheless be considered a derivative work that betrays itself by certain imperfections added by the translator.

In addition to McGann and Finneran, other digital humanist scholars and experts such as C. Michael Sperberg-McQueen, Gregory Ulmer, Allen Renear, and Dino Buzzetti have offered their insights on such varied topics as digital markup specifically for use in the humanities; the difficulties in creating facsimile images for Internet publication; the metaphorical minefield of obtaining permission to use original images—an important and worrisome topic that Lawrence Lessig discusses at length in *Free Culture*; and the importance of documenting every step of the way in creating an electronic scholarly edition, from conception to production to publication. Besides practical and theoretical considerations and difficulties, these scholars have explored the history of the computer as a technology on its own and in conjunction with textual analysis, as well as specific examples of projects that have come to fruition over years (and sometimes decades) of work. Such discussions and examples lay an important foundation for all of new media studies.

New media inarguably equates to new means by which to present content. When writing became established, new methods were developed to allow information storage and retrieval, including clay tablets, palimpsests, scrolls, and books. The technologies of printing, including the printing press, allowed more standardized and more easily distributed books. The telegraph, radio, and television transmitted information without the need for any more physical space than the single device required. Katherine Hayles, Janet Murray, Jeff Rice, and Jay David Bolter all similarly (and validly) argue that without a medium, content cannot be presented. Murray believes that all media, whether print, television, or Internet, contain examples of great and

atrocious content: it is the *meaning* that makes the difference (273–274). On the other hand,

Bolter, Rice, and Hayles, especially, seem to think that the content is reliant upon the form, or

medium; as Rice argues in *The Rhetoric of Cool*, "the *content* of media, like video games and

television, does not affect us so much as the *form* and *rhetoric* of media do" (7), invoking a link

to both traditional and digital rhetoric. Though a clear consensus may not be achieved among

these scholars, it seems clear that new media—specifically the Internet, here—offers scholars a

new means of representing and disseminating information.

Digital representations of texts offer us new means of looking at old information. To

consider a text as computational data is to allow machines the ability to process the text. As

Buzzetti argues, "a digital representation is data, and data is processable. Data is the

representation of information in a form that can be processed by a machine. And this is a point

worth insisting upon; for the means of rendering a text – spoken, written, printed, digital –

affords a different and distinctive approach to seizing it" (46). However, Buzzetti believes that

modern digital editions do not take full advantage of the digital mode. Scholars are either too set

in the literate mode of textual analysis to change, or are uninformed about what capabilities the

technology can offer. For Buzzetti, "the true rationale of a genuine digital edition consists

precisely in taking advantage of the digital form of representation to improve our critical

engagement with the text through effective computational processing" (46). While many current

scholarly archives and databases simply present images and transcriptions of handwritten or

printed texts, more is possible, and it is my hope that the current endeavor embodies that

possibility.

Countless prototypes and examples of scholarly new-media genres exist online; these

include, in no particular order, NINES, the Ivanhoe "ludic playspace," and the Rossetti Archive,

all of which involve McGann; the Mark Twain Project; the Walt Whitman Archive; the Willa Cather Archive; the Charles Brockden Brown Electronic Archive; the Women Writer's Project, housed formerly at Brown University and currently at Northeastern University; and "Paul the Not-So-Simple" by David Birnbaum of the University of Pittsburgh. Many of these projects are ongoing and have received several rounds of funding from the National Endowment for the Humanities, while others are becoming increasingly difficult to access with modern technology—an all-important consideration editors should acknowledge in their planning process.

NINES, an acronym for **N**etworked **I**nfrastructure for **N**ineteenth-Century **E**lectronic **S**cholarship, serves as a research and resource hub for digital scholarship on texts originating in the nineteenth century. Rather than being itself an edition dedicated to a certain text or author, it more broadly serves as a communication nexus, peer-reviewing body, and support system for digital scholars working with material from the nineteenth century. It is a self-described "scholarly organization devoted to forging links between the material archive of the nineteenth century and the digital research environment of the twenty-first." NINES is also the home to Ivanhoe (frequently stylized in all capitals as IVANHOE), a "game" developed under the auspices of Jerome McGann for the purposes of fostering "critical awareness of the methods and perspectives through which we understand and study humanities documents" ("Ivanhoe"). While repeatedly referred to as a "game," IVANHOE is more accurately a shared virtual space in which participants can interact with one another and with a digital representation of texts. It shares some features of many commercial online games, including the ability of multiple users to connect simultaneously to a central server, where they can interact with the environment and see the results of each other's actions. Unlike most such games, however, there is no predefined

"goal" or set of rules apart from what the users themselves might determine. Unfortunately, the game has been inaccessible since 2012, with no certainty regarding future access, whether in a live or archival state.

In addition to NINES and IVANHOE, Jerome McGann has a guiding hand in the Rossetti Archive, an archetype of the modern developing genre of the "digital edition." The Rossetti Archive "provides students and scholars with access to all of [Dante Gabriel Rossetti]'s pictorial and textual works and to a large contextual corpus of materials, most drawn from the period when [his] work first appeared and established its reputation" ("Rossetti"). Evidence of the time-consuming nature of the development of such a collection exists in the fact that, while started in 1993, the Rossetti Archive was not in a complete state until 15 years later in 2008.

Similarly, the Mark Twain Project Online is an online archive showcasing collected works written by Mark Twain, the ultimate purpose of which "is to produce a digital critical edition, fully annotated, of everything Mark Twain wrote." The particular importance of this project to my own is not manifest in the content as much as the means of presenting it. Specifically, its customizable document interface—developed at the same institution, the University of California-Berkeley—"provides a powerful research and reading experience" ("Mark Twain Project"), and is one I chose to present my own text (described in more detail in Chapters Three and Four below).

The Walt Whitman Archive, like Rossetti and Mark Twain above, serves as a further example of what might reasonably be termed a "classic" digital scholarly edition. It is the work of digital scholars Ed Folsom and Kenneth Price, who conceived of the edition in the 1990s, and believed the lack of physical constraints offered by a web page were well suited to an edition of Whitman's work since "fixed forms of print do not adequately capture his incessant revisions"

("Whitman"). This belief is certainly not limited to the work of Whitman, and could be a cogent argument for many of the scholarly editions currently in existence.

Continuing the investigation of single-author digital collections is the Willa Cather Archive, also located at the University of Nebraska-Lincoln, and founded in 1997. The Cather Archive, however, unlike the other projects discussed above, faces the additional challenge of copyright concerns about the work published by Cather after 1922 (Jewell). Copyright and digital publication constitutes a concern for all scholars today, and is an issue that requires delicate and careful consideration, especially in the case of a digital edition that seeks to publish as many works of an author as possible.

Finally, the Charles Brockden Brown Electronic Archive is, once more, an edition dedicated to the works of a single literary figure. In this case, the early American novelist Charles Brockden Brown is the focus, and it is this project that has provided the largest influence in my current endeavor due to my personal involvement with its growth and development over several years. Housed at the University of Central Florida, it is a nexus for scholarship on the author, and coordinates the efforts of academics interested in Brown from all around the world.

The Women Writer's Project is a group responsible for the publication of *Women Writers Online*, a collection of digitally encoded texts written by pre-Victorian women. Begun at Brown University in the late 1980s, it has been under steady development and serves as a foundational model of producing texts for publication in a digital environment. Of note in this regard is its close association with the development and adoption of digital text-encoding technologies and presentation platforms. Unlike most other projects described here, however, texts from the Women Writers Online collection are available as a subscription to institutions and individuals ("WWP"). Nevertheless, the approach undertaken by the Women Writers Project serves as an

imitable one for the production of other digital scholarly editions, and this project is at the forefront of modern digital text collections.

Unlike McGann's Rossetti Archive and other similar collections of works by a single author, "Paul the Not-So-Simple" serves as an example of a digital version of a single tripartite text, a collation of three translations of the same work, along with explanatory notes, all cross-referenced within the same web page and generated from a source XML file. This text, an exercise in comparative translation at the University of Pittsburgh, is billed as "a technologically innovative electronic edition of the Old Church Slavonic Life of 'St. Paul the Simple' from the *Codex Suprasliensis*." Its technological innovation and particular influence on my edition are discussed further in Chapter Four.

In the field of classical literature, there are several digital behemoths, but distinctly few projects dedicated to any one text or author. Besides Roberto Busa's *Index*, the largest collections in the digital classics include the Perseus Project, the Open Greek and Latin Project, and the Thesaurus Linguae Graecae, all of which contain huge collections of ancient texts that have been digitized, and are examined in more detail below in Chapter Four.

No single one of these examples is built on any standardized model or pre-existing template; rather, each one has a unique history with a unique planning process, and the current state of each one reflects that fact. No interface is the same, and no type of information represented is identical. Each is an example of what a digital edition can be. Unfortunately, despite the positives of worldwide access and electronic design, the current state of McGann's IVANHOE serves as an exemplar of the difficulty in working in the ever-shifting cyber-environment that is the Internet: after 2007—quite recent in terms of a human lifespan, certainly—it has become impossible to play the game on "modern browsers," and the last update

to the site from January 2012 suggests that no current work is being done to make the game accessible. This fact raises a troubling concern for the "great scholarly revolution" (McGann, "Radiant"), of course: if information from as recent a year as 2007 is already inaccessible, what kind of future does that suggest for efforts in the digital humanities? Granted, Ivanhoe is not the same kind of information as the other more regularly maintained scholarly databases such as the Rossetti Archive or the Mark Twain Project, but it remains visible as a cautionary reminder, at least. Nor again has it been consigned to oblivion; such is one of the benefits of the hypertext that McGann himself mentions: "Unlike a traditional book or set of books, the hypertext need never be 'complete'" (*Radiant* 71).

It is also important to keep standards in mind: it cannot be a solution to rely on proprietary methods of data storage, retrieval, or transfer. Such a necessity is echoed again and again by luminaries like Burnard, Sperberg-McQueen, and McGann, and why the eXtensible Markup Language (XML) is so often chosen by digital humanists to build digital editions or databases. XML, as a plain-text format, provides a better guarantee of future access, and is hierarchically structured to allow relatively easy transformation to any newer, more-robust standard that may be developed. Father Busa would likely have appreciated this foresight. It is also XML that serves as the bedrock for nearly all the aforementioned scholarly archives. Even though they do not all adhere to the same schemata, or "flavors," of XML, each one ensures future viability by following a consistent and pre-planned structure, a necessity to determine and document when planning scholarly databases.

Of course, academics should not shift their entire focus to preservation. As Daniel Cohen and Roy Rosenzweig argue, "In order to preserve, you must first create." Scholarly inquiry is undoubtedly experiencing change, arguably more than it has in hundreds of years; besides

expanding access to more people and areas than ever before, the Internet is opening up new ways of both creating and preserving content. Roberto Busa's early version of the *Index* was as much a scholarly database as the Mark Twain Project; it simply existed in a less variable form. Variability (as Lev Manovich argues is one of the defining traits of new media) and deformance, though, define this "great scholarly revolution," and we must learn to balance variability, access, and preservation enough so that what we create can be archived, and what is archived remains accessible for the future.

CHAPTER THREE: WHAT IS MARKUP?

In any text, there exist two separate, fundamental dynamics, without either of which a text cannot exist: the meaning and the representation of the meaning. While the meaning can exist simply as thought, it is not until that thought is represented through some means that it becomes a text both consumable and interpretable. At the point at which a text is capable of perception by an audience, it can be infused with external meaning and *marked up*; an author can mark up their own text in as simple a manner as giving it punctuation, but an audience can mark up the same text by appending to it a new meaning the author may not have foreseen. Such "markup" is not restricted to physical marks on a piece of paper, though such marks are responsible for the term itself. Indeed, with changes in modes of representation have also come new means of markup, and the newest forms underlie the most fundamental drive of the current project. While the digital age has brought with it significant new markup technologies, the eXtensible Markup Language (XML) is the one used by this project to describe and subsequently represent content in a digital environment. Finally, while all forms of markup provide methods for making statements about a text, such interpretations bring a longstanding tradition into an electronic age—that of "ekphrasis," or the description in one medium of a text in another.

"Markup" itself is a word with a broad enough array of definitions to demand refinement: in regards to the current discussion, and in line with several scholars, I propose markup to be a means of imposing interpretation upon a text. Different technologies that mediate language— from speech, to writing, to print, to digital—each offer a different approach to textual interpretation, and each kind of text can be marked up. Oral markup, for example, has resulted in the change of stories from one mouth to another. As the *Guidelines* of the Text Encoding

Initiative (TEI) eloquently state, "markup" is "any means of making explicit an interpretation of a text" (xxxiii).

In the oral mode of communication, markup can be seen in the re-telling of stories, a process that results in subtle variations each time a text is spoken—each variation, in effect, creates a new interpretation upon the original text. In the literate mode, be it chirographic or print, markup can be seen in the editing and notating of a text. More recently, digital tools have enabled a wider degree of markup, perhaps most notably with descriptive technologies including hypertext and digital metadata. In addition to Walter Ong's language apparatus of orality and literacy, Gregory Ulmer posits a third—electracy. Electracy expands humanity's means of expression in the same way that orality witnessed the technology of spoken language, and literacy saw the invention of writing, scholarly institutions, and new interpretive methods allowed by such. For its part, electracy is defined by the electronic encoding and transmission of data. Importantly, each apparatus—orality, literacy, and electracy—brings with it new technologies, new means of viewing and representing the universe, and new institutions to perpetuate and expand our knowledge (Ulmer xxiii-xxiv).

It seems apparent that the word "markup" can signify any one of many different concepts, but where does it originate? According to Tim Berners-Lee, acknowledged inventor of the World Wide Web, "HTML is the language of the Web, and XML has become a favoured way of structuring information. But not many people realise that these languages owe their origins to typesetting" ("Mark-up"). Berners-Lee traces this typesetting influence back to 1967, when "printer William Tunnicliffe suggested at a conference of the Canadian Government Printing Office that publishers should encode texts with generic markup," whereby parts of a text would be identified by their function rather than for their appearance. Eventually, this idea led to

the development of the markup standard GenCode, itself a foundation for later developments including the Generalized Markup Language (GML) of Charles Goldfarb, Edward Mosher, and Raymond Lorie, which Goldfarb himself bills as "a means of allowing the text editing, formatting, and information retrieval subsystems to share documents" in an IBM research project commissioned and adopted for managing legal documents. Crucially, "GML introduced the concept of a formally-defined document type with an explicit nested element structure" (Goldfarb 568). GML served as a basis for the subsequent Standard Generalized Markup Language (SGML)—developed by an ISO working group chaired by Goldfarb—which Berners-Lee thereafter used as a starting point for the widely recognized markup language of the Web, HTML.

Even limited to textual markup, several adjectives can precede the word to refer to the different effects that markup applies to a given text, among them "procedural," "descriptive," and "referential"—though some of these types have the potential to overlap. Technical editors Carolyn Rude and Angela Eaton distinguish among several types of textual markup based on the mode of the text and its status in a production cycle. "Editorial markup," for example, "consists of directions for the development and production of [a] document" (Rude and Eaton 43), including suggestions for content or visual design revision of a printed document, while "structural markup . . . identifies the parts (structure) of the text" in order to separate "document structure from the appearance of the document" in an electronic file (60). Of such procedural markup, Goldfarb argues that it is "inflexible . . . [in that,] if the user decides to change the style of his document, . . . he will need to repeat the markup process to reflect the changes" (7). Such inflexibility is a problem inherent to editorial markup. Lou Burnard, in his exploration of SGML's place in the humanities, agrees, defining the origin of markup very narrowly as

specifically "editorial" markup, or "annotation or other marks within a text intended to instruct a compositor or typist how a particular passage should be printed or laid out" (42). This seems inadequate, however, since markup can instruct anyone—not just compositors or typists—about any aspect of a text.

> James Coombs, Allen Renear, and Steven DeRose in a 1987 article on the subject argue: Whenever an author writes anything, he or she 'marks it up.' For example, spaces between words indicate word boundaries, commas indicate phrase boundaries, and periods indicate sentence boundaries. This fact is widely ignored; indeed, markup is usually treated as an unfortunate requirement of using electronic text-processing systems, that is, as something to be avoided. A careful analysis, however, reveals that authors regularly use two types of markup in their manuscripts: punctuational, for example, placing periods at ends of sentences; and presentational, for example, numbering pages. Thus, markup cannot be escaped because our writing systems require it. (934)

That is, *nearly everything about a written text outside of its meaning-content is markup*, and markup is used ubiquitously. Further, and more broadly, Coombs et al. define five "types of markup" that might apply variously to texts in different modes, be they spoken, printed, or digital:

1. punctuational: "the use of a closed set of marks to provide primarily syntactic information about written utterances"—though ubiquitous, markup in the form of punctuation is vulnerable to consistency and stylistic variations;

2. presentational: "includes horizontal and vertical spacing, folios, page breaks, enumeration of lists and notes, and a host of ad hoc symbols and devices"—like punctuation,

presentational markup exists primarily for the purpose of human reading and

comprehension;

3. procedural: "consists of commands indicating how text should be formatted"—

procedural markup includes editorial marks that indicate changes in formatting that one

might wish to be made to a text;

4. descriptive: "indicates what a particular text formatter should do; descriptive markup

indicates what a text element is or, in different terms, declares that a portion of a text

stream is a member of a particular class"—in the realm of digital markup, start-tags and

end-tags that define the boundaries of elements serve to describe the content within them

as possessing certain characteristics; and

5. referential: "refers to entities external to the document and is replaced by those entities

during processing"—in digital markup, character entity references like "—" fit this

category, whereby the definition of what is being referenced—in this case the em-dash—

may be changed or tweaked for formatting purposes, while the content of the referring

document remains unchanged without a loss in meaning or legibility. (935–937)

Though these types of markup laid out by Coombs, Renear, and DeRose can be useful for

categorization, their definitions have not gone unchallenged in the intervening years, nor has the

terminology remained static. As Syd Bauman mentions in "Interchange vs. Interoperability,"

"what we used to call 'descriptive markup' is now better described as 'indicative logical

markup,'" as suggested by Allen Renear in a 2000 essay on the complications of markup

terminology. In particular, Renear attempts to problematize the terminology by associating the

uses of markup with Speech Act Theory in linguistics: markup can be "imperative," "indicative,"

or "performative," he says (417). While there is some merit to refining terminology, not all of Renear's distinctions seem useful. The <p> of HTML and the <p> of XML-TEI each describes a paragraph, which certainly throws into question the usefulness of the term "descriptive markup"—one of Renear's main points in this article, considering the differences between the intentions of HTML and XML. How, then, should we distinguish between an HTML paragraph that is meant to be displayed a certain way, and an XML paragraph that simply *is a paragraph*, where nothing specific is meant to be said about how it appears? This problem is a useful one to consider, and one which seems to have a ready solution.

The goal of using "indicative logical markup" is, according to Bauman, one of the primary reasons that we encode texts, that is, "to help us analyze or explore the text in a manner congruent with our research or teaching." Marking up a text for instances of poetic tropes is more correctly *indicative* than it is solely *descriptive*: each XML element dedicated to poetic figures of speech indicates where it occurs and information about it, but it also indicates that there is something meaningful about that text and that it has been marked up for a reason. Again, turning to Bauman, we can determine that the goal of analyzing a text can be met by "develop[ing] a markup system that separates out the parts of the document you find interesting and describes that which is interesting about them in sufficient detail."

According to the above list, several processes and technologies can be categorized as types of markup: marginalia, commentary, the editorial process, LaTeX, XML, HTML—each can pertain to at least one of the five types and allows a broad interpretation of what constitutes markup. For humans, a paragraph might be recognized through convention and practice by the markup of white space and punctuation. Computers, however, must be explicitly directed to recognize such textual components.

The all-important part of describing an electronic text by means of markup is known as "encoding." In a 1999 article for *Computers and the Humanities*, Christopher Welty and Nancy Ide define "text-encoding" generally as "the practice of marking up text with tags that indicate a section of text should be interpreted or rendered in a particular way" (61). Encoded text, then, is text that possesses self-contained information about the text itself in order to facilitate processing and transmission by computer programs—important features for a digital text's accessibility and preservation. By being encoded, a text crosses a crucial threshold of identity: it transcends the world of characters-and-spaces and joins the world of electronic data, wherein it gains content specifically *about* those characters and spaces.

Regarding the creation and storage of data, Christine Borgman asserts that the word "data" itself refers generally to large, individual textual units. More specifically, she says, "Data are outputs of research, inputs to scholarly publications, and inputs to subsequent research and learning. Thus they are the foundation of scholarship" (115). Data here are independent and ostensibly whole blocks of information in the way that a book, an article, or a manuscript is a single instance of data. The quotation in question comes from Borgman's discussion of the so-called "data deluge": while it is indeed becoming increasingly feasible to store, access, and analyze large quantities of these data, no systematic or standardized method or infrastructure yet exists to store or disseminate them, meaning that textual scholarship and its products, while increasingly pursuable, are not universally reachable (115).

In their discussion of where markup in general—and XML in particular—belongs in FRBR categorization, Allen Renear, Christopher Phillippe, Pat Lawton, and David Dubin argue that "an XML document is generally understood to be a combination of text (or other data content) and XML markup." Operating on this assertion, "data" takes on a different definition

from Borgman's: while Borgman believes that "data" means an entire, discrete textual artifact, to Renear et al., "data" encompasses the meaning-bearing content contained *within* a textual artifact, specifically the content that comes in the form of words, images, or other pieces of content that can be interpreted by a reader or viewer (in agreement with the OHCO model). This view of data is arguably one of a finer granularity than Borgman's, though both assessments of data possess merits depending on the context in which they are discussed or analyzed. Particularly, problems arise at least in part due to the lack of accepted definition of what actually makes up a text; using Christian Wittern's contextually-specific definitions discussed above, it seems clear that a similar disconnect occurs with the word "data"; even so, "data" and "text" in this context seem almost interchangeable: Borgman's "text" is arguably defined as "the object studied," while Renear's is "character data." With this distinction, it will be easier to understand the many varying uses of the same words.

If we consider the content of a text as "data," then information we impose upon that content essentially becomes data *about* that data, commonly referred to as "metadata." Unlike HTML, whose tags "might indicate a preferred typeface, point size, font color, and even something of the arrangement on the page of a poetic stanza," XML metadata indicates something specifically structural or interpretive about a piece of the text (Freistat and Jones 111). For example, to mark a string of text as italicized text in HTML, the start-tag <i> and its matching end-tag </i> (both of which combined along with everything contained between them comprise an "element") are used. By XML's extensible nature, though, scholars can say much more about a text than its structural elements; a reason can be given in the metadata for *why* a string of text is italicized, such as *emphasis* indicated with an <emph> element. In this case,

while HTML is concerned primarily with the straightforward fact that the content should be displayed in italics, XML indicates the *reason* the text is different from the text around it.

In fact, virtually any feature of a text can be marked interpretively since this standard of digital markup is customizable. As Underberg and McDaniel state, "XML emerged as part of the overall development of humanities computing, concerned as it was for expanding the potential usefulness of a digital collection beyond that of the original researcher's intentions" (48). The digital marking up of texts is important because, while any given text may benefit from digitization through searchability and preservation, such digitization allows for future work using the same source material, essentially laying a groundwork upon which future researchers can build their own interpretations.

According to the Dublin Core Metadata Initiative, "Metadata has been with us since the first librarian made a list of the items on a shelf of handwritten scrolls" ("Using Dublin Core"). Metadata, whether it be a list on paper or an SQL database on a web server, is essential to interpretive performance. It must be possible to categorize and filter the vast amounts of data that humanity creates in order to generate meaningful information from the chaos.

To differentiate between data and metadata, Cynthia Haynes offers the distinction that "data" is "primary and visible," while "metadata" is "ancillary and invisible." However, according to Haynes, that distinction has increasingly blurred as manipulation of HTML <meta> tags for the sake of search engine optimization increased, becoming the networked data equivalent of *deus ex machina* (Haynes). While only tangentially relevant to structured markup data like XML, such manipulation exposes the rhetorical presence of an author and their motives even within the usually hidden metadata.

According to Daniel Headrick's comparison of data and information described above, it might seem more reasonable instead to call "metadata" "metainformation" since the markup itself is not entirely arbitrary (as it represents some specific, identifiable pieces of information), and it provides another layer of interpretation beyond the base level of content.

Digital Markup

The metadata of most consequence to this project is the digital markup tags inserted around the content of Beresford's text. XML, of course, was not born as a sudden idea; it comes out of a distinguished line of ever-developing standards. XML and its predecessors operate under the method of separating textual content from visual presentation. Lou Burnard, for one, claims that SGML "describes only the formal properties and inter-relations of the components of a document" (41); it is "decidedly unhelpful about how texts are to be reproduced . . . Its strength is that by separating the notion of what the text actually is from how the text is rendered, it makes possible the use of the same text by many different kinds of processor" (41–42). This claim supplements Goldfarb's earlier assertion that "descriptive markup like [SGML] . . . is employed to assist the reader's comprehension by emphasizing the structural attributes of the document and its elements" like chapters, emphasized phrases, and lists (9). Further, Burnard continues, SGML and its descendants operate on rules that can be set by its users: "A markup language must specify what markup is allowed and whereabouts, what markup is required, how markup is to be distinguished from text, and what the markup means." SGML only does the first three, and "additional semantic information is needed" in order for the markup to be useful (42). Hence, more defined guidelines like the TEI's came out of a need for more specific use cases.

The eXtensible Markup Language, in which this project is constructed, possesses two separate levels of constraint for adherence to the standard, and gives it the ability to be parsed by

multiple people and systems: one of structure and syntax, called "well-formedness," and one of vocabulary, or "validity." As a rule, valid markup is already well-formed, but it is possible to encounter well-formed markup that is invalid. For XML, well-formed markup is a more fundamental requirement, and is markup that adheres to the structural limits of the standard by meeting several basic criteria: there is one root element, all open-tags are matched to corresponding close-tags, and the hierarchical tree-model is preserved. For the earlier SGML as "a rigorously articulated logic for marking the structural parts and relations of textual documents" (McGann 4), stress is placed on "conformance" rather than validity or well-formedness (Goldfarb 478). SGML is essentially the progenitor of many of the markup languages commonly employed today, including XML and HTML.

A well-formed XML document on its own, though, must offer some specifications in order for its purpose to be understood by more than the creator of the document. To assist in such interchange, an XML document can be validated against a custom set of constraints that might include allowed elements and attributes, as well as where those particular elements are allowed to appear, and in which particular order. This fundamental capability of XML is actually referenced by the 'X' for "eXtensible." Despite XML's resemblance to HTML, for example, the latter is *not* extensible: unlike HTML, XML "does not consist of a fixed set of tags" (Burnard and Bauman xxvii). Despite the potentially limitless customizations that such extensibility offers, it is nevertheless possible (and often necessary) to create automated methods of ensuring adherence to custom rules and vocabularies. For this purpose, digital markup parsers can compare marked-up documents against various validation methods, including "schemata" (singular "schema") and "document type definitions" (DTD). These markup tools can constrain and define elements, attributes, and rules.

For the purpose of common, shared validation schema for digital humanities scholars, the

XML vocabulary of the Text Encoding Initiative was born. Providing a continually updated

collection of validation scenarios, the TEI Consortium works to keep an up-to-date and relevant

set of agreed-upon definitions for humanities texts. Now, years later, in response to the evolving

needs of the digital scholarly community, the TEI *Guidelines* are updated and published

regularly every year to reflect new technologies, new goals, and new challenges that arise as

more people adopt the approach of digital markup. Susan Hockey, in praise of the TEI, writes,

> If one humanities computing activity is to be highlighted above all others, in my view it
>
> must be the TEI. It represents the most significant intellectual advances that have been
>
> made in our area, and has influenced the markup community as a whole . . . In many
>
> ways the TEI was ahead of its time, as only with the rapid adoption of XML in the last
>
> two to three years has the need for descriptive markup been recognized by a wider
>
> community. Meanwhile, the community of markup theorists that has developed from the
>
> TEI continues to ask challenging questions on the representation of knowledge. (16–17)

Critically, though, Nancy Ide and Michael Sperberg-McQueen are careful to note that the TEI

offers *guidelines* rather than strict *standards*; indeed, as they write only a year after the

establishment of the TEI, they recognize the necessity of adaptability with markup languages,

and that no single standard can effactually be enforced—nor should it. They argue, "It was

recognized from the outset that the Guidelines will be successful only if they prove useful to

those who are actually encoding texts" (13). Certainly it can be appreciated that what proves

"useful" will change over time, and as technologies change it is important for scholars to

appreciate that the guidelines have "been developed not by any standards agency, funding body,

or other authority, but by and for the research community, [so] researchers who do not find it

useful remain free to ignore it entirely and to develop their own encoding scheme whenever they see fit" (8). Of course, a major benefit of conforming to guidelines laid out by the TEI is its relatively widespread adoption; it is always possible to provide a custom set of arbitrary definitions for other scholars to follow and participate in one's work, but that work will be more accessible if there already exists some common ground among participants.

While the situation of text encoding has vastly improved in terms of acceptance and support in the years since GML and SGML, nevertheless it remains a young and fluid pursuit, especially in the humanities. Groups like the Text Encoding Initiative have made significant progress in merging the traditional work of scholars with the digital, and in bringing awareness to some of the possibilities offered by the technology. It is of course important to note that TEI-XML is not the only form of digital markup suitable for digital textual representation, nor does it lack its critics; indeed, a significant reality to acknowledge regarding XML is the fact that it has limitations and is by no means a universally accepted technology. Beyond the potential difficulty introduced by applying a structural model of encoding to texts that might lack such a rigid structure, it is interesting that one of the most cogent criticisms of the TEI comes from one of its most staunch supporters, Sperberg-McQueen. Of it, he argues, "The intellectual integrity of materials encoded with the TEI encoding scheme is harder to guarantee [than plain text]. With the TEI, as without it, integrity remains inescapably the responsibility of the creator of an edition; all that the TEI can do is to provide the mechanisms needed to allow textual critics to create intellectually serious electronic editions using the TEI encoding scheme." In other words, it is *because* of its customizable nature that the TEI (and XML more generally) has the potential to become more of a hindrance than an ally. Nevertheless, it is the focus of this project, and as George Landow argues, "One of the fundamental strengths of XML, of course, lies in its creation

of a single electronic text that can lend itself to many forms of both print and electronic presentation" (107)—it serves as a versatile and approachable "single source." As long as texts are represented faithfully and consistently, and markup choices are documented thoroughly, many alternative markup approaches are viable.

Markup beyond the Purely Technical: Humanities Computing

While a more technical discussion of XML and its use within the project follows in subsequent chapters, here are discussed the reasons *behind* the use of markup—with a specific focus on XML—for study in the humanities.

In the 1980s, SGML was the de facto markup standard for document encoding, but it was insufficient for less technically inclined researchers who were beginning to develop electronic scholarly editions. In 1988 came one of the original documents calling for "a markup standard for encoding literary texts . . . The existence of such a markup standard would greatly facilitate the interchange of texts between different researchers, archives, and textual analysis packages, and between publishers and researchers" (Barnard et al. 265). Its conclusion was that SGML seemed to be the best basis for the development of a new standard, but was itself not suitable for researchers not familiar with the technology. Also important, Barnard et al. note, is the necessity of markup "minimization" to reduce the amount of tagging necessary by an encoder to "a tolerable level" compared to the overly complex SGML for such purposes (275). Such desires— and warnings—have remained consistent since 1988. While markup specifically designed by and for humanists has come into use, the same guiding principles have endured.

Christine Borgman's claims of "[i]mprovements in searching, analysis, and visualization tools" (115) encompass a wide variety of digital technologies and projects. In this same vein, Thomas Rommel asserts that, as a direct result of the rise of computer use in the humanities,

"textual analysis as well as the ensuing interpretation of a text as a whole can be based on a complete survey of all passages that promise results, no matter how long the text is." Moreover, he continues, "Comparative approaches spanning large literary corpora have become possible, and the proliferation of primary texts in electronic form has contributed significantly to the corpus of available digital texts" (89). That is, in recent decades, it has become possible to quickly analyze not only very large individual texts, but also very large collections of texts. No complete survey of all these technologies and texts seems even remotely feasible, but there certainly exist notable examples of approaches and projects that scholars have undertaken to take advantage of the continually improving access to digital technology.

Echoing Borgman's suggestion that "improvements" by means of computers assist scholars in interpretive work, Rommel details three specific benefits made possible by literary computing since the 1960s and 1970s:

1. "virtually unlimited access to high-quality electronic texts;"
2. "sophisticated software that lets the user define the terms of analysis rather than vice versa;" and
3. "powerful computing equipment that supplies unlimited computing power and storage capacity" (93).

Certainly these three benefits are vaguely phrased, but they hold up under scrutiny: as time progresses, the price-to-storage-capacity ratio continues to decrease, and computer processing power continues to improve. However, as Rommel also acknowledges, "literary computing still remains a marginal pursuit" despite the above "impressive advances" (93). In *Radiant Textuality*, Jerome McGann explores this sentiment in more depth: "[T]he general field

of humanities education and scholarship will not take the use of digital technology seriously until one demonstrates how its tools improve the ways we explore and explain aesthetic works—until, that is, they expand our interpretational procedures" (xii). Being the lead editor behind the Rossetti Archive, McGann has already attempted to demonstrate the way in which digital tools can improve humanities scholarship. At the time of the publication of *Radiant Textuality*, the Rossetti Archive was already available in its first incarnation, so McGann was thoroughly invested in demonstrating to the field at large that digital technologies had something significant to offer. In fact, some of the most notable examples of "improvements" enabled by digital technologies include online scholarly editions and archives, of which Jerome McGann's Rossetti Archive is but one instance.

Whereas textual formatting can indicate a great deal about meaning and authorial intention, structural markup distinguishes form and content, even in instances of the same font indicators. Consider italicized text, for example. While *emphasized words* and the title of the book *Radiant Textuality* may both be represented by italic font, the markup for each is different based on meaning. In the former, one might use the TEI element <emph> for emphasis, while the latter is tagged as a book <title>. In this instance, the distinction is that of semantic meaning compared to typographical appearance conventions: the former is an interpretive tag while the latter is a descriptive one. As the TEI *Guidelines* are clear to state about the difference, "[I]t may often be difficult to make a clear distinction between details relating purely to the rendition of information and those relating to the information itself" (459). Following on this notion, Sarah Arroyo contends that practitioners of good markup should mark up a text based on *what* the information is rather than *how* it should be displayed, avoiding shallower "physical" or "visual" tags with "nonsemantic meaning" such as HTML's <b> for bold text in preference of "logical"

or "structural" tags like <strong>. Both of these preceding tags will produce the same visual result in most conventional browsers, but the *meaning* assigned to the element's text gains a layer of interpretation with the latter tag—there exists a specific meaning that a markup author wishes to convey through a logical, semantic, non-ornamental element. In sum, "The crucial difference between XML and other pre-existing markup languages such as HTML is that it uses tags to assign meaning to the actual content being marked up. This allows applications to be developed that truly separate the data content from the formatting and structure aspects of the document" (36). Again: XML is designed to address the content—*not* the form—of a text.

In nearly any example of markup in practice, one common rallying point is the decision of how the content and the form of a text interact with one another, and how a visualization method should treat either (or both). In truth, who gets to decide isn't even necessarily clear; in a scholarly edition like McGann's, the choice of how far apart to keep content and form is the decision of the editor. On the other hand, the same final "output" can be achieved with several various methods. Furthermore, "content vs. form" is no simple discussion; McGann himself, in "The Rossetti Archive and Image-Based Image Editing," states, "From a traditional literary scholar's point of view, SGML markup limits itself in several important ways. Most significant, it does not adequately recognize the radical difference between an aesthetic and an informational textual structure" (147). To McGann, then, using digital markup—in particular SGML, the Standard Generalized Markup Language—is insufficient by itself to indicate the separation of content and form. He describes the tension that must be addressed between the duties of a critical editor and those of a facsimile editor.

The OHCO thesis of DeRose et al. is especially useful in relation to the use of markup for digital representation. The series of content objects that make up any text "occur in a certain

order," and "smaller content objects do not cross the boundaries of larger ones; thus a paragraph will not begin in one chapter and end in the next. For this reason, the structure of a document is a hierarchical one, like a tree or taxonomy" (6). Such a model is a foundation of descriptive markup, and while it precedes the XML specification, it essentially describes it, whereby XML "elements" are analogous to "content objects." It is important to reiterate that, in this model, the content objects—paragraphs, quotations, and other blocks of meaning—are distinct from their appearance; in the OHCO model, a text is defined primarily by its content and structure rather than its rendering or form.

Dave Clark observes that the separation of content and form is "taken as a given" while at the same time it receives "little critical scrutiny." Clark's argument centers around the fact that many scholarly editions are built on standards that already inherently separate their content and their form. XML, for example, can be harnessed to keep content and form very distant or very close. Consider the following possibilities:

1. XML with in-built Cascading Style Sheets (CSS) language;
2. XML with separate paired visualization coding using CSS; or
3. XML with separate paired transformation coding using eXtensible Stylesheet Language (XSL) transformations.

In the first case, a single document contains all the necessary information to describe both content and form. This approach can be useful for small presentations or individual files that one might wish to display using a suitable program such as a web browser. Because of redundancy and the possibility of conflicting CSS definitions, however, any system with multiple documents would be much better suited to the second situation in which the content—all the textual data—is

contained in one file while the form—the visualization information—is contained elsewhere, in one or more separate CSS files. In this manner, consistency is maintained across any number of documents that refer to the CSS definition, and any changes in the CSS influence all content documents. Indeed, this approach is a standard practice for websites.

Finally, the third example is one step removed from the second, and is largely how electronic archives operate. As in the second case, XML content is stored separately. Transformation stylesheets, written in another XML standard called XSL, contain instructions to automatically produce entirely separate documents by transforming the structure and content of the original (Kay 3), which can be nearly any other plain-text file type—including other XML and CSS combinations as mentioned above. Typically, though, this practice is employed to convert XML into a more visually appealing HTML website, where instructions in the XSL convert XML elements like <emph> (for emphasis) into strictly stylistic tags like <i> (for italic). This third example is typically employed as part of a larger system that Clark calls a CMS, or "content management system," and is how my edition operates. Specific examples and XSL rules are detailed below in Chapter Six. For Clark, a CMS is

> a system that approaches the problem of content management by using markup, metadata, and tools to break documents into component parts, to a level of granularity (e.g., paragraph level, sentence level, word level) set by organizationally defined information models, and labeling each part with metadata that describe its meaning and relationships to other content. The same content can then be automatically assembled in different genres, with different presentations, and in different media. (39)

In essence, Clark's definition of a CMS describes the approach behind many electronic scholarly editions, including McGann's Rossetti Archive. In short, "Separating content from presentation

and systematizing presentation affords the easy maintenance of visual consistency" (55). This positive outlook on the CMS influences Clark to argue that content and form separation is a "look to the future" by "making the presentation of Web documents independent of hardware platform, software, culture, disability, etc." (40). Though idyllic sounding, this hope seems, unfortunately, idealistic, especially in the context of humanities computing.

As demonstrated by Clark's argument, a great strength of a non-proprietary format like XML is its ability to be rendered by multiple systems. My artifact, for example, was built and tested primarily within the California Digital Library's eXtensible Text Framework (XTF) system as well as John Walsh's TEI Boilerplate implementation. By no means are these the only methods or platforms to represent a TEI-encoded text. In fact, the nature of open and customizable markup is the possibility of interchangeability. An open and well-defined markup schema allows many programs and systems to represent the text without the need for specialized or proprietary software. Indeed, it is even possible to produce not only a readable rendition for a web browser, but even an aesthetically pleasing one using CSS rules built into the markup file itself.

In his brief summary of SGML, Tim Berners-Lee explains, "An SGML document is marked up in a way which says nothing about the representation of the document on paper or a screen" ("SGML"). A defining concept in digital scholarship is the separation of content and form, including whether—and if so, how—to distinguish those layers of a text. For some scholars, such as those working with chirography and original manuscripts, the form of the text may be as important to consider as the content, while others' interests might be solely focused on what meaning the textual content conveys, wholly unrelated to the artifact's physical characteristics. In either case, a digital text must have adequate means of representing the

features of the artifact it hopes to mediate. Further, Jay David Bolter claims, "the goal of representation has been transparent presentation. The medium is supposed to function as a window through which the viewer can see the objects represented" (25). In this sense, the medium is analogous to the form of the text, and it should merely assist in the consumption of the content; it should not, itself, be the focus of the text. Certainly this belief is not universally held; it is the goal of many avante-garde genres, including early hypertext electronic literature, to make the reader blatantly aware of the medium, causing it to be as much a part of the text as the meaning-content itself. In scholarly databases of pre-digital authors, however, an approach that elevates the mode of representation to the level of the content would most likely be considered distracting.

In the realms of digital markup, HTML is specifically designed to describe and manipulate a text's form whereas XML's purpose is less clear. In a system that employs XML markup, the XML itself usually serves as the system's content, its semantic level of representation, while form is manifested by the transformative markup of stylesheets, such as CSS and XSLT. In this arrangement, the visual arrangement of a document deserves as much attention as the information itself since the arrangement is mainly responsible for separating the two. Even a coded document like an HTML or XML one reflects deliberate choices by an author (Applen and McDaniel 131–132). XML is used to describe the *what*, and XSLT/CSS to describe the *how*.

On the other hand, "The worst aspect of visual systems," says Leslie Lamport in 1988, "is that they subvert the process of communicating ideas by encouraging the writer to concentrate on form rather than content" (9). That is, representation of information should be a secondary concern, and not a primary concern of an editor. Using logical systems—in this case, descriptive

markup—the content can be represented in many ways after it is produced, a process often referred to as "single-sourcing." Interleaving the content with the form ensures a narrow scope of technologies that will properly render the text, and greatly increases the difficulty of working with the content in the future—clearly not ideal for scholars who wish to maintain a text's longevity. The separation of content and form facilitates the "interchange of documents between different systems" (Berners-Lee, "SGML") by keeping the information about the medium separate from the data with semantic meaning.

As part of Dave Clark's argument that this separation is "taken as a given," he argues that "separating content from presentation can not only save time but allow for rapid reuse and repurposing of content. A single piece of content, properly marked and stored, can automatically and simultaneously appear in user manuals, help files, and press releases that can in turn be automatically altered to appear in print, on the Web, or on mobile devices" (36). The scenarios Clark describes here are further examples of single sourcing, by virtue of their using a single digital source to produce a variety of output materials. "Separating content from presentation and systematizing presentation affords the easy maintenance of visual consistency" (Clark 55), an important goal in both presentation and preservation.

In undertaking a digital project, it is essential to plan ahead. For example, are white space and indentation parts of the content or form of a particular document? How much of it should be represented as exact? Does the creator want a facsimile representation or just to retain the content? That is, how much of the form should feature as relevant in the coding? A middle ground is perhaps possible via the acquisition of facsimile images; even if "only" the document's structure—the content—is coded, providing images of the original artifact can offer a context far beyond that of a remediated transcription.

Many of the model projects that inspired this work function on such an approach: that is, while not all of the presentational features like font and column breaks are rendered, such features can still be seen by readers in the images of the original. Indeed, it has been said of facsimile images, "The primary purpose of a facsimile in editing is to provide editors and textual scholars with a reliable version of the source document or documents" (Kiernan 262). This both saves the time and effort required of coding said features and allows a separate evaluation of the content and the form. Regardless of the decision of how much content-versus-form to represent, such a decision is necessary early on for the success of a project. Further, when it comes to planning, what is the future of the project? Will other scholars contribute to it over years? If so, it is crucial that they know the system employed at the beginning and how to make it work with whatever system they will be using.

Of course, the above technologies and conventions were not born of the great chaotic void that separates academic institutions and sometimes even exists between departments; XML as it stands today is the product of decades of development in descriptive markup. When the Standard Generalized Markup Language was conceived in 1986 (Burnard 41), the exigence of its creation was the hope of ensuring consistency, both for the ease of markup processing programs and—at least as importantly—for the improved ability of human readers and scholars to understand what other scholars have chosen to encode (47). At this point, it becomes important to distinguish between two key concepts that have always been at the heart of markup scholarship and have regained a focus in recent discourse: "interchange" and "interoperability."

Syd Bauman, in a presentation for the annual Balisage markup conference in 2011, discusses the two in great detail. He begins the analysis by reasserting the above benefits of digital technology described by Rommel, as well as adding that encoding texts digitally with

markup helps digital humanists "share the text with others, whether for reading, similar analysis, or for analysis we might not have thought of or even imagined." The act of sharing, in this case, becomes the focus of Bauman's investigation; that is, he breaks sharing down into two types, either interchange or interoperability. In either type, data are transferred by some means—for example, through download or direct copy transfer—but what happens once the data are transmitted determines which type of sharing occurs.

In the case of interchange, human interaction ensues: data cannot simply be transferred and placed into a new system without some kind of manipulation based on discussion and understanding. As Borgman says, "making content that was created for one audience useful to another is a complex problem" (10). For example, should I wish to incorporate a previously coded XML document into the Charles Brockden Brown Electronic Archive, I would need 1) to know the conventions under which the creators of said XML document operated; 2) to know the conventions that our XTF instance requires; and 3) to edit said XML document so that our instance of XTF can properly index and display it. The overall process requires some interpretation and balance, unlike the other type of sharing: interoperability.

With interoperability, no human intervention is required: to continue with the previous XTF example, this would be the equivalent of downloading another scholar's XML file, placing it into the XTF data directory, and the process is done—the new file can be rendered identically to all the other files without any changes whatsoever. As Bauman is quick to note, however, "interoperability is difficult" and should not be the overarching goal of any project. Indeed, he says, "For humanistic scholarship variation of expression is a necessity." He claims further that "interoperability and expressiveness are competing goals constantly in tension with each other"; the stricter the standards become to allow pure interoperability, the less room is left for

individual or project-specific interpretation. Rather than striving for universal interoperability, then, interchange is both a more worthy and more achievable goal. If nothing else, it necessitates some form of communication among scholars, and is likely to spark continued discourse and inspire new practices across disciplines and institutions.

Standards can certainly offer the possibility of interchange and, in some cases, interoperability. Unfortunately for Dave Clark, his "look to the future" (cited above) of widespread adoption of content management systems is essentially a call for a worldwide interoperability standard, a tall order indeed, even without taking into account the differences that exist within disciplines and across departments. The only possible way to achieve such an ideal would be the widespread adoption of strict standards, which would strangle innovation and customization in text encoding, as Bauman asserts. Besides the possibilities of interchange and interoperability, less stringent standards for humanistic metadata ensure that data creation and analysis is not proprietarily bound: whether data be taken to mean entire texts or individual pieces of character data, the only way to ensure that many scholars can work with said data is to create an agreed-upon set of conventions, which is indeed what standardized markup languages hope to do. Adriaan Van der Weel, in his chapter "The Concept of Markup," avers that markup, just like writing systems, is governed by conventions. If everyone used his or her own arbitrary conventions in writing, little would be intelligible or communicable outside of a very narrow set of users, and the same is certainly true of markup.

As noted earlier, Lou Burnard argues, "A markup language must specify what markup is allowed and whereabouts, what markup is required, how markup is to be distinguished from text, and what the markup means" (42). This statement is made in the context of supporting the adoption of a standardized digital markup by literary scholars. In short, Burnard believes that a

markup language should have its basics laid out so that diverse users can employ it and understand how other users may have employed it differently. His desires translated to the development of the Text Encoding Initiative, a set of guidelines for digital markup to be used by humanities scholars. As Bauman concludes, digital humanists should aim for a balance of interchange and interoperability, which "requires a lot of adherence to standards (e.g., TEI), but eschews mindless adherence that curtails expression of our ideas about our texts." In this manner, all can understand the underlying markup syntax and grammar, and refer to documentation to explain project- or institution-specific customization choices.

In fact, when it comes to the actual representation of the document, digital markup makes the interchange of documents viable; that is, while so-called "high-level markup" like SGML and XML delineates the structure of a text, it "says nothing about the representation of the document on paper or a screen" (Berners-Lee, "SGML"). Ideally, such structural markup indicates little to nothing about the form (that is, the formatting and layout inherent to a specific system) of a document, and everything about the content itself. Further, this fact "is invaluable when it comes to interchange of documents between different systems, providing different views of a document, extracting information about it, and for machine processing in general."

To lend further support to the increasing necessity of digital literacy, Katherine Hayles in *Electronic Literature* points out that all published writing now comes out of a digital file; thus, even "printed" artifacts are affected by code. Specifically, she says, "So essential is digitality to contemporary processes of composition, storage, and production that print should properly be considered a particular form of output for digital files rather than a medium separate from digital instantiation" (Hayles, *Electronic Literature* 159).

It might be asked why the use of digital markup should be considered by non-computer scientists. Besides the apparent answer—that in the ever-changing realm of academics, just as in everyday life, computers are becoming more of a necessity—preservation, access, and single-sourcing are each enabled by the digital representation of texts. Given the ability to create a single text that serves multiple needs and can be produced in multiple formats, it is surprising that so few scholars actually engage in the practice.
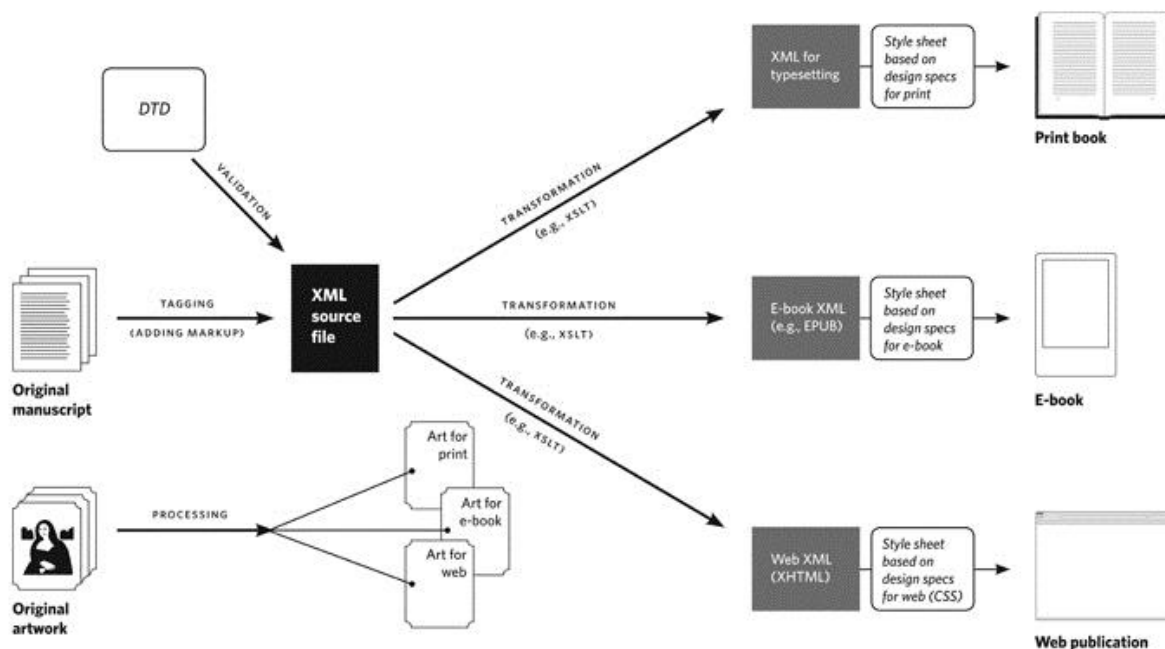
Figure 1: The *Chicago Manual of Style*'s XML workflow diagram, providing a summary of how an XML document can function as a single-source file

http://www.chicagomanualofstyle.org/tools_workflow.html

Recall George Landow's claim in *Hypertext 3.0* that "One of the fundamental strengths of XML, of course, lies in its creation of a single electronic text that can lend itself to many forms of both print and electronic presentation" (107). Beyond this benefit, by employing a non-proprietary binary format like XML for the purpose, no restrictions in terms of software

licensing exist: a typical XML file can be viewed and manipulated on any operating system with a standard text editor. Since no specialized software is required, the threshold for learning and understanding the technology can be substantially reduced.

Text in Code

There exists an interesting layer of abstraction in the production of a digital edition, especially one created with customized structural tags such as this one. The associated marked-up edition is both a case-in-point as well as an original contribution; it provides my own poetical analysis of a translation of a famous epic poem as well as an example of how digital markup technology might be applied for preservation and interpretation—that is, for the creation of digital scholarly editions. Moreover, this very document serves as a meta-discussion of the process it describes: creating digitally marked-up works. It is born not only as a digital file, but as a structured XML document. That is, it was written *as* an XML document, and its production did not take place within word-processing software.

While working in a purely digital environment is convenient, efficient, and relatively unhindered compared to physical writing, several scholars have considered the burgeoning dilemma this practice presents. In *Electronic Literature*, Katherine Hayles concisely claims that "Literature in the twenty-first century is computational" (*Electronic Literature* 43). More recently, Matthew Kirschenbaum and Doug Reside have expanded on Hayles' statement, contending that "a writer working today will not and *cannot* be studied in the future in the same way as writers of the past, because the basic material evidence of their authorial activity— manuscripts and drafts, working notes, correspondence, journals—is, like all textual production, increasingly migrating to the electronic realm" (260). In most contexts, the only portion of the writing process that future scholars will have to study is the final product itself, without any of

the working drafts, notes, or research that went into that product. Incremental versioning

software exists, of course, but by no means is its application ubiquitous. Not only do the practical

problems of capturing and preserving electronic ephemera exist, though; so, too, does the

consideration of *what* will be worth collecting and evaluating in the future. As Kirschenbaum

and Reside continue, what types of data on an author's electronic storage media are relevant or

too personal when exploring the context and exigence of a born-digital document (262)? How

much data is too much? These questions do not yet have definitive answers, especially as very

little work has been conducted on the study of digital-age scholars and authors.

Markup as Digital Ekphrasis: A Meta-Discussion of the Metadata

Consideration of an author's writing process, be it one's own or someone else's, is

arguably an interpretive endeavor. Such consideration comprises an assessment of a single point

or span of time, not necessarily in contextual isolation, but certainly as a focal point that pushes

other considerations to the side while the interpretation occurs. To examine a specific span of

time is also to pause consideration of all else. As Svetlana Alpers argues, description is a

"pictorial point"; more pointedly, ekphrastic description is a "singular combination of an

attention to imitation or description with a suspension of narrative action" (15). Thus, in the

description of an item, be it the armor of Aeneas or the XML tags surrounding a rhetorical figure

employed by James Beresford in his translation of the *Aeneid*, narrative action ceases while said

description occurs.

James Heffernan and W. J. T. Mitchell after him describe "ekphrasis" as "the verbal

representation of visual representation" (Heffernan 3; Mitchell 152). Mitchell continues,

"Ekphrastic poems speak to, for, or about works of visual art in the way that texts in general

speak about anything else" (159). Both Mitchell and Heffernan offer unique depth and insight to

the concept in terms of poetry and physical art including paintings and sculpture, but at its core, ekphrasis applies to the broadest definitions of text—those creative works that are capable of being interpreted. While Heffernan, Mitchell, and others are primarily concerned with verbal representations of physical works, even a painting or a sculpture itself is ekphrastic as it attempts to capture or represent a single event in stasis. Prolific and highly regarded Virgil scholar Michael C. J. Putnam distills the essence of ekphrasis as "art describes art" (ix). Whether applying to sculpture, epic poetry, or XML, this definition carries beyond its traditional role and into the digital age. As long as remediation is involved, it is reasonable to claim that ekphrasis is present as well.

Heffernan further defines a useful subset of ekphrasis that he calls "notional ekphrasis," a representation of an imaginary work or object that causes a reader "to see how the poem reconstructs it, how the poet's word seeks to gain its mastery over the painter's image" (7). Notional ekphrasis comprises the representation of imaginary features, but as such gives more power to the representational words than the object since there is no "original" against which to compare. In short, the verbal description employed in any instance of ekphrasis "actively describes what is not there" (Bartsch ii). In the case of ancient epic poetry, wherein most ekphrastic descriptions are of imaginary or temporally distant objects and events, nearly all instances of ekphrasis will be notional.

Putnam expands, "Ekphrasis is an encapsulation of visualized art wherein a verbal medium strives both to delineate a spatial object and, even while beholden to narrative's temporality, miraculously to capture the instantaneity of a viewer's perception" (208). Moreover, in particular to the *Aeneid*, "All Virgil's examples of notional ekphrasis . . . help us contemplate . . . the larger poem for which they stand as synechdoches. We pause as we become visionaries

before art's stillness" (208–209), even as the surrounding narrative does not cease while the audience contemplates the artful description. Putnam's eloquence here elucidates some scholarly opinion on Virgil's employment of ekphrastic representation, and raises another foundational tenet of the device.

In its traditional form as engaged by Heffernan, Mitchell, and Putnam, ekphrasis "originates as a seemingly ornamental adjunct to a larger text, a descriptive digression from the main line of epic narrative" (Heffernan 137); that is, when an author engages in ekphrasis, the main narrative is paused while the object of the ekphrastic endeavor is realized. When the shield of Achilles (Heffernan's and Mitchell's exemplar from the *Iliad*) or the shield of Aeneas is described in vivid detail, the narrative cannot progress until the description is complete. Essentially, the authors have decided to dedicate their craft to *showing* the audience the details of an object in order to represent it more faithfully, even when—or perhaps *especially* when—the object can never be seen. Putnam delves deeper into classical ekphrasis, especially as it pertains to the *Aeneid*, arguing that through ekphrasis, "we linger, not to escape the story's flow but to deepen our understanding of its meaning, to watch metaphor operating on a grand scale" (ix). Whether the object exists or not is moot; notional ekphrasis serves the cases of the shields by giving absolute power to the representing medium to explore the object and, indirectly, its importance.

Branching off of Heffernan's core definition, Mitchell constructs a hierarchy of "phases" or "moments" of ekphrasis: 1) ekphrastic indifference, 2) ekphrastic hope, and 3) ekphrastic fear. Each phase relies on the previous as a foundation. Compared to visual representation, such as in photography, Mitchell avers that a verbal representation simply cannot convey as much information about an object. Rather, words "may refer to an object, describe it, invoke it, but it

58

can never bring its visual presence before us in the way pictures do. Words can 'cite', but never 'sight', their objects" (152). "Indifference" summarizes this notion: true depiction is impossible.

"Hope," on the other hand, is "the phase when the impossibility of ekphrasis is overcome in imagination or metaphor," or where language is used to evoke an image in the "mind's eye" (Mitchell 152). Imagery and notional ekphrasis, then, are forms of ekphrastic hope.

Finally, the phase of "fear" is defined by the worry that, if ekphrasis is abused or is *too* descriptive, then something is lost. If the verbal description is too good, as it were, then we might as well be looking at the object itself; however, at that point we also lose the need for the description, and the need for both essentially disappears (155). In the end, Mitchell believes an ekphrastic balance is to be most desired in any representation.

Digital Ekphrasis

Ekphrasis perseveres in a new medium by the fact that digital markup can be as ekphrastic in nature as a poetic description of Achilles' shield. A metalanguage like XML evinces ekphrasis by virtue of the markup's attempt to represent another text via a coding language; features of another document, oftentimes one that exists in a separate medium, are portrayed by the code. Valentine Cunningham further argues of ekphrasis that it is defined by "that pausing, in some fashion, for thought before, and/or about, some nonverbal work of art, or craft, a *poiema* without words, some more or less aestheticized made object, or set of made objects" (57).

Not only is ekphrasis a pause to enable extra consideration; it is also the provision of extra detail about something—thus is markup ekphrastic in much the same way as a parenthetical phrase set off from the rest of a sentence by punctuation. To avoid Mitchell's notion of ekphrastic fear, though, just as with words on a page, the amount of markup and what it says

59

about the text it seeks to represent must be balanced for fear of being too unruly, cumbersome, or simply too detailed to be helpful.

In taking the time to view from afar the processes and theory that make up this project, I am arguably engaging in an ekphrastic endeavor. Not only is the discussion of the project ekphrastic—that is, it interrupts the narrative to examine details about a textual object—but the markup itself is ekphrastic by calling attention to parts of the text that I have structurally identified with custom XML tags. While Mitchell, Heffernan, and others are content to limit ekphrasis to verbal representations of the visual, it seems reasonable to argue that ekphrasis can encompass remediation more broadly; in short, any representation in one medium of an object that exists in another medium is an example of ekphrasis. Under such a revised definition, then, we can open the door for digital humanities to engage more fully with ekphrasis. By classifying a text marked up in XML as an ekphrasis of a printed novel, a new branch is allowed to sprout from the tree of textual analysis and literary criticism.

CHAPTER FOUR: MARKUP IN PRACTICE

Scholarly editions serve as an enduring foundation of critical engagement in the humanities. However, given the rapid growth of new technologies, old methods are less successful as the pace of change and the amount of information available to interpret continue to increase. In order to survive in the digital age, the scholarly edition must adapt. Structural markup technologies like XML are a means to this end, as exemplified by several well-regarded digital editions that employ them and serve as inspiration for the current edition. Using said editions and approaches as paragons of the new vision of the scholarly edition, I am hopeful that academic acceptance of the technologies will increase, and with it innovative ways to apply those technologies.

Why should humanists be bothered with XML literacy in particular? Given such an increasing reliance on the "born digital," humanists are increasingly realizing that interacting with their texts with newer technologies like digital markup in mind can be a useful approach. As Martha Nell Smith—the Founding Director of the Maryland Institute for Technology in the Humanities (MITH)—argues of the transition of the humanities from print to electronic, "In order to move the scholarly monograph into the digital realm, humanists need to embrace the new technologies developing in digital humanities communities of practice, technologies making not only new work but new ways of working possible, especially those that will not become obsolete" (320). Besides the practical use of preservation, digitization and XML markup together allow scholarly interpretation to flourish with fewer constraints than traditional methods. Nor are scholars limited to a small vocabulary predefined arbitrarily. Instead, they can set out to define their own tags for analysis, such as identifying and singling out specific poetical tropes or figures

of speech in an epic poem. The technology allows this to happen *around* the textual artifact itself, and can be represented in myriad ways.

The Electronic Text and the Digital Edition

In the mid-1990s, not long after what might be called the "codification" of the TEI, C. Michael Sperberg-McQueen insisted that "electronic scholarly editions must meet three fundamental requirements: accessibility without needless technical barriers to use; longevity; and intellectual integrity" (41). These defining requirements are just as important now as they were then, and other scholars have reiterated similar desires of digital scholarship in the intervening years.

A significant advantage the electronic version of any scholarly edition possesses is its lack of physical constraint. Kenneth Price, co-editor of the Walt Whitman Archive with Ed Folsom, agrees. In "Electronic Scholarly Editions," he says, "One distinguishing feature of electronic editions is their capaciousness: scholars are no longer limited by what they can fit on a page or afford to produce within the economics of print publishing." Rather than being bound by the confines of a paper volume, which is limited not only by physical space but also by increasing cost proportional to size, digital editions are limited only by the electronic resources available to the scholar. As computer technology undergoes constant improvements in efficiency, storage space and processing power are less limiting factors than they were even a decade prior. Further, there is little reason to expect such efficiency improvements to diminish. As time goes on, the technology capable of creating, storing, and rendering electronic texts and images will only become more accessible.

On the other hand, Martha Smith supplements, "Scholarly electronic editions are not simply a matter of bits and bytes and how they have been encoded to link to one another and to

appear in particular browsers. Dynamic rather than static, deeply encoded, they are designed to enable readers to achieve better understanding of texts and open up possibilities for more sophisticated interpretations" (312). That is, while electronic texts and technologies undoubtedly enable a wider participation in knowledge creation, they are not merely electronic replacements for the traditional physical edition. Instead, the technologies used to produce them also enable deeper engagement with texts that is simply not possible in a physical manifestation of a work.

Also unlike physical scholarly editions, digital ones are able to operate with fewer limitations. For example, Smith states, "In contrast to the constrained visibilities of book representations, access to questions of editorial fidelity and therefore to the editorial process itself is much more obtainable in an electronic edition featuring images of all documents edited as well as their translations into typography" (311). Many current digital editions, like the Charles Brockden Brown Electronic Archive and Rossetti Archive examined below, offer textual transcriptions as well as access to facsimile images of the original documents. Some editions, like the Perseus Project, also provide editorial notes, which are not relegated to footnotes or endnotes, but which can take up as much or as little room as the user desires.

While there exists great opportunity in the realm of digital editions, Dino Buzzetti expresses some concern that humanists fail to see digital editions as anything beyond a digital representation of a previously physical text (45–46). His concerns are certainly valid: using XML with a schema, such as this edition does, for example, serves as a means of allowing textual scholarship and criticism to move beyond the physical page. However, without a specific focus, it becomes easy to understand the potential for resistance to digital endeavors as simply using new technology for its own sake. Thankfully, digital projects exist that have demonstrated the wider possibilities of digital editions crafted with structural markup.

Several projects may serve as examples of markup in use by humanities scholars, with varying degrees of complexity: the Brockden Brown Electronic Archive and the Mark Twain Project Online each serve as examples of how XSL transformations can mold raw XML data into presentable, interconnected databases; Paul the Not-So-Simple demonstrates the possibility of cross-referencing various editions of a single text in multiple languages, again showing some possibilities of XSLT; and the Petrarchive evinces the elegant simplicity of the TEI Boilerplate platform. The preceding projects certainly do not comprise an exhaustive list; however, each of them to some degree has influenced my approach in the current endeavor with the work of James Beresford and his translation of the *Aeneid*.

The first of these projects, "Paul the Not-So-Simple," demonstrates different translations of the same text within multiple linked frames, each frame containing different information: one the original Old Church Slavonic, one Ancient Greek, one Modern English, and a fourth frame for textual notes and commentary. Notably, each frame contains hyperlinked content that causes the other three frames to jump to the corresponding information when clicked by the user. Such an interactive layout enables a powerful and useful approach to textual analysis and comparison.

Figure 2: An interface screenshot from the "Paul the Not-So-Simple" homepage

The above image demonstrates the linked-frame layout that served as an early inspiration for my edition, in which I planned to compare the English and Latin texts side by side. Each frame provides either a translation of the text or linguistic commentary about it.

http://paul.obdurodon.org/ohrid/

A second digital edition of influence to my Beresford edition is the Mark Twain Project Online. Housed at the University of California in Berkeley, the MTPO showcases a customized instance of the eXtensible Text Framework, which was developed at the same institution. This edition presents a polished interface in combination with a robust set of source documents that easily demonstrates some of what can be accomplished with digital scholarship. Notable features

include the choice to include or exclude editorial notes as part of the visible text, as well as to

add specific citations to a temporary collection to analyze and compare different sections of the

text.
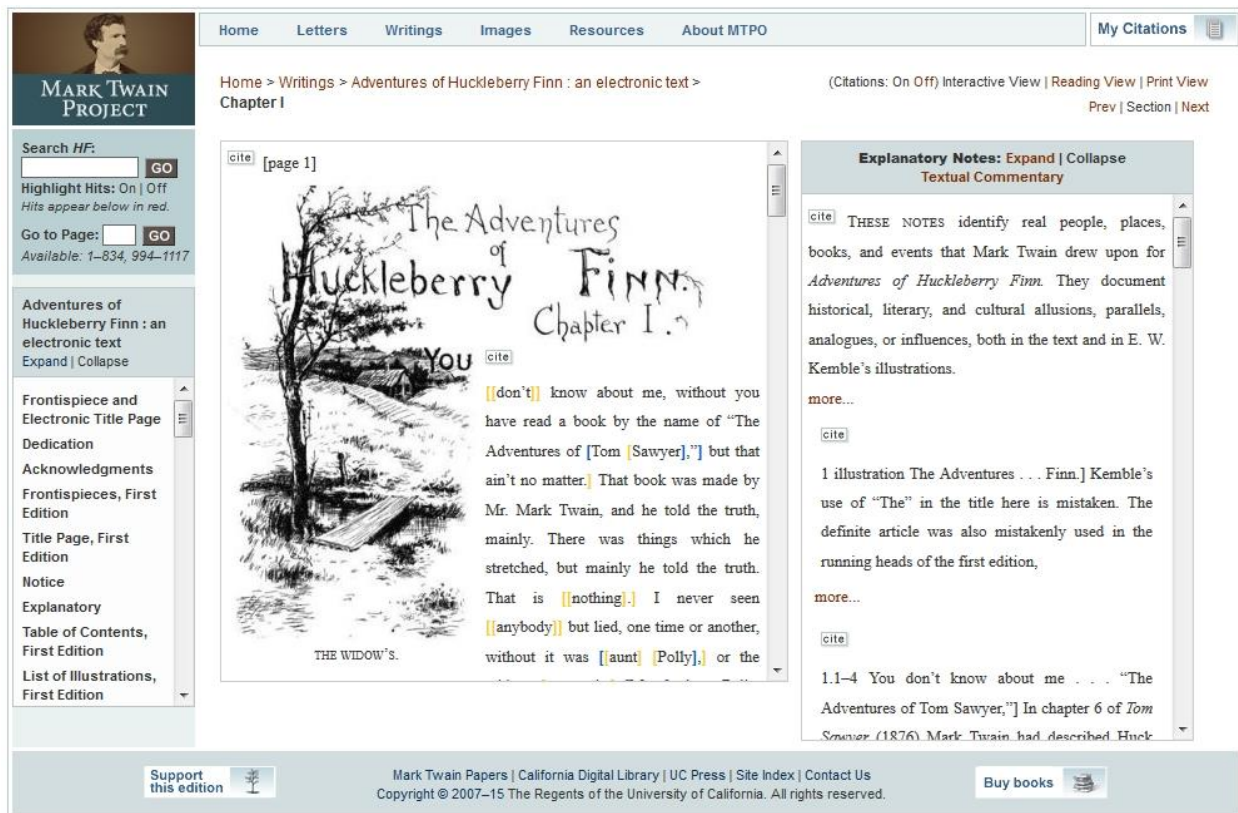
Figure 3: *Adventures of Huckleberry Finn* on display at the Mark Twain Project Online

Represented here is part of the first chapter of Mark Twain's *Adventures of Huckleberry Finn*. At

use in this edition is the XTF platform, customized with the ability for readers to switch between

"Reading View" and "Interactive View," the latter of which is shown here. The interactive view

enables in the center frame the display of editorial note markers, which, when clicked on, take

the user to the corresponding textual note. The reading view shows the same main frames, but

disables the clickable hyperlinks in the main text frame, providing fewer distractions should a

user simply want to read the text without editorial markup.

The Charles Brockden Brown Electronic Archive collects in one location the various collected writings of the eponymous early American author. The collection makes use of XTF to index and render texts in several genres and in numerous source physical states, including printed publications as well as handwritten manuscripts. Each text provides associated facsimile images so that the digitized text can be compared to the original. Additionally, very little editorial intervention was applied to the digitized texts, which influenced my own decision to take a conservative editorial approach to the Beresford text.
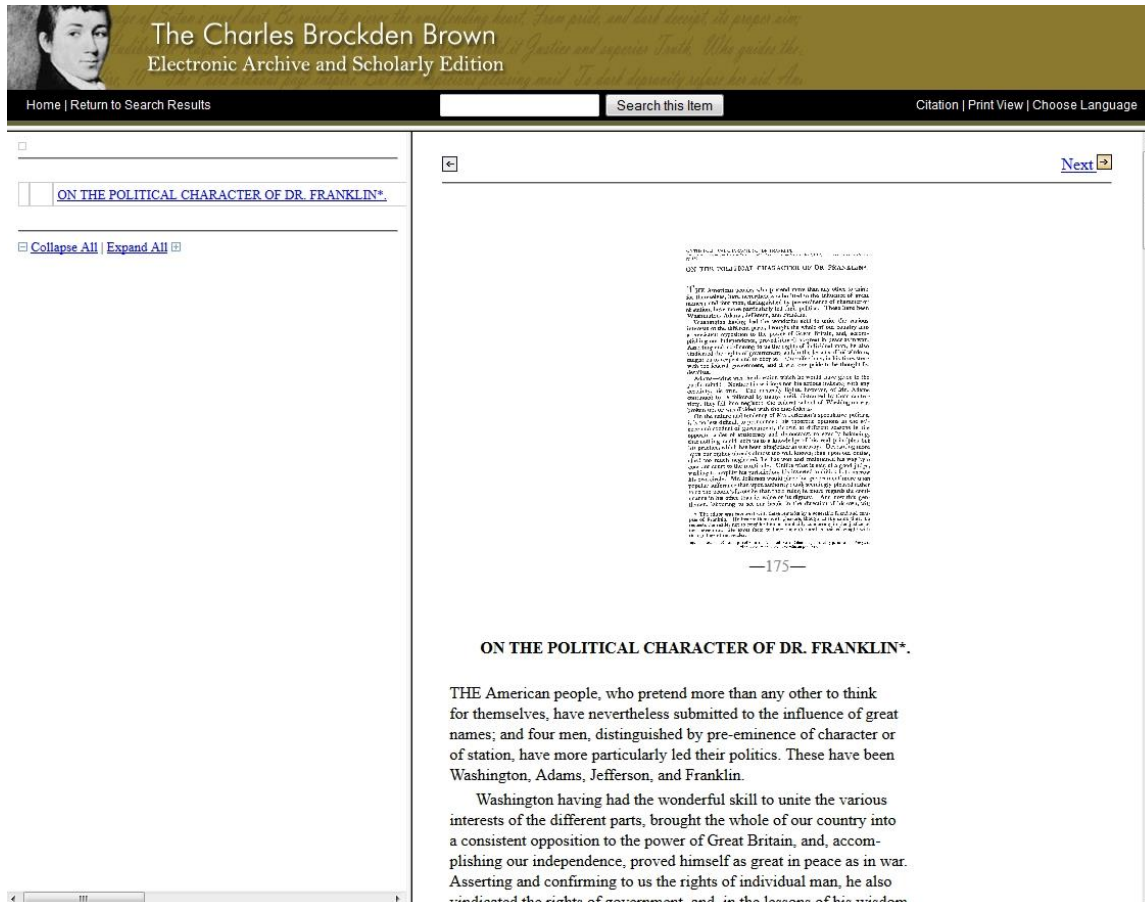
Figure 4: A sample document from the Charles Brockden Brown Electronic Archive and Scholarly Edition

Shown here is a document rendered by a different implementation of the XTF platform. As with other digital editions, the facsimile image is provided alongside the transcription of each text for nearly 1,000 texts of all manner by the Early American author Charles Brockden Brown, including essays on various cultural and philosophical topics, poems, historical pamphlets, and novels (Kamrath et al.). The Brockden Brown digital edition is of special importance and exerts a greater influence due to my personal involvement with it as a TEI coder and technical assistant; in fact, it is because of its implementation of the XTF platform that I was initially willing and able to employ the same platform for my own work.

http://www.brockdenbrown.cah.ucf.edu

Developed and maintained at Indiana University, the Petrarchive functions as an ongoing effort to encode and represent the Italian poet Petrarch's *Rerum vulgarium fragmenta*. After meeting the goals of a 2013 grant to create a functioning prototype for a selection of the poems, work continues to complete the collection. Notable from this project is its use and effective customization of TEI Boilerplate, the content management system developed by John Walsh for use by the archive. The system affords viewers the ability to choose between diplomatic transcriptions and edited views of the text, as well as provides facsimile page images and a browsable image gallery (Storey and Walsh). Other benefits of this system and its implementation in my current project are discussed below.

Petrarch's *Rerum vulgarium fragmenta*: Charta 1 verso

text view [diplomatic transcription ▼]

charta 1 verso

5   Quando io mouo i sospiri a chiamar uoi   El nome che nel cor mi scrisse amore /

     Laudando sincomincia udir di fore   Il suon de primi dolci accenti suoi.

   5   Vostro stato real chencontro poi /   Raddoppia alalta impresa il mio ualore.

     Ma taci grida il fin che farle honore   E daltri homeri soma / che da tuoi.

   9   Cosi laudare & reuerire insegna   La uoce stessa pur chaltri ui chiami

     O dogni reuerença et donor degna.   Se non che forse apollo si disdegna.

     Cha parlar de suoi sēpre uerdi rami   Lingua mortal presūptuosa uegna.

6   Si trauiato el folle mi desio   A seguitar costei chen fuga e uolta.

     Et de lacci damor leggiera & sciolta   Vola di nançi allento correr mio.

   5   Che quāto richiamando piu lenuio   Per la secura strada men mascolta.

     Ne mi uale spronarlo o dargli uolta.   Chamor p̄ sua natura il fa restio.

   9   Et poi chel fren p̄ força a se raccoglie /   Imi rimango in signoria di lui /

     Che mal mio grado a morte mi trasporta.   Sol p̄ uenir allauro onde si coglie

     Acerbo frutto. che le piaghe altrui   Gustando afflige piu che nō conforta.

7   La gola el somno et lotiose piume   Anno del mondo ogni uertu sbādita.

     Onde dal corso suo quasi smarrita   Nostra natura uinta dal costume.

Figure 5: A sample image from the Petrarchive

The Petrarchive showcases the use of John Walsh's TEI Boilerplate project for displaying

multiple views of a text.

http://dcl.slis.indiana.edu/petrarchive/

Finally, the Poetess Archive at Texas A&M University encompasses the time period surrounding Beresford's own work: the so-called "poetess" tradition, or "the extraordinarily popular, but much criticized, flowery poetry written in Britain and America between 1750 and 1900" ("Poetess"). In this archive, editor Laura Mandell has gathered a TEI bibliography of those British and American poets who fit the poetess tradition. However, beyond simply serving as a repository of information about various poets within a 150-year timespan who wrote in a certain style, the Poetess Archive is also notable for presenting at the 2012 Digital Humanities conference in Hamburg the work of Manish Chaturvedi of Miami University of Ohio: a visualization tool called Myopia. Most crucial regarding Myopia is its ability to render poetic features of a TEI-encoded text in an interactive and visually engaging manner, including rhythm, meter, and even poetic tropes, albeit to a limited extent. While at a first look this tool might seem ideal for my current project, it is unfortunately most suited to client-side presentation only; that is, it requires the installation of a significant software support framework, including Python as well as a three-dimensional graphical rendering engine that might be used for video games. Such software enables powerful visualization for analysis, but lacks the benefits of easy installation and, hence, wide online deployment. Nevertheless, in order to work, Myopia requires a customized TEI schema not unlike my own, wherein XML elements such as <assonance> and <consonance> are employed. As of the present, the Poetess Archive has only encoded poetic tropes in John Keats's "Ode on a Grecian Urn" for display using Myopia.

While the preceding projects influenced or parallel the approach I have taken here, their content is largely irrelevant to it; it is instead their form and layout I have considered in the creation of my own edition. On the other hand, there exist several digital projects that intersect both the classical content of Beresford's text *and* the genre of the digital edition, perhaps the

most renowned of which is the venerable Perseus Digital Library at Tufts University. The Perseus Project provides an extensive offering of Latin, Ancient Greek, Arabic, Germanic, and other texts prior to the twentieth century, all encoded in TEI-compliant XML, many with companion-encoded English translations. The extensive collection contains more than 160 million words, with the greatest concentration of those being in the classical texts (nearly 69 million words) and nineteenth-century American writings of various genres (over 58 million words). The Perseus presentation of the *Aeneid* includes English translations by John Dryden (1697) and Theodore C. Williams (1910), as well as textual notes by John Conington (1876) and Georgius Thilo (1881, in Latin). The spirit of Perseus's non-English collection is similar to the early goal of my project, wherein I would provide a source text and its translation in a single view for comparison. In this regard, my original goal would have largely replicated the same endeavor as Perseus, albeit with a different English translation.

Figure 6: A sample view offered by the Perseus Digital Library

The above image shows a Latin stanza in focus with two English translations to the right side.

Though not all features are visible in this view, other facets of the Perseus interface include notes

by previous editors, linguistic analyses, references, and other contextual content that might

belong in any scholarly edition.

http://www.perseus.tufts.edu/hopper/

In addition to Perseus, the Open Greek and Latin Project housed at the University of

Leipzig aims "to represent every source text produced in Classical Greek or Latin from antiquity

through the present, including texts preserved in manuscript tradition as well as on inscriptions,

papyri, ostraca and other written artifacts" ("Open"). While the repository is not static and is continually expanding, collections include original Greek and Latin texts by classical authors such as Catullus, Ovid, Plato, Sophocles, and Xenophon, as well as translations in various languages of many texts. Unlike Perseus or other mainstays of the digital edition genre, this project's main goal is one of preservation—certainly worthy, but notably lacking in interpretive endeavor beyond the structural coding. In fact, the collection features "A collection of machine-corrected XML versions of classical authors and works, freely available to download and reuse." Some of these texts include facing-page translations, which mirrors the initial goal of my own project that was changed in favor of providing facsimile and interpreted versions of the same text. The most fundamental structural coding has been carried out at the Open Greek and Latin Project, and very little more. This apparent lack, however, offers exciting opportunities for future coding, analysis, and comparison. Thanks to this source of texts, a vast amount of TEI-compliant classical texts is readily available for scholarly interchange and is under constant augmentation.

Similarly, the Thesaurus Linguae Graecae (TLG) at the University of California-Irvine exemplifies an extensive collection of digitized Ancient Greek texts ranging from the earliest extant versions of Homer to the fall of Constantinople and the Eastern Roman Empire in 1453, especially notable for being an early adopter of using digital technologies in the preservation of texts since 1972 ("TLG"). Under continuous development since its first planning meeting, the project has seen the rise of both home computing and the Internet, and it has embraced generations of change in electronic media, from magnetic tapes, to CD-ROMs, to its present online system. Unlike the Perseus and Open Greek and Latin projects, though, the TLG is only available through subscription.

Given the incredibly diverse and voluminous amount of source texts these digital classics projects contain, the scope of mine compared to them differs significantly. Even so, they all share some goals and features that bind them together despite their apparent incongruity. For one, they bring the features of ancient texts into a modern medium. Further, they offer multiple options for viewing the same source text based on user interaction. Finally, they combine multiple interpretations of an ancient text into a new product that advances a new analysis of long-studied texts. My digital version of Beresford's text, on the other hand, does not focus on comparing translations or providing historical editorial notes; instead, the focus is squarely on a single translator's implementation of English verse conventions in his translation of a classical source text—a much narrower focus than the *en-masse* gathering approach of the Open Greek and Latin Project or the Thesaurus Linguae Graecae. Also unlike these others, my own project very explicitly highlights the interpretive markup I have employed; that is, while I hope to offer a facsimile reproduction, I also use digital markup to call attention to streams of the text to which I have assigned my own interpretations.

These projects certainly do not constitute the entirety of what the digital classics offer, though they may be the most visible. Given the decreasing barriers to powerful computing technology, as well as increasing digital literacy, small projects have more opportunity to flourish. Open-source projects like Perseus, the Open Greek and Latin Project, and my own Beresford translation are critical in spreading knowledge and technological literacy so that other projects of this type might burgeon. For future expansion and interchange, the documentation of markup choices is essential, especially for a project operating with a custom-defined vocabulary like my own and the several listed before. For this project in particular, my custom schema articulates which choices I have made, and this essay expands upon it. Should I or another

scholar wish to work from its current point in future years, it should be possible to continue with its current constraints without the need for trial and error. However, a balance is necessary. While "Imposing markup on texts can create long-term resources that will benefit both students and scholars," and while it is possible to code to a near-infinite degree of fineness in any text, "Poorly done markup, however, consumes valuable resources, results in seldom used materials, and creates suspicion among scholars about the benefits of using markup" (Durusau 309). It is necessary, therefore, to have a goal in mind when deciding on a markup vocabulary and approach.

Durusau's concerns are certainly valid. Digitization is not intended to be a cure-all: besides the hurdles of learning a new digital technology, scholars have to consider longevity in digital environments. Just as paper manuscripts can suffer the physical wears of time and the environment, digital files are subject to bitrot, hardware failure, and obsolescence. That is, the *medium* is more of a problem—which is alleviated, but not solved, by redundancy—than a UTF-8 encoded text. In this regard, Syd Bauman claims that open formats like TeX or XML satisfy the need to avoid proprietary formats as mentioned above—an important consideration, especially given the short spans of time that commonly make proprietary formats unusable ("Interchange vs. Interoperability"). Bitrot and hardware failure can at least be mitigated—though not entirely conquered—by redundancy and file checking. Emerging hard drive formats attempt to negate the problems of bitrot by enabling file consistency checking, but this technology, as of now, is not largely deployed. Until it becomes more common, redundancy in multiple remote locations serves as the best means of digital longevity assurance.

A Discussion of the Associated Project Files

The project is composed of a system of several interdependent files. While the Beresford edition serves as a single-source document, the associated files provide both context and specific representation rules for said document.

- thesis.xml: this document, providing exigence and discussion of the project as a whole;

- beresford.xml: the XML source of the edition itself—the encoded text of James Beresford's translation of the Aeneid, which is coded for rendering by several content management systems;

- aeneis.xml: the Latin text of the Aeneid, sourced from the *Latin Library*—though my project has since strayed from the goal of providing a side-by-side comparison of the Latin and English translation, this file has been coded as a TEI document;

- main.xml: the composite TEI corpus of all project texts, including this thesis, the Latin text 'aeneis.xml,' and the translation 'beresford.xml';

- encodingDesc.xml: the shared resource for header information for all project files—this file includes definitions for rendering and font formatting rules, as well as a TEI definition for the metrical pattern of 'beresford.xml';

- WFDSchema.odd: the "One Document Does it All" (ODD) file defining what I have determined to be relevant elements and attributes for inclusion—each figure in particular is discussed in detail below in Chapter Six; and

- WFDSchema.rng: the RELAX NG schema automatically generated from WFDSchema.odd for use with programs like oXygen and Eclipse—this file defines rules

for XML editing programs to assist with my XML coding by ensuring consistency according to those custom rules.

The platforms used to render the project—both the California Digital Library's XTF and John Walsh's TEI Boilerplate—require a certain degree of customization before providing any useful representation of the Beresford text. Each platform possesses a specific set of files that I have customized to fit my needs, the most crucial of which are discussed in more detail in Chapter Six.

Educating Students and Scholars about Markup

Realizing the possibilities enabled by the merger of literary analysis and digital tools creates a new method of poetic interpretation. This project can serve as an example pedagogical tool for interpretive poetry analysis: using it—or a similar approach to another text—as an example can help students and scholars visualize the potential of digital humanities in a curriculum. The approach taken by this project could further serve a variety of audiences, including students and scholars who wish to visualize the structure of poetic devices, rhythm, and meter ascribed to a poem by its author. Here, the very focus is the manipulation of the form *around* the content in order to make specific parts of the content stand out. When the figures themselves stand out in colors that are separate from the words of the text, it becomes possible to notice patterns and frequencies that might otherwise be far more obscure or subtle while reading the original or the facsimile representation of the text.

CHAPTER FIVE: AN ORIGINAL: THE LEGACY AND ONGOING DEFORMANCE OF THE AENEID

Performing and Representing the Aeneid

The *Aeneid* itself has been a classic for two millennia, with countless translations, iterations, re-imaginations, and critical editions in every medium since the time of its inception. Disagreement abounds regarding the goal of Virgil in the creation of his epic, whether he intended to glorify the order and might of the newly founded Roman Empire, or whether he meant to subvert the new regime. Such dissonance has existed long enough for adherents of the former view to be called "optimists" and the latter "pessimists" (Perkell 14–15). While not falling in either camp, this new digital edition is simply another link in that long deformative chain, yet it still brings something new. In his own argument for his edition, James Beresford critically engages his contemporaries and predecessors by asserting that their impetus may have been worthy, but so many efforts were misguided by translators who attempted to force something of themselves into the text, making it less genuine to the original meaning. In particular he criticizes the use of rhyme where it does not exist, as by attempting to force English rhyming conventions onto a text and language that rarely employed rhyme, one necessarily twists the text into a form that takes on meanings inherent to the idioms of the translating language (Beresford viii). Even as he argues this, though, Beresford employs the textual features common to English poetry of his day, including an iambic pentameter rhythm and meter as well as numerous poetic figures.

In the same volume as Christine Perkell above, herself a classicist whose research focuses on epic poetry, come William Anderson and Gary Miles, each of whom have dedicated much time and effort to classical studies, and who each offer insight into the *Aeneid* and its context. In

particular, Anderson provides a history of the epic as it has been translated and rendered in

English, while Miles analyzes the *Aeneid* as a foundation story of Rome. Consequently, as a

foundation story, the *Aeneid* is quintessential in describing the origins of a particular culture

(Miles 231); no other work of Classical literature is its equal in so thoroughly defining the

cultural identity of Rome during the transition period from the late Republic to the early Empire.

It is a common assumption of the optimists that the *Aeneid* is an analogue for the time of strife

and upheaval through which Virgil himself lived, with hopes of a new, more peaceful age within

view (Perkell 5–7).

Many interpretations exist for the exigence of the *Aeneid*, and we will likely never have a

complete understanding of all of Virgil's decisions and motivations. For the purposes of this

project, the broad meaning, the grand interpretation, while always important, is largely irrelevant

to the textual encoding of rhetorical figures. It may be that my coding of Beresford's translation

enables or assists others in their own interpretations, but it is not my intention here to engage in

the *why* of Virgil's epic, and whether he was pro- or anti-Imperialism, a regretful poet longing

for the heyday of the Republic and coerced into propaganda, or a staunch Imperialist touting the

fledgling glory of the Roman Empire and its first emperor, Augustus.

Regardless of Virgil's cause or sentiments, his work comes to us in an uncertain

condition. That is, there exists no canonical version of the Latin text; it is also a part of the text's

legend that Virgil himself wanted the manuscript burned upon his death. Thankfully, his request

was never fulfilled, but we are nevertheless left with a work of poetry in an indeterminate state

(Perkell 5).

Even the "source" Latin I initially acquired for this project, at the freely available Latin

Library website, is acknowledged to be "drawn from different sources," producing a bricolage

text that has been cobbled together from different editions and even time periods. As Perkell argues, "Although most scholars today do believe that we have the ending V[i]rgil intended, the persistence of questions points to the problematic nature of the poem's abrupt and, in some respects, unresolved closure" (5). That is, the *Aeneid* was never truly a complete work. What we have comes to us through the mouths and hands of various editors and scribes who have finished half-completed lines and perhaps even corrected some discrepancies in the original Latin.

James Beresford's translation is by no means the earliest or even the most famous. The first line-by-line translation of the *Aeneid* into English comes to us from Gavin Douglas, a Scottish bishop, politician, and poet whose 1512–1513 enterprise remained largely unknown during his life. The first British English translation to gain popular success—started by Thomas Phaer and finished by Thomas Twyne—came in 1573, where it inspired a long line of other attempts in other meters and with varying fidelity to the original (Anderson 285–286). Arguably, it is this line of inspiration that led to Beresford's own version as well as the enduring "greats" of previous centuries.

The most successful and influential of these translations, even through the present, is inarguably John Dryden's version of 1697, a translation that employs rhyming couplets, which "was held to be the proper measure for heroic poetry" (Brower 271) at the time. Dryden's success also persisted because, according to William Anderson, "It took the tense times of World War II in Britain to concentrate attention once again on the *Aeneid* and its special kind of unromantic heroism as a parallel for the savagery and stress of modern warfare" (286). Therefore, it was not until the mid-twentieth century, with a general ennui surrounding the long-used rhyming heroic couplet and a new scholastic focus on classical civilizations, that new translations could vie for the influence that Dryden's translation enjoyed for more than 200 years

(Anderson 287). With a renewed interest in Classics, an explosion of translations flourished in the latter half of the twentieth century, with a loosening of the restrictions of rhythm and meter. In his writing during the last year of the twentieth century, William Anderson is convinced that "it is still too early to name the next major translation, but there are writers and publishers who compete for that title right now" (287). In terms of twentieth-century translations, there is undoubtedly no small number from which to choose that might compete for the most influential. These names include, but are certainly not limited to, H. R. Fairclough (1916), Rolfe Humphries (1951), C. Day Lewis (1952), W. F. Jackson Knight (1956), Allen Mandelbaum (1971), Robert Fitzgerald (1983), C. H. Sisson (1986), and David West (1991). This list includes translations in both prose and verse, each with varying approaches to fidelity in meter and meaning.

Within the present century, Robert Fagles (2006) and Sarah Ruden (2008) stand out as examples, both of whose translations offer renditions in Modern English while still attempting to preserve some form of metrical and idiomatic fidelity. In the postscript to his 2006 translation, Fagles claims to "have tried to find a middle ground . . . between the features of an ancient author and the expectations of a contemporary reader" by balancing the "literal" and the "literary" (388). As a result, Fagles has opted to avoid a strict rhythm and meter likely to be found in Dryden's and his contemporaries' translations for the trade-off of looser translation. That is, by not restricting himself to five iambic feet in every line as Dryden, Beresford, and others have done, Fagles gives himself the verbal space to translate Latin idioms that comprise one or two words into corresponding English that might require considerably more words, but nevertheless maintain what Fagles feels is a true-to-spirit representation of the Latin meaning. Perhaps most importantly in Fagles's edition is his explicit acknowledgment of a slew of more recent translations in both prose and verse for which he gives thanks (including many of the ones

listed above), but also meaningfully argues that every edition of the *Aeneid*—including all of its

iterations and translations—influences the others: "So there may always be room for one

translation more," he says, "especially as idioms and eras change" (401).

Unlike Fagles's intentional balance of the "literal" and the "literary," Sarah Ruden's 2008

edition is notable for its intentional "line-by-line, metrical translation" (Ruden xi)—that is, the

poem is translated into an English rhythm and meter using the same number of lines as the Latin

from which Ruden draws. Ruden, like the earlier English translators, attempts to abide by a

rhythm and meter because, as she admits, by using prose instead of verse she "would have been

tempted, in the typical way of a modern thinker, to squeeze out all the plausible meaning of

V[i]rgil's juicy lines, dilute them with [her] own explanatory verbiage, and offer a watery drink

without much actual flavor of the *Aeneid* left" (x). While she employs a stricter scansion than

Fagles, Ruden's translation by no means adheres to rigorous iambic rhythm, and frequently

strays from using ten syllables per line—all this despite her claim to have chosen iambic

pentameter (ix). Ruden, then, seems to believe that fidelity is best achieved by maintaining both

verse and meaning, and hopes to limit the injection of meanings and idioms from her target

language (English iambic pentameter) that could influence and twist the spirit of the original text

(Latin dactylic hexameter). Her crucial argument rests in her claim that "correspondence in form

is vital to authentic rendering, especially of Greek and Roman poetry." In the end, though, she

expresses a similar dissatisfaction with her own work as Virgil did with his when he purportedly

"clamored for his manuscript to be burned" (xi); "We cannot match in reality," she says

regretfully, "our vision of what we need to create from our minds" (xii). Though my own edition

is not a translation, I have found myself in frequent agreement with Ruden's sentiments,

especially of the discrepancy often found between the idea and the reality—the "Muse" and the

representation. Ruden's opinions also echo the choices and arguments James Beresford makes of

his own translation, addressed below.

Finally, it took these more recent textual explorations even to contemplate the possibility

that the *Aeneid*, or any work for that matter, is able to exist as more than any individual artifact at

a specific point in time. As poetry critic and classical scholar Reuben Brower asserts in a 1972

article, "[T]here are only texts for interpretation, whether the text is written or oral, a piece of

behavior - a dance or a cockfight - a drawing or painting, a sculptured stone, or a terracotta pot"

(270). Such a definition of a text—that is, any product of human expression that can be

interpreted—attempts to break down the division between disciplines, between those who

critique paint on canvas and those who contemplate words on paper. In practicing what he

argues, Brower himself seeks to explore and compare two verbal texts with a painting,

demonstrating that what constitutes a "text" is more than its traditional expectations.

In the preface to his translation, James Beresford justifies the choices he makes in

balancing the spirit of the work with fidelity to the author's original language. Being poetic

pieces, every choice—both of the author and the translator—are specific in order to fit the

rhythm, meter, and context. As he asserts, with some chagrin, "It seems to have been wholly

overlooked by the bulk of our Translators, that the great principle which should actuate them, is a

wish to extend and perpetuate their Author's renown; and this, as must be evident, can only be

effected by the closest imitation which it is in their power to produce" (Beresford vii). Further,

he continues, "Whoever, then, considers himself as free, at one time, to force upon his Author

thoughts and words in which he has no property; at another to rob him of those which are his

own, is unqualified for the trust he has undertaken, since he is not, what it is alone his business to

be, a faithful Representer" (Beresford vii). That is, to Beresford, fidelity to the author's original

meaning is of utmost importance so that the translator neither injects their own thoughts upon the work being translated nor takes anything away. In his postscript, he maintains this argument, believing that the ideal translation is one that offers the greatest fidelity, which, while he states admiration for other translations like Dryden's, he clearly does not believe his predecessors have adequately accomplished (xix-xx). Such desire for faithful representation is echoed repeatedly, even through the present—as discussed in regard to Ruden's translation—but also notably in the context of digital representation by Kevin Kiernan, who advises the use of facsimiles to assist in reliable and faithful editing (262), hearkening again to Smith's argument for digital editions.

It seems with these assertions that Beresford hopes to represent as faithfully as possible Virgil's thoughts and to avoid the "deforming" part of deformance. However, his work is still an edition of an earlier work; moreover, it is a translation that creates an entirely new perspective on the work. While Beresford is hopeful that he has largely succeeded in not imposing "thoughts and words in which he has no property," Jerome McGann's notion of deformance is unavoidable, and would suggest that no translation or other interpretation is possible without obvious transformation—no matter how faithful it is to the previous language.

McGann is equally adamant in asserting that "there is no such thing as an unmarked text" (McGann 138), echoing a common topic of discourse in digital humanities, especially markup theory. As mentioned above, Text Encoding Initiative *Guidelines* provide a brief historical synopsis of markup, claiming that "the word *markup* has been used to describe annotation or other marks within a text intended to instruct a compositor or typist how a particular passage should be printed or laid out." However, "As the formatting and printing of texts was automated, the term was extended to cover all sorts of special codes inserted into electronic texts to govern formatting, printing, or other processing . . . Generalizing from that sense, we define markup, or

(synonymously) encoding, as any means of making explicit an interpretation of a text. *Of course, all printed texts are implicitly encoded (or marked up) in this sense*" (Burnard and Bauman xxvii, emphasis mine).

As evident from the emphasized statement from the TEI *Guidelines*, McGann is not alone in believing that all texts are marked up: seemingly simple features of a text such as punctuation, line breaks, and the white space between words are all forms of markup that help humans read, or interpret, a text (xxvii). Even "scriptio continua," the style of writing that lacks white space and punctuation between words, bears markup in the form of line breaks and enlarged letters that serve as locational delimiters (for example, the start of a text). In short, according to such arguments, there can be no non-deformed text: if a text is marked up—which all texts are—it exhibits the results of interpretation.

In a co-authored article with Dino Buzzetti, McGann goes into further depth about the nature of markup-as-interpretation. In agreement with the TEI *Guidelines* and McGann's earlier claims in *Radiant Textuality*, he and Buzzetti assert, "Any explicit feature of a text can be conceived as a mark" (Buzzetti and McGann 60), and, "It is through markup that textual structures show up explicitly and become processable" to either human or computer readers (64). That is, once the text has been marked up (with punctuation, for example), syntax and meaning can be derived by a human reader based on the location of specific marks—or even lack of marks. For a computer processing the text as data, markup like structured open- and close-tags allow the text to become parseable. It is for this consideration of text-as-data that the TEI *Guidelines* "make recommendations about suitable ways of representing those features of textual resources which need to be identified explicitly in order to facilitate processing by computer programs" (xviii).

Alternatively, Buzzetti and McGann establish markup as an intermediate part of a text by arguing that "markup can be viewed as a sort of diacritical mark" that simultaneously is part of the text and describes it (64). Markup creates a connection between the text and its "interpreters," be they authors, editors, or readers. Each interpreter becomes a part of McGann's notion of deformance; editors and authors specifically "reconstruct a complex documentary record of textual makings and remakings, in which their own scholarly work directly participates" (Buzzetti and McGann 71). That is, there is no isolated text: besides being part of a greater conversation, editing and creating content in the digital medium allows quick duplication and distribution that is "beyond recall by the editor" (Berrie et al. 269). Phill Berrie et al. continue, "The arduous business of entering, proofreading, amending, and consequently reproofing a transcription containing the new interpretation can seemingly be avoided in the electronic medium, but in fact a new state of the text will have been created," thus enabling the introduction of new errors that might not have existed in a previous edition (270). Errors, then, become part of the scholarly discourse and must be addressed in future discussions.

In particular regard to the needs of classicists, Greg Crane argues that their approach towards the Classics has largely been the same for millennia in the development of lexica, encyclopedias, commentaries, and editions; even in the past, they "established standard, persistent citations schemes for most major authors." It is no surprise then that "[t]he adoption of electronic methods thus reflects a very old impulse within the field of classics" (46). Most importantly, according to Crane, electronic media are not necessarily to be embraced because they are new, "but because they allow us to pursue more effectively intellectual avenues than had been feasible with paper" (46). Moreover, due to the increasing availability of general digital tools, classicists "are better able now than ever before to share infrastructure with our colleagues

not only in the humanities but in the rest of the academy as well" (Crane 55). In the case of this project, Crane's sentiment is particularly pertinent as I have employed customized versions of very general digital humanities tools like XTF and TEI Boilerplate.

A (Con)Textual History: The Aeneid *in 1794*

In 1794, at the age of 30, James Beresford successfully published his first work, *The Æneid of Virgil*, an English translation of the original into blank verse. According to his 1841 obituary in *The Gentleman's Magazine*, it seems Beresford garnered little fame for this work; instead, it was for his 1806 satire entitled *The Miseries of Human Life* that Beresford earned the most attention, as well as "an abundance of imitators" (548). Including his *Miseries* and the *Aeneid*, at least 13 independent works are attributed to Beresford, as well as several contributions to periodical publications. Though certainly not as significant to his success as an author as his *Miseries,* Beresford's translation is his first foray into publishing and represents a late eighteenth-century approach to classical translation.

The physical artifact that serves as my source text is in a precarious state, which functions partially as an impetus for this project. Paper is fleeting; it is my hope that the digital form will allow the physical work to live on. Beyond the desire for preservation, however, is a desire for analysis: Beresford's incipient entry into publication provides a window into British literature of the late 1700s, and delivers an example of neoclassicism and poetic formalism. These latter categories coincide with my own interests, and I believe that creating a digital textual analysis of the work can produce valuable insight via little-used methods. Perhaps at least as important as the result is my desire to validate the approach.

While the physical object I digitized suffers from external degradation, the pages are legible and intact.

THE

ÆNEID OF VIRGIL,

TRANSLATED INTO BLANK VERSE

BY

JAMES BERESFORD,

FELLOW OF MERTON COLLEGE, OXFORD.

———————

——CASUM INSONTIS MECUM INDIGNABAR AMICI,

NEC TACUI—— DEMENS!   VIRGIL.

———————

LONDON:

PRINTED FOR J. JOHNSON, ST. PAUL'S CHURCH YARD.

1794.
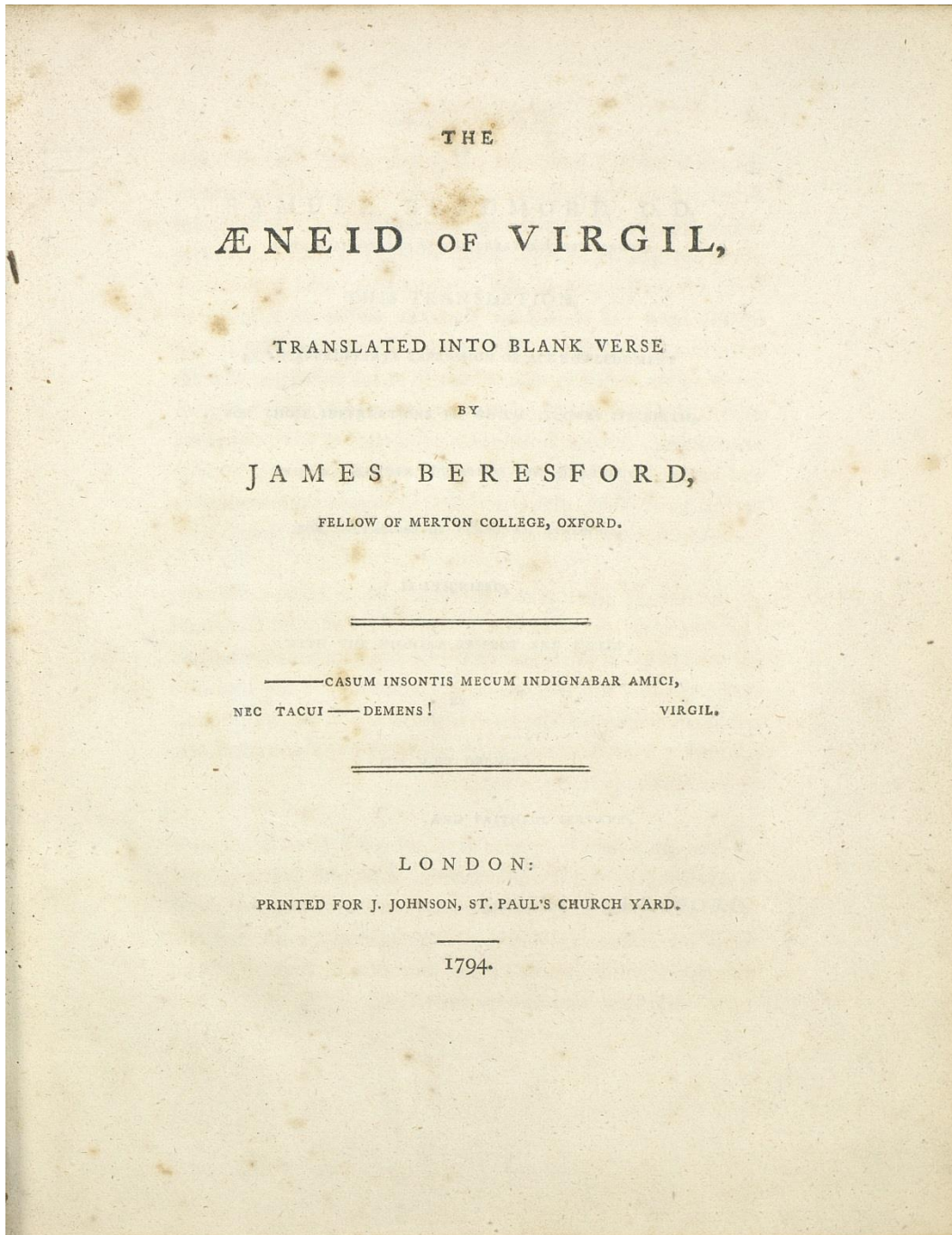
Figure 7: The title page of the Beresford translation

The title page of the source text elucidates the title, subtitle, author, epigraph, and imprint

information, including the publisher and date of publication.

Æ N E I D.

B O O K I.

Arms and the Man I sing, who first, Fate-driv'n
From Ilium's coast, thence to Italia came,
And the Lavinian shore: wide-toss'd was he
By stress of Pow'rs above on flood and field,
Through fell Saturnia's memorable rage;
Nor less, beside, in battle much endur'd,
Until he rais'd his city, and his Gods
In Latium lodg'd: whence rose the Latin race,
And Alban Sires, and walls of Rome sublime.

   O Muse! the springs reveal; what Pow'r profan'd—
Or brooding what offence, could Heav'n's high Queen
Compel the man for piety renown'd
To roam, the sport of many a hard mischance,
And many a sorrow meet, and many a toil?
In heav'nly bosoms burns such quenchless ire?

   An ancient town there stood, by Tyrians held,
'Gainst Tyber's mouth and the Ausonian shore
In counterview, far distant; rich in gold,

B

Figure 8: The first page of the first book of the Beresford translation

Visible in the above images are an abundance of features common to the textual artifact, among

which are multiple levels of headings; the common indentation of stanzas subsequent to the first

of each book; small-capital font formatting on the epigraph and the "rms" of the first word, "Arms"; the long-s character; the Æ ligature; the gathering symbol at the footer of the page (in this case "B"); and the typesetting used throughout the work. Also visible in these examples is the overall condition of the surviving artifact, which is very good considering its age of more than 200 years. Besides the detached covers and the coloration of the pages, very little about the book's state has suffered.

In his prefatory content, Beresford cites or alludes to "numerous predecessors" (viii) and other "principal Translators" (v), including particularly John Dryden and his eminent 1697 edition. Beresford is under no illusion that his own will surpass the fame and success of that edition, but he makes worthy claims supporting his own work and his editorial choices, including to forego the rhyming couplet form that Dryden employed.

Also in his preface and his postscript, Beresford discusses the decisions and challenges he encountered in his translation of the Latin source to English. In the former, he largely addresses the guiding philosophies of his work and engages in copious self-deprecation; most importantly, he argues that an unfortunate penchant of his contemporaries and predecessors is a desire to improve their own esteem rather than "to extend and perpetuate their Author's renown" (Beresford vii). Such renown, he continues, can only be achieved by "the closest imitation which it is in their power to produce." The translator who seeks "to force upon his Author thoughts and words in which he has no property" is unqualified to be "a faithful Representer" (vii). Such condemnation is telling both of Beresford's opinion of others' approaches as well as of his own: he intends to produce an accurate rendition of Virgil's work, even as he chastises himself, begging that any perceived "objections" the reader may have of the translation are not "weighty enough to crush [him]" (viii). Further, Beresford hopes that the final published work is his best

attempt at remaining faithful to the original "without change, addition, or retrenchment," while yet omitting that which would, "in the English, degenerate into feeble, ungraceful, or barbarous redundancy" (xi). Also omitted, he says, are "the obscure or disputed passages of my Author" (ix). Instead, he has made the conscious decision to emend disputed passages silently so as not to "encumber or interrupt the reader with a detail of various readings or various interpretations" (x). It is worth noting, then, that those lines that Beresford has translated have been selectively filtered from the set of all lines that *could* be attributed to Virgil. Such choices are clearly not oversights, but rather intentional and acknowledged editorial intervention.

I have considered each of the book's features listed above, and addressed them variously. Because I, like Beresford, strive for a faithful representation, I have chosen to maintain vowel ligatures and the long-s character—though this latter choice proved an interesting challenge, as discussed in detail later. Font formatting is another intentional choice of Beresford's, and I have likewise coded any variations diligently. The different "content objects," be they stanzas, lines, headings, paragraphs, or otherwise, are all coded with an eye for differentiation; that is, if it seemed to me that Beresford (or his publisher) made an intentional choice in the presentation of the text, I attempted to represent that choice in my interpretive markup.

In all, Beresford presents a humble persona, desirous of a fair judgement even as he himself proclaims that his translation is no paragon of the craft, but is instead compared to a cast of a statue, "enabling certain persons to conceive what they have no opportunity of viewing for themselves" (ix). Such strictness is reiterated in the postscript, wherein Beresford lists the principles of a good translator, asserting that the "cardinal duty" of such a role is that of fidelity (xx). Nor is Beresford shy to admit his errors; he dedicates an entire page to noting errors in his work and how they should be emended by the reader.

E R R A T A.

N. B. The reader is particularly defired to alter thofe Errata to which an Afterifk is prefixed.

Page 6, line 10, from bottom, (Eurus,) *dele* parenthefis.
    21, — 16, transfer the comma from *rear,* to *walls.*
    23, — 3, *for* lab'rers hands, *read* labourers; and between this line and the next add this:
        " And works elaborate of the Artift's hand."
    45, — 8, *for* black, *read* blank.
    47, — 12, transfer the comma from *you* to *ftart.*
    88, — 2, from bottom, *for* ancient, *read* was old.
* 102, — laft, *for* Pæan, *read* Ææan.
  130, — 15, *for* Trojans', *read* Trojan.
  141, — 9, *for* untried, *read* uneffay'd.
  154, — 8, *for* when, *read* where.
   —, — 9, *for* hafte, *read* haft.
  160, — 9, from bottom, transfer the comma from *while* to *tomb.*
  161, — 1, transfer the comma from *Italia* to *explore.*
  162, — 11, *for* and, *read* had.
* 181, — 3, *for* eager, *read* ready bow.
  202, — 5, *for* to, *read* I.
  301, — 10, *for* their, *read* thine.
  338, — 5, from bottom, *for* fixing deep its, *read* with the fplinter'd.
  339, — 3, *for* car, *read* ear.
  344, — 9, from bottom, between *high* and *heads,* insert *thofe.*
  370, — 11, *for* attendant choir, *read* attendant train.
  419, — 9, *dele befide.*
  422, — 5, *for* heeding, *read* needing.
  430, — 16, *before* to, *read* thou.
  432, — 1, *before* by, *read* is.
* 480, — 7, from bottom, *for* furvive, *read* furviv'd.
* 493, — 2, from bottom, *dele* the whole line.

Figure 9: Beresford's "Errata" page

The errata page notes in detail the words and their locations that should be corrected by the

reader to be more faithful in translation. Some corrections seem very minor, such as moving the

placement of commas, but clearly were of enough meaning to Beresford to bear inclusion in this list of admitted mistakes. Others seem more significant, including the substitution of words or phrases or even the addition and deletion of entire lines. In my coding, I did not make these emendations, instead choosing to represent the book as it appears.

Based on Beresford's subscriber list, it seems clear that his arguments were successful. Comprised of 11 pages, this list includes all those people who contributed to the production of his translation, of whom several ordered multiple copies. Perhaps most notable within this list is the Prince of Wales, George IV, who became the King of England in 1820.

LIST OF SUBSCRIBERS.

His ROYAL HIGHNESS THE PRINCE OF WALES.

A.

REV. George Anguiſh.
Rev. J. R. Arding, Briſtol.   2 Copies.
Mr. Auber, Hoddeſdon.
Mr. Peter Auber, Hammet Street.
Rev. John Auber, M. A. Aldenham.
Mrs. John Auber.
Mrs. Judith Auber, Cheſhunt.
Mrs. Elizabeth Auber, Cheſhunt.
Miſs Catharine Auber, Cheſhunt.
Mr. Iſaac Auber, Croydon.
Mr. Stephen Auber, Elder Street.
William Alderſey, Eſq. Wimpole Street.
Thomas Audley, Eſq. Lynn. 2 Copies.
A. Bath.
Anonymous.   5 Copies.
Anonymous, by Lieutenant Colonel Alexander Mac Kenzie, 78th Regiment.
———— by the ſame.
———— by the ſame.
———— by the ſame.
———— by the ſame.
———— by the ſame.
———— by the ſame.

Anonymous, by the ſame.
———— by the ſame.
Thomas Audley, Eſq.

B.

Scrope Berdmore, D. D. Warden of Merton College, Oxford.
Samuel Berdmore, D. D. late head Maſter of the Charterhouſe School.
Mr. Samuel Berdmore, Jeſus College, Cambridge.
—— Burton, D. D. Canon of Chriſt Church, Oxford.
—— Bathurſt, D. D. Canon of Chriſt Church, Oxford.
Rev. Charles Bathurſt, Fellow of New College, Oxford.
Rev. George Baxter, Brother of St. Catharine's, Muſcovy Court, Tower Hill.
Rev. Robert Baxter, Fellow of Jeſus College, Oxford.
George Trotman Baxter, LL. B. Trinity College, Cambridge.
Rev. Charles Barton, St. Andrew's, Holborn.

d

Figure 10: The first page of Beresford's list of subscribers

Several individuals commissioned multiple copies; featured prominently at the top of the list is

"His Royal Highness the Prince of Wales."

Knowing Beresford's context and guiding principles in his own words is helpful to my own digital edition of his translation, especially in regards to his intentional lack of attempt at maintaining a rhyme scheme; this piece of knowledge directly influenced my approach at coding rhyme, discussed in the next chapter. His editorial decisions—namely, to avoid forced rhyme, to interpret Virgil's disputed passages silently, and to hold as true as possible to the original text while still bending the English language around Latin idioms—is information that not every editor of a scholarly edition is fortunate enough to have access to.

CHAPTER SIX: THE PRESENT ADVENTURE: THE PROJECT'S INFLUENCE, GOALS, AND PROCEDURE

"Editing scholarly editions of poetry has never been an enterprise for the faint of heart"
(Freistat and Jones 105)

The project sits at the intersection of the paths of several disciplines, and would not be possible without exactly the right amount of indecision in my own past. Computers had always been central to my life, but it was early in my undergraduate career that I decided that my skills were ill-suited to computer engineering and software development. It was thus that the interdisciplinary calling of digital humanities came to me as a particular windfall. My personal interests in the Classics, structured poetry, and knowledge production and preservation all finally found a mutual home. This project attempts to combine these seemingly loosely related topics into one cohesive textual artifact by undertaking a from-scratch facsimile reproduction of the 1794 blank-verse rendition of the *Aeneid* by James Beresford and by comparing my own poetical analysis and facsimile reproduction of the 1794 Beresford version with the possibilities offered by other digital archive projects.

Inspired primarily by David Birnbaum's "Paul the Not-So-Simple" project, I initially intended to produce a dynamic, side-by-side comparison of Beresford's work and the original Latin with which users could interact, such as by clicking relevant "hotspots"—for example, stanza markers—in one text that would jump to the corresponding location in the other text. This approach would enable users to compare the same area of multiple versions or translations of a work, referred to in Freistat and Jones's "The Poem and the Network" as a "dynamic collation" (115). This format is intimately familiar to scholars of foreign languages, wherein parallel translations often serve as a means of presenting the source text and its line-by-line translation on the facing page.

On the other hand, Fraistat and Jones echo the sentiment expressed in the epigraph above—as well as the cautionary sentiments of Durusau and Lavagnino—by stating, "Poetry as a form . . . tests and sometimes strains the resources of any textual encoding system." Further, and equally as important, they claim, "Editors of poetry in the electronic medium need to possess a clear imagination of the front end, the interface or desired physical display of the text, when going into an editing project, in order efficiently to plan the logical, structural markup—and vice versa" (106). While I had experience with XTF before beginning this project, I had not undertaken a digital parallel, line-by-line translation, nor did I fully consider the relatively few benefits compared to the relatively many hurdles requisite of such an endeavor.

When I began the procedure of digitizing Beresford's work, I was as yet unaware that the work would be released as part of the Eighteenth Century Collections Online (ECCO) project. However, said digitization engages in no interpretive discourse, providing merely a structural representation. Moreover, I believe my own transcription fostered a closeness with the text that would have been missing with someone else's work. Further, it had been my intention to take advantage of optical character recognition (OCR) software to perform the majority of the transcription, but the software to which I had access (Nuance Communications' OmniPage Pro) proved wholly incapable of interpreting many of the features of the text, especially the differences between the long-s character "ſ" and other characters with ascenders such as "f," "t," and "l." As a result, hand-coding still has its place in a situation wherein high-end OCR software is not feasible to purchase by a single person working on a single book. Even with such software, proofing of the output would be desired, and keying and coding by hand enables greater familiarity with the text than automatic generation.

A crucial point of consideration in textual editing is choosing what is of interest in a work. While virtually everything can be represented, an editor must decide what is *of interest or use* to represent. As I began hand-keying the text, I initially decided I would encode every instance of long-s as it appeared. It did not take long, however, for me to realize I had dictated far more work than necessary for myself, and created only more difficulty in text-search functionality. Instead, I modernized every instance of long-s and instead created XSL rules to convert standard "s" into "ſ" where appropriate. This approach allowed for faster typing in addition to alternate renditions without any significant additional work or coding, though it must also be acknowledged that any variations in the original cannot be represented by this approach.

Besides typographical considerations, it is important to question which textual features are worth singling out. For example, is enjambement a poetic feature worth marking up with the @enjamb attribute on each <l> element where it occurs? I cannot expect to identify every poetic device present in Beresford's translation of a twelve-book epic poem, nor even to mark up every example of those devices I *do* select. Such a task is virtually impossible for a single coder in such a timeframe as this project. Indeed, nearly everything can be represented with indicative markup like XML, but there is a point at which the work outweighs the benefit. As Berrie et al. argue, by adding interpretative markup like TEI-conformant XML, "a new state of the text will have been created" (270), thus creating the possibility of introducing new errors that were not in the previous version. Buzzetti and McGann echo similar concerns, arguing that scholars who edit texts (and indeed, anyone who marks up any kind of text) "reconstruct a complex documentary record of textual makings and remakings, in which their own scholarly work directly participates" (71). In short, while new interpretations may offer new value to scholarly discussions, employing greater amounts of markup also creates more potential for error.

At a more fundamental level even than the XML markup is the content and rendering of the original text. My own opinion in the creation of the edition follows the philosophy of Beresford—while I created a version of the text that contains textual analysis, it is my hope that a facsimile version, true to the original, is accessible to users. In order to preserve fidelity to the original Beresford book, to what extent would I have to encode font and typography? As Wesley Raabe investigates in his own dissertation, "every part of the text [is] potentially significant," to the extent that he dedicates careful, intentional thought to "ligatures, end-of-line hyphenation, italic or roman punctuation marks, and apostrophes" (207). I encountered each of these considerations in my own transcription and encoding, and ultimately had to make editorial decisions based on each.

A "Traditional" Poetical Analysis

Since Beresford's text is a translation of a work of classical literature, there is a considerable overlap of the figures inherent to the original and figures injected by the translator—intentionally or otherwise. Some of the figures, especially synchysis and chiasmus, are more relevant to the context of Latin and Ancient Greek poetry than Modern and Early Modern English, but nevertheless I believe are worthy of indication in Beresford's translation given their noticeable inclusion as well as their creation of a connection that spans the 17 centuries between the works.

As I initially read Beresford's translation, I took note of certain recurring poetic devices and made a list of those most frequent and most interesting to me. Admittedly, this selection process was arbitrary to some degree since there exist far more figures and devices than I have indicated. Keeping in mind the warnings of Berrie, McGann, and others, though, I had to cull a reasonable amount from the larger set to avoid both insurmountable amounts of work as well as

unruly schema definitions and a tangled nest of resultant markup. In addition to selecting figures most interesting to me, I attempted to reach a consensus among several established sources, including William Harmon's *Handbook to Literature* and M. H. Abrams and Geoffrey Harpham's *Glossary of Literary Terms*. Rather than relying on a single author's definition, I have instead attempted to contribute my own understanding of each selected figure, synthesizing an agreement from my own knowledge and practice alongside the definitions of established authors.

While traditional poetic analyses rely on the dichotomy between "schemes" and "tropes," such an approach seems too broad for this project, and is less fitting for the classical context of the source text. In the end, I established three overarching categories for the various figures: those of sound, those of speech, and those of thought. My figures of sound and speech might equate to "schemes" while my figures of thought are analogous to "tropes."

For all items of interest, it is important to note that I do not claim to have identified every single instance—*especially* in the cases of lines' real meters—nor is it likely that another reader would identify the same figures in the same places I have. Indeed, this is a hallmark of interpretation: each reading reflects a unique hermeneutic endeavor.

Figures of Sound

Figures of sound include those manipulations of words that change how a syllable, word, or phrase is pronounced; such changes might be omissions or additions of syllables, or the repetition of certain sounds in proximity. Most frequently, figures of sound are employed to fit the poem's blank verse meter—unrhymed iambic pentameter—though the repetition category does not necessarily serve such a goal. Instead, sound repetition—including alliteration, consonance, and assonance—serves to draw attention to the phrases and lines in which they

occur, and might even hope to imitate a sound being described. As an example, the phrase "of mountain magnitude" in Book 2, line 21, exhibits alliteration of the "m" sound (and, additionally, is a good example of hyperbole), drawing the reader's attention to the immense size of the Trojan Horse.

Rhyme occurs scarcely, and where it does, it is likely unintentional. Regarding it, Beresford explicitly states in his preface (quoting John Dryden before him) that "what it adds to sweetness, it takes from sense, and he who loses least by it may be called a gainer" (viii). Thus rhyme is a secondary concern for him, as his primary motive is to preserve the underlying meaning of Virgil's epic. Based on his stated opinion, it is likely that the few instances where rhyme occurs are happenstance. Nevertheless, where it occurs, it has been indicated in the edition.

Included in my particular figures of interest in the sound category are alliteration, assonance, consonance, elision (also known as "synalepha"), hypermonosyllable, metaplasm, and sibilance. This category is by far the most common and frequent of all poetic figures in the work, especially given the fact that metaplasm is evident any time a syllable is added to or left out of a line in order to fit the ten-syllable meter. This occurrence is so frequent that no page exists in the edition on which metaplasm does not occur.

The figures of repeating sound, namely, those of alliteration, consonance, sibilance, and assonance, often serve to draw the reader's attention to the places in which they occur, or less frequently—especially in the case of sibilance—to mimic the sound of an action being described.

While alliteration is one of the more commonly recognized devices that occurs in abundance, even outside the realm of poetry, it bears defining to justify my encoding. Specifically, William Harmon defines it as "the repetition of initial identical consonant sounds or

any vowel sounds in successive or closely associated syllables" (13). On the other hand, Abrams and Harpham define alliteration much more openly as "the repetition of a speech sound in a sequence of nearby words . . . when the recurrent sound is made emphatic because it begins a word or a stressed syllable within a word" (12). In my mind, alliteration is most pronounced as Abrams and Harpham mention, in those words where the repeating consonant or vowel sound is on the *stressed* syllable of a foot. An example of this is evident in the case of "By stress of Pow'rs above on flood and field" (Beresford I.4), wherein "flood" and "field" alliterate not only by proximity and by possession of the same initial letter, but also by virtue of falling on the stressed syllable of their respective metrical feet. Alliteration is certainly possible and occurs frequently mid-word and on unstressed syllables, but the effect is arguably more pronounced at the beginnings of words and on stressed syllables.

In my early identification of figures, I erroneously treated consonance as what is really more subtle alliteration. For example, in the line "Appropriated praise and glory won" (Beresford V.302), I marked the phrase "Appropriated praise" as consonance of the "pr" sound. In fact, consonance is a feature for identifying rhyme. As Harmon states, consonance is "The relation between words in which the final consonants in the stressed syllables agree but the vowels that precede them differ, as 'add / read,' 'mill / ball,' and 'torn / burn'" (107). My erroneous coding of these figures has been preserved in the current edition, but future changes to the ODD and the coding will account for this.

I chose "sibilance" to stand as a specific subset of alliteration that makes use of sibilant characters and groupings such as "s," soft-"c," "z," "sh," and "ch." The sounds signified by these characters are of the same linguistic category and occasionally occur in Beresford's translation as an onomatopoeia, as in the case of "*Hisses so* fierce with *sudden swarms of snakes*" (Beresford

VII.607). In this case particularly Beresford clearly seeks to mimic the sound of the hissing snakes in the very language of the line.

Finally, other common figures of sound are those that directly affect the rhythm and meter of the poem—hypermonosyllable, elision, and metaplasm. Each of these devices shares the purpose of removing syllables from within or between words in order to fit the ten-syllable requirement of blank verse that Beresford employs.

The first, hypermonosyllable, comprises any word such as "hour," "power," or "heaven" that is typically scanned as two syllables, but is instead scanned as a single syllable in order to fit the needs of the particular line of verse. Very frequently Beresford clearly indicates when he intends such words to be read as a hypermonosyllable by using an apostrophe to omit a vowel, as in "On growing hills of water up to *Heav'n*" (III.800). However, these cases are not always so clearly marked. For instance, line 826 of Book III reads, "Groaning, and Heaven involves with wreathing fumes." Here the line cannot fit the blank verse without "Heaven" scanning as a single syllable, even though its actual occurrence as a hypermonosyllable goes unmarked.

The next meter-affecting device, elision (also called synalepha), "is most often accomplished by the omission of a final vowel preceding an initial vowel, as 'th' orient' for 'the orient'" (Harmon 167). As before, this figure is employed most frequently for the purpose of omitting a syllable to fit the meter. Consider Book I, line 39 as an example: "*Th' award* of Paris, and her charms disdain'd." Here, instead of using three syllables by writing "The award," Beresford essentially blends the consecutive "e" and "a" into a single utterance. Again, however, not every instance in Beresford's translation is as obvious as this example. The consecutive words "the infernal" in the following excerpt evince some of the difficulty inherent in identifying

every example of elision: "Here full athwart the hounds *the infernal* maid / Spread sudden madness . . ." (Beresford VII.655–656).

The final of these three figures affecting rhythm and meter is metaplasm—broadly defined by William Harmon as "a general term for almost any alteration of words or patterns" (297). Vague as this definition is, it bears refinement. In particular, "metaplasm" itself does not possess any specific defining characteristic; instead, metaplasm consists of six defined subsets:

1. prothesis: addition to the beginning of a word, as in "adown" instead of "down";

2. epenthesis: addition to the middle of a word—I did not identify any of this type of metaplasm within Beresford's translation;

3. paragoge: addition to the end of a word, as in "foughten" instead of "fought";

4. aphaeresis: removal from the beginning of a word, as in "midst" instead of "amidst";

5. syncope: removal from the middle of a word, as in "o'er" instead of "over"; and

6. apocope: removal from the end of a word, as in "oft" instead of "often"

Each of the above listed items is a specific case of metaplasm, and each acts to add or remove a syllable from any given line of poetry. There are nearly 500 lines of poetry identified with more than one instance of metaplasm occurring on a single line, with one notable example being "Was *toss'd round ev'ry* shore, and thou full *oft*" (Beresford I.894) with a total of four.

A discussion of rhythm and meter is fitting to follow figures of sound as, arguably, rhythm and meter themselves are figures of sound.

The Rhythm and Meter of Beresford's Translation

Beresford chose to use blank verse for his translation, as many of his predecessors and contemporaries were wont to do. This particular rhythm and meter is composed of lines of

unrhymed iambic pentameter, indicated in the encoding description of the TEI file in the <metDecl> element. In lieu of traditional scansion marks, stressed syllables are indicated by the capital letter "S," unstressed by "U," foot boundaries by the vertical pipe symbol "|," and line boundaries by the forward slash "/." By these conventions, a line of iambic pentameter is rendered as US|US|US|US|US/.

While iambic pentameter is the acknowledged rhythm and meter of the translation, maintaining an exact adherence to the pattern would not only be incredibly difficult, but also incredibly repetitive in the more than 13,000 lines of poetry that comprise the translation. Where variations occur, the TEI Verse module supplies a method of indicating lines that differ from the defined "normal" by means of the attribute @real, which can be attached to any individual line. Using the same metrical symbols mentioned above (S, U, |, and /), the "real" meter of each line can be indicated. While I cannot claim to have marked every deviation, I have done so where it is obvious to me that "iambic pentameter" does not fit the rhythm or meter of the line.

As discussed in the preceding section, several poetic figures can directly influence the rhythm and meter of any given line, namely, elision, metaplasm, and hypermonosyllable. These three figures serve as powerful tools for Beresford to vary the rhythm of his lines, and he employs them liberally. Though each figure of sound contributes as a direct manipulation of the poetic meter, figures of speech and thought are no less responsible pieces in the overall assembly of the poem.

Figures of Speech

Rhetorical figures of speech are perhaps most concisely defined as "wordplay": this category comprises a wide variety of devices, such as repeated words or phrases, omitted words or phrases, varied syntax patterns, interrupted words, and instances of exaggeration and

circumlocution. In this category are the figures of anaphora, apostrophe, asyndeton and polysydeton, chiasmus and synchysis, diacope, hyperbole, litotes, tmesis, and zeugma. Each of these various figures can be grouped according to function within the text.

Under the umbrella of word repetition falls anaphora, diacope, and polysyndeton. In some cases, these three even have the potential to overlap. The first is a rhetorical device defined by repetition of certain words or phrases in close proximity, especially at the beginning of several concurrent clauses. One of the earliest instances in Beresford's translation of this figure occurs in the repetition of "Who now" in the following lines: "*Who now* shall Juno's sacred name adore? / *Who now* with suppliant gifts her altars crown?" (I.67–68). In a similar vein, diacope consists of the same word or phrase repeated in close proximity with a brief interruption, usually with the intention of expressing tension. A particularly strong case of tension building through diacope occurs in the following lines: ". . . and now the flames throughout the walls / Are louder heard; and *near, and still more near*, / The storming conflagration rolls along / Its fiery torrent . . ." (Beresford II.1064–1067).

Of final note in this category of repetition is polysyndeton, "The use of more conjunctions than is normal" (Harmon 373). This figure can, like diacope, assist in building tension, though it can also serve to make a list of items or occurrences seem longer than it truly is, as in Book IX, lines 865–868: "Like that warrior old / Apollo mov'd in all, in voice, *and* hue, / *And* hoary hairs, *and* arms of dreadful clang; / *And* then the keen Iulus thus bespeaks . . ." Here each trait of the warrior Butes, whose form the god Apollo is mimicking, is preceded by the coordinating conjunction "and," thus causing each trait to receive additional scrutiny and to become part of a litany of features.

The opposite of polysyndeton, asyndeton features in the form of words omitted from the text, particularly conjunctions. As polysndeton is defined by an overabundance of conjunctions, so asyndeton features a distinct lack where they would otherwise be expected; for instance, Book II, line 872 contains a list of four items separated only by punctuation: "*Her nuptial joys, her parents, children, home*, / Shall she salute . . ." Aeneas' anger is here built by a long list of considerations of the joys Helen might still experience while his city lies in ruins and she—in his mind—is the entire reason that his friends, family, and home have been destroyed. In these considerations, asyndeton helps to augment nearly half a page of the positive traits of Helen's existence counterbalanced by the suffering of Troy and its inhabitants.

Two figures in particular comprise the category of exaggeration and periphrastic understatement: hyperbole and litotes. The first is most easily demonstrated through a description of the Trojan Horse's being "of mountain magnitude" (Beresford II.21). Hyperbole is recognizable by blatant or even impossible overstatement, also evidenced by "proud machines that vie in height with Heav'n" (IV.119), or the warrior Nisus being deemed "more swift than winds or lightning's wing" (V.421).

On the other hand, litotes seeks to bring particular attention to a phrase by negating its opposite, as is the case with falling arrows that are "*Not inexpert* of harm" (V.811). Litotes in this instance draws attention to the danger to which the gods' messenger Iris exposes herself as she casts aside her disguise on the battlefield. Another intriguing instance of litotes actually occurs as part of a larger hyperbole, demonstrating the possible intricacies of poetic figures as a whole, but especially of these two of exaggeration and understatement: "*Not less* their number than of waves that roll / In Libya's floods . . ." (VII.971–972). In this case, "not less" most likely means "at least" if not "more than." When taken as part of the larger phrase and surrounding

context, a great multitude is meant to be portrayed, and the litotes contributes even more to the overall hyperbolic description of the situation.

A more complex assortment of linguistic subtlety occurs with chiasmus and synchysis, closely related verse figures identified by word patterning. In each, a set of words follows a pattern-matching scheme whereby at least two linguistic features are each matched to other surrounding words that fill a similar role. Consider lines 664 and 665 from Book II: "**On the left** arm protective <u>they oppose</u>, / The battlements <u>they grapple</u> **with the right**." In the indicated lines, the bold pairs "On the left . . . with the right" and the underlined pairs "they oppose . . . they grapple" each share a role in the sentence that is mirrored by its corresponding match. That is, the underlined words consist of the subject and verb of a clause while the bold indicate *on which side* the verb occurs. In classical poetry, chiasmus is said to follow the pattern *abba*, whereas synchysis follows an interlocking *abab* pattern. These elaborate constructions were common in classical poetry, and Beresford clearly does his best to replicate the practice, with 22 indicated instances of chiasmus and 20 of synchysis.

Two other figures of wordplay are also evident in the text. Tmesis, Greek for "cutting," is evident wherever "a word is cut into two parts between which other verbal matter . . . is inserted" (Harmon 477). Most commonly in Beresford this feature occurs in words that end in "-soever," such as "what change soever" (X.151) and "how late soe'er" (VII.672). Also involving wordplay, zeugma is a particularly dextrous use of a verb that "has two or more objects on different levels, such as concrete and abstract" (Harmon 509), or is an expression "in which a single word stands in the same grammatical relation to two or more other words, but with an obvious shift in its significance" (Abrams and Harpham 347); such is the case of a javelin that "snatch'd his voice and life" (Beresford X.467).

Finally in the category of figures of speech, Beresford liberally employs apostrophe—"a direct and explicit address either to an absent person or to an abstract or nonhuman entity" (Abrams and Harpham 345). The second stanza of the entire poem begins with apostrophe, with an invocation to a muse (presumably Calliope, whose domain was epic poetry) on line 10 of the first Book. Though seemingly straightforward, this figure poses some unexpected challenges. Given the historical and mythological context of the story, not every address to a deity is necessarily an instance of apostrophe. Considering the fact that many of the deities feature as actors in the plot, it cannot be assumed that every direct address to them is to an entity who is not present. As such, I reserved my coding of apostrophe to those instances where the narrator himself directs his speech to a deity, or where other nonhuman entities are addressed directly, as can be seen when Priam proclaims, "I now obtest, *eternal fires of Heav'n*!" (Beresford II.226). I have indicated at least 90 instances in which Beresford uses this figure in such a manner.

Figures of Thought

A poetic figure of thought is a word or group of words that to some extent is not meant to be understood literally. Also known as "tropes," these figures involve turns of phrase that include comparison, predication, and associative substitution. In Beresford's text, I have identified four such figures, of which simile is the most frequently employed. In addition to simile, the three other figures of thought I have indicated are metaphor, metonymy, and synechdoche.

Simile and metaphor are arguably the most easily recognizable figures of thought due to their widespread use throughout literature. Both express comparisons between unlike items. In the case of simile, the comparison between objects is directly expressed through constructions using words such as "like" or "as" (Harmon 445). A clear use of simile falls on lines 334 through 336 of Book II: ". . . such his bellowing cries / *As rears the bull when from the altar's side* /

Wounded he flies abroad . . ." Metaphor, which is more subtle than simile, still accomplishes a similar goal. However, instead of directly expressing a relationship between unlike objects, renames it as a predicate or as an appositive, as in the case of ". . . stood she there / *Impenetrable flint or Parian rock*" (Beresford VI.617–618). In this example, "she" is subsequently described as possessing the qualities of "[i]mpenetrable flint" or "Parian rock."

The next two figures of thought, metonymy and synechdoche, belong to a group defined by association. In the former, metonymy, "the literal term for one thing is applied to another with which it has become closely associated because of a recurrent relation in common experience" (Abrams and Harpham 134). In "At the first breath of wind, with drowsy swell / *Ocean* awakes . . ." (Beresford VII.722–723), "Ocean" is metonymy for the Greek god Poseidon due to his association with the ocean. Similarly, in "In solid darkness thus three doubtful *suns* / We roam the seas, as many starless nights" (III.295—296). In this case, "suns" directly takes the place of the word "days" due to the association of sun with the day (as moon with the night).

Finally, synechdoche is a form of association that occurs when "a part of something is used to signify the whole" (Abrams and Harpham 135), as is the case in "*These eyes* beheld fell Neoptolemus" (Beresford II.742). "These eyes" qualify for synechdoche as they stand for the person to whom the eyes belong—in this case, Aeneas.

These final four figures are few and far between when compared to the earlier figures of sound and speech. Part of this fact is due to my more intense focus on the others, but it could also be attributed to Beresford's own stated editorial approach, whereby he attempts to remain faithful to the spirit of Virgil's text. Injecting tropes where perhaps they did not exist does not make for an editorially conservative edition.

Choosing a Markup Scheme

      As addressed by George Landow, consistency is a paramount consideration when deciding on a markup approach for text encoding (107). Even if the encoding scheme is completely arbitrary, consistency ensures its survivability and interchangeability. Thankfully, enough scholars have concerted their efforts into developing viable approaches to representing textual artifacts so that arbitrary approaches are unnecessary. Partly due to the significant status of the TEI in humanities computing, my familiarity with it, and also because of the qualifications it possesses for studies and analyses in literature, I selected the TEI as my encoding model.

      However, simply using the TEI without customization—unfiltered, as it were—is hardly an ideal approach. The so-called "tei_all" schema contains every possible element and attribute defined by the TEI Consortium (hence the "all" portion of its moniker), and as such is far more exhaustive than necessary for any narrow project. As a result, I defined a custom schema to assist my coding, including not only TEI's verse module, but also custom element definitions for poetic figures. These custom definitions, though, cause my schema and documents created with it not to be strictly TEI-conformant documents; that is, they would fail to validate against the tei_all schema. Particularly, the Beresford coding and this very document in its raw XML form would fail to validate due to my very specific customization of elements for rhyme and bibliographic entries.

      In particular, I have required of every <rhyme> element a specific type, limited to "eye," "imperfect," and "perfect" rhymes. This requirement is not only missing from the default TEI schema, but my requirements defy the requirements of the default by adding unsupported attributes. It is similar with my definition of the monographic element <monogr> within a bibliographic entry: my own definition requires MLA-specific information, such as the medium

of publication and the date the resource was accessed. Even should I not have redefined the requirements of several TEI elements, my own custom poetry figure element, <wfd:poetryFigure>, precludes the validation of any document against tei_all. While this element exists within the bounds of my own custom namespace (wfd:) and as such does not necessarily "break" TEI conformance, but rather extends it, attempting to code a document with my elements *without* using my custom schema would produce an error for every instance. These impediments to validation, though, are not impediments to consistency and functionality. Had I possessed deeper understanding of the minutiae of TEI conformance when I began coding, I might have approached my customization differently, but my own schema nevertheless functions as I intended.

Another important consideration, especially given the prosodic conventions of the time period during which Beresford wrote his translation, is whether to evaluate meter as written or as pronounced. For instance, in Book 5, line 26 reads, "No pow'r is ours to oppose or win our way." As written, it would be of the pattern US|US|UUS|US|US/, which deviates from the standard ten-syllable line of blank verse. Deviation is certainly to be expected, especially in a work of nearly 14,000 lines as Beresford's is; most likely, though, "to oppose" is elided and meant to be pronounced as two syllables by dropping the first "o." On the other hand, Beresford elsewhere indicates such readings as "t' oppose." This situation is certainly not isolated; such uncertainty is indicative of the difficulty inherent in choosing how to represent a text, and also reinforces the hermeneutic nature of textual analysis.

Digital Textual Analysis

By virtue of being an electronic text that contains binary character data, my edition already enables greater ease of analysis than a physical one, and certain content management

systems like XTF enhance that benefit by making use of data indexing for powerful character-, word-, and phrase-level searching. Compared to non-digitized editions, electronic texts allow users to interact with the data in any manner they see fit, including finding patterns, word frequencies, punctuation peculiarities, and more. Since the document is also marked up in TEI-XML, even more information can be gleaned, such as finding the five most frequent words limited only to prose paragraphs ("the," "of," "to," "in," and "and," in descending order) or only to verse lines ("the," "and," "of," "to," and "with"); counting the number of stanzas that have fewer than ten lines (43, in this case); or even limiting one's search to every first line of each stanza—of the 348 stanzas that occur in Beresford's translation, for example, the name "Æneas" occurs on the first line of 24 of those stanzas, nine of which occur in Book XII and six in Book V.

Having access to a completely XML-encoded poetic text also facilitates computational analysis of those figures that have been indicated by my customized markup. Carrying out such an analysis of my coded figures allows me to determine with little effort that Beresford most frequently employed figures of sound (approximately 6,950), followed by speech (approximately 900), and trailed distantly by thought (approximately 140). Each of the figures and their examples discussed in depth in earlier sections are easy to find precisely due to my marking of them. Using XSL in combination with my Beresford XML file, it is also possible to extract a vast array of information regarding the coded poetic figures, such as determining which book possesses the highest number of coded figures (Book II with 876), what the average number of figures is per book (just over 697); or any number of other various objective analytical interests related to my interpretive markup. Granted, this analysis relies on my own coding and interpretation, and I acknowledge that my coding cannot account for every instance of every

113

figure given the timeframe and the fact that I alone was responsible for coding the text. The numbers also come from a fixed point in time of the edition, even though I do not intend for the edition to remain static. Nevertheless, admitting fluctuation and the fact that these numbers likely do not account for every instance of every figure, it seems reasonable to claim that sound is by far the most prevalent type of poetic manipulation employed by Beresford. This claim makes sense considering his prolific use of metaplasm in order to fit each line to the requisite 10 syllables, as well as his frequent employment of alliteration.

Though such an analysis is reflective of only a select few figures at one point in the life of a coded edition, the practice and ability enabled by this digital edition is indicative of future prospects and possibilities. Applying the same schema to other translations would also enable cross-comparison among different editions of the same text, which would provide an additional layer of interpretation. For instance, by comparing the same lines of poetry across various translations, it would be possible to see at a glance how different translators approach both the literal Latin-to-English translation and the poetic use of figurative language. Do multiple translators employ alliteration at a certain point in the text? Do they attempt to replicate any of Virgil's Latin poetic devices in English? It would not be a difficult task to visually compare multiple editions side-by-side at the same points in the text in much the same way that Paul the Not-So-Simple functions, thereby allowing visually guided answers to these questions.

Creating a Digital Edition

The MLA-CSE guidelines for creating a scholarly edition assert that "[t]he scholarly edition's basic task is to present a reliable text: scholarly editions make clear what they promise and keep their promises" ("Guidelines"). That is, before interpretation is practicable, an accurate representation is necessary. Reliability is certainly bolstered by high-quality facsimile images,

attention to detail in planning, and coding consistency, all of which allow one to keep the promise of fidelity to a source text.

The primary goal of acquiring facsimile images is to provide reliability (Kiernan 262); reliability is necessary in transcription, in coding, and in reference. Because of the importance and reliance upon source images, consistency is paramount, much as the case with coding. Picture quality, dpi, and dimensions are all important factors to consider in this regard, and for preservation purposes should be sufficient to stand in place of the artifact as accurately as possible.

The process of keyboard transcribing the work is both the most important and the most time consuming. At the outset of this project, it was my intention to employ OCR software and supplement it with corrections where it had trouble; as it turns out, relying on the OCR would have been more trouble than keyboarding the entire work was since the OCR technology available to me was less than adequate. The prevalence of the long-s character, in particular, resulted in a plethora of mis-identified characters, even after "training" the software. As a result, my entire process diverted early from my planned schedule, and I ended up manually reproducing the entirety of the text. Despite the extra work, however, I feel that keyboarding the piece gave me a much greater familiarity with the text than automated software could have.

Once keyboarded, the next step in the production of the edition is coding. In my edition, the ODD defines the allowed list of figures to code based on the selected figures of thought, sound, and speech. It was my intention from the beginning to limit strictly the amount of customization I would impose on the default TEI set of verse elements. As such, I believe my solution of adding only a single verse-supplementing element—to which can be appended several narrowing attributes—is a strong solution to achieving my coding goals. Not only does

115

the creation of a strict customization promote interchangeability by limiting the amount of necessary coding familiarity for future modification, but it also enables a gentle learning curve with a well-defined approach to coding poetic figures—further enhanced by virtue of said custom element existing within its own namespace in order to keep it separate from the predefined TEI modules.

As the complete TEI schema is constructed out of several component modules, including one dedicated specifically to verse, an XML document that employs an unmodified TEI schema already possesses flexibility and robust support for verse coding. Nevertheless, the default TEI is not able to account for every situation, and it is for such a reason that customization is both possible and encouraged (Burnard and Bauman 668). Therefore, balancing the TEI's predefined verse elements with supplemental custom elements and attributes enables greater relevance to any given project. In the case of my own project, a newly defined element with attributes supports the coding of rhetorical figures while certain tweaks to the default TEI verse module allow more specific focus on features of interest. Particularly, one feature to which I returned on multiple occasions is that of rhyme; since the TEI verse module already contains a <rhyme> element, rather than defining a new value for rhyme, I instead opted to modify the TEI's constraints for that element to better match the characteristics of rhyme that were of interest in the Beresford translation. On the other hand, another viable option to tackle customized markup for rhyme and to provide better conformance with tei_all would have been to use my already-defined custom element, wfd:poetryFigure, and add a "rhyme" value to the @wfd:sound attribute, thereby defining any text surrounded by this element as a poetic figure of sound, of type "rhyme." This decision is the best example of my struggles with developing a custom schema to use for the project.

Means of Representing the Work

Employing a customizable and non-proprietary method of text encoding allows virtually countless choices for rendering the coded work. Moreover, "the amount of information encoded in a full-scale TEI text is difficult to represent in print," so "rendering a TEI text often means re-arranging or transforming it in some way" (Rahtz 310–311). The method of transformation Rahtz describes, using an eXtensible Stylesheet Language Transformation (XSLT), is essential to turning a customized document like this edition into something designed for display in a system more complex than simple CSS rules are capable of. With XSLT, elements can be omitted, re-arranged, and modified by a program or browser regardless of the layout of the original TEI, which CSS alone does not enable. Due to the limitations of certain technologies operating on their own, current systems employ a combination of several technologies simultaneously in order to exhibit the desired textual features. In particular, XTF and TEI Boilerplate transform the TEI into HTML using XSLT, and the HTML's display is ultimately governed by CSS rules.

Figure 11: A sample rendering of Beresford's text in XTF

       The image above shows the first page of Beresford's translation of Book II of the *Aeneid* as rendered by my customized installation of XTF. The left frame provides the book's table of contents, the middle frame the facsimile representation, and the right frame the text showing the figures I have tagged. In this instance, blue underlines indicate figures of sound, including alliteration, and italicized daggers "†" are inserted at locations where metaplasm occurs. One particular nuance to notice, and discussed earlier, is the fact that the facsimile frame renders long-s characters where they belong whereas the interpreted ignores the long-s in favor of modern "s." Both frames are generated from the same source XML, and it is only the XSL that follows a set of rules to determine which "s" characters to transform into "ſ."

— 39 —

ÆNEID.

BOOK II.

1 ALL filent fat with fixed eyes intent;
    Then from his lofty couch the Prince began:
    Sorrows, O Queen! paft utterance fevere
    Thou bidft me to revive, how Grecia's fons
    The ftrength and lamentable realms of Troy
    Up-rooted; miferies by thefe eyes beheld,
    And whofe great part I was. What man is he
    Of the Dolopian hoft or Myrmidon,
    What foldier e'en of hard Ulyffes' band,
    Who, telling fuch difaftrous fates as thefe,
    Could hold from tears?—And now the dewy night
    Adown the fky precipitate defcends,
    And finking ftars folicit to repofe.
    Yet if fuch ardour fill thy mind to learn
    Our various tribulations, and in brief
    To hear the laft and tragic toil of Troy,
    Though yet at the remembrance paft my foul
18 Thrills, and with forry back recoils, I yield.
     4

Figure 12: A focus of the facsimile frame of the XTF interface

In this image focusing on the leftmost text frame, the facsimile text is visible. In the facsimile

rendering, poetic figures are omitted, and the long-s characters are rendered as they appear in the

book.

```
1104 ▽    <xsl:template match="text()" mode="long-s-1">
1105 ▽       <xsl:analyze-string select="." regex="(dre|glo)ss('d)"><!-- regex rule for double long-s: convert any
1105    double 's' preceded by the pattern 'dre' or 'glo' and followed the pattern "'d" (matches: ad[dress'd] |
1105    [gloss'd] ) -->
1106 ▽          <xsl:matching-substring>
1107                <xsl:value-of select="regex-group(1)"/><xsl:text>&#x17f;&#x17f;</xsl:text><xsl:value-of
1107    select="regex-group(2)"/>
1108          </xsl:matching-substring>
1109 ▽          <xsl:non-matching-substring>
1110                <xsl:value-of select="."/>
1111          </xsl:non-matching-substring>
1112       </xsl:analyze-string>
1113    </xsl:template>
1114
1115 ▽    <xsl:template match="text()" mode="long-s-2">
1116 ▽       <xsl:analyze-string select="." regex="ss(['aeiou][^d\W])"><!-- regex rule for double long-s: convert any
1116    double 's' followed by an apostrophe or a vowel, followed by any letter besides 'd' (matches: rece[sses] |
1116    le[ss'n]ing | to[ssin]g | pui[ssan]t | le[sson]s ) -->
1117 ▽          <xsl:matching-substring>
1118                <xsl:text>&#x17f;&#x17f;</xsl:text><xsl:value-of select="regex-group(1)"/>
1119          </xsl:matching-substring>
1120 ▽          <xsl:non-matching-substring>
1121                <xsl:value-of select="."/>
1122          </xsl:non-matching-substring>
1123       </xsl:analyze-string>
1124    </xsl:template>
1125
1126 ▽    <xsl:template match="text()" mode="long-s-3"><!-- regex rule for double long-s: convert any double 's'
1126    preceded by an 'A' or 'a' or 'o' and followed by any letter (matches: [assi]duous | m[assy] | [Assa]racus |
1126    gl[ossy] ) -->
1127 ▽          <xsl:analyze-string select="." regex="([Aao])ss(\w)">
1128 ▽             <xsl:matching-substring>
1129                   <xsl:value-of select="regex-group(1)"/><xsl:text>&#x17f;&#x17f;</xsl:text><xsl:value-of
1129    select="regex-group(2)"/>
1130             </xsl:matching-substring>
1131 ▽             <xsl:non-matching-substring>
1132                   <xsl:value-of select="."/>
1133             </xsl:non-matching-substring>
1134          </xsl:analyze-string>
1135    </xsl:template>
1136
```

Figure 13: Some of the XSL rules for converting "s" to "ſ"

As the long-s characters were not actually coded as part of the text, it bears explanation of how

they have been rendered in the project. Once I determined the various locations where long-s was

used instead of "s," I produced XSL templates (above) to match sections of text that followed

those rules and convert the modern "s" to "ſ."

— 39 —
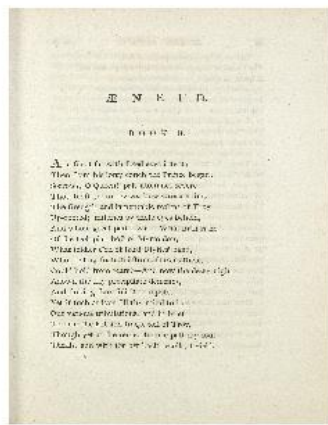
ÆNEID.

BOOK II.

1 ALL silent sat with fixed eyes intent;
   Then from his lofty couch the Prince began:
   Sorrows, O Queen! past utterance severe
   Thou bidst me to revive, how Grecia's sons
   The strength and lamentable realms of Troy
   Up-rooted; miseries by these eyes beheld,
   And whose great part I was. What man is he
   Of the Dolopian host or Myrmidon,
   What soldier e'en[†] of hard Ulysses' band,
   Who, telling such disastrous fates as these,
   Could hold from tears?—And now the dewy night
   Adown[†] the sky precipitate descends,
   And sinking stars solicit to repose.
   Yet if such ardour fill thy mind to learn
   Our various tribulations, and in brief
   To hear the last and tragic toil of Troy,
   Though yet at the remembrance past my soul
18 Thrills, and with sorry back recoils, I yield.
   4

Figure 14: A focus of the interpreted frame of the XTF interface

Here, the rightmost frame is highlighted, showing the interpreted version of the text; poetic

figures and features are highlighted as a visible part of the text, and long-s is converted to

modern "s" for improved readability. When figures are moused over, information appears to indicate how each section of text has been coded.

Particular XTF files most heavily customized for my XML coding include component.xsl, teiDocFormatter.xsl, and docFormatterCommon.xsl. These three files contain all the rules for converting the descriptive XML into procedural HTML and displaying the resulting transformed text within the set of frames that serve as the presentation area for XTF.

```
198 ▽    <xsl:template match="*:l">
199 ▽       <xsl:choose>
200 ▽          <xsl:when test="parent::*:lg">
201 ▽             <tr>
202 ▽                <td width="30">
203 ▽                   <xsl:choose>
204 ▽                      <xsl:when test="$interpret!='1'">
205 ▽                         <xsl:choose>
206 ▽                            <xsl:when test="@n">
207                                <span class="run-head"><xsl:value-of select="@n"/></span>
208                             </xsl:when>
209                             <xsl:otherwise><xsl:text> </xsl:text></xsl:otherwise>
210                         </xsl:choose>
211                      </xsl:when>
212 ▽                      <xsl:when test="$interpret='1'">
213 ▽                         <xsl:choose>
214 ▽                            <xsl:when test="@n and not(@real)">
215                                <span class="run-head"><xsl:value-of select="@n"/></span>
216                             </xsl:when>
217 ▽                            <xsl:when test="@real">
218                                <span class="irreg-line" title="Irregular line of pattern
218    {./@real}"><b><xsl:text>[*]</xsl:text></b></span>
219                             </xsl:when>
220                             <xsl:otherwise><xsl:text> </xsl:text></xsl:otherwise>
221                         </xsl:choose>
222                      </xsl:when>
223                   </xsl:choose>
224                </td>
225 ▽                <td>
226                   <xsl:if test="parent::*:lg[preceding-sibling::*[1][self::*:lg]] and
226    .[not(preceding-sibling::*:l)]"><xsl:text>   </xsl:text></xsl:if>
227                   <xsl:apply-templates/>
228                </td>
229             </tr>
230          </xsl:when>
231 ▽          <xsl:when test="parent::*:note">
232             <span class="fnote-l"><xsl:apply-templates/></span><br/>
233          </xsl:when>
234 ▽          <xsl:otherwise>
235 ▽             <xsl:if test="@n">
236                <span class="run-head"><xsl:value-of select="@n"/></span>
237             </xsl:if>
238             <xsl:apply-templates/><br/>
239          </xsl:otherwise>
240       </xsl:choose>
241    </xsl:template>
```

Figure 15: A sample transformation template from XTF's component.xsl

Shown here is the set of rules governing a single line of poetry for display in both the facsimile and interpreted frames. The stylesheet checks the conditions of each line and, given the matching rule it encounters, applies a specific transformation to that line. Rules exist for both main content frames of XTF—facsimile and interpreted—and a special rule is defined for the first line of each new stanza *after* the first one, since the initial stanza of each book is not indented.

Like the XTF content management system, TEI Boilerplate contains a system of interconnected files for transforming and displaying a source XML document. Unlike XTF, however, the TEI Boilerplate leaves a much lighter digital footprint: its only requirement is a modern web browser, and it can be run in either a server or client environment. That is, while it can be placed on a webserver for access via the Internet, TEI Boilerplate can also be run on a local computer without a server framework, or even without an Internet connection. This system allows quick visualization with very little set-up time. With a TEI schema, a user can define basic XSL transformations and CSS rules to highlight desired structural elements with just the assistance of a ubiquitous web browser.

ÆNEID.
BOOK II.

ALL silent sat with fixed eyes intent;
Then from his lofty couch the Prince began:
Sorrows, O Queen! past utterance severe
Thou bidst me to revive, how Grecia's sons
The strength and lamentable realms of Troy
Up-rooted; miseries by these eyes beheld,
And whose great part I was. What man is he
Of the Dolopian host or Myrmidon,
What soldier e'en of hard Ulysses' band,
Who, telling such disastrous fates as these,
Could hold from tears?—And now the dewy night
Adown the sky precipitate descends,
And sinking stars solicit to repose.
Yet if such ardour fill thy mind to learn
Our various tribulations, and in brief
To hear the last and tragic toil of Troy,
Though yet at the remembrance past my soul
Thrills, and with sorry back recoils, I yield.

Broken in battle and repuls'd by fate,
The Chiefs of Greece, when years on years had roll'd,
A horse, at length, of mountain magnitude
Up-build through Pallas' art divine, and weave
Its ribs with beams of fir: 'tis feign'd a vow
For safe return. Such rumour wanders wide.
In its blind body a destin'd band select
They shut secreted, and with armed hosts
Its womb, capacious dungeon, fill throughout.

Figure 16: A sample rendering of Beresford's text in TEI Boilerplate

Shown here is the same source XML as rendered in TEI Boilerplate. In this case, only a single

version of the text is shown at a time when the user selects from multiple sets of CSS rules. In

this customized installation, it is possible for the user to switch between transcribed and interpreted text, each one displaying similar information to the frames described above in XTF. The XSL in TEI Boilerplate is limited to XSLT 1.0 for display in web browsers, which is significantly less powerful than XTF's XSLT 2.0 and offers fewer options for transformation. Thus it is not possible to configure rules for displaying the long-s character. Such limitations, however, are counter-balanced by the fact that Boilerplate is a light-weight software package that can be run even without a network connection, and merely requires a web browser to display online, unlike XTF, which requires a web server and additional servlet packages that support Java programming—in my implementation, the Java servlet is Apache Tomcat.

Files relevant to displaying XML files within TEI Boilerplate that I modified or created include teibp.xsl, custom.xsl, custom_interp.css, and custom_trans.css. The two former files govern the transformation of XML documents into HTML, while the two latter files govern the display of the HTML output.

```
21 ▽    <xsl:template match="wfd:poetryFigure">
22 ▽        <xsl:choose>
23 ▽            <xsl:when test="@wfd:sound">
24 ▽                <poetryFigure>
25                      <xsl:attribute name="class">sound</xsl:attribute>
26                      <xsl:attribute name="title"><xsl:value-of select="@wfd:sound"/></xsl:attribute>
27                      <xsl:apply-templates/>
28                  </poetryFigure>
29            </xsl:when>
30 ▽            <xsl:when test="@wfd:speech">
31 ▽                <poetryFigure>
32                      <xsl:attribute name="class">speech</xsl:attribute>
33                      <xsl:attribute name="title"><xsl:value-of select="@wfd:speech"/></xsl:attribute>
34                      <xsl:apply-templates/>
35                  </poetryFigure>
36            </xsl:when>
37 ▽            <xsl:when test="@wfd:thought">
38 ▽                <poetryFigure>
39                      <xsl:attribute name="class">thought</xsl:attribute>
40                      <xsl:attribute name="title"><xsl:value-of select="@wfd:thought"/></xsl:attribute>
41                      <xsl:apply-templates/>
42                  </poetryFigure>
43            </xsl:when>
44        </xsl:choose>
45    </xsl:template>
```

Figure 17: The template rule for TEI Boilerplate contained with custom.xsl

These lines govern what the content management system will do when it encounters my custom

element within the source XML. In this case, it will create a new element called <poetryFigure>

that can then be styled using CSS.

```
1   /* Place project-specific CSS here */
2 ▽ fw {
3   display:none;
4   }
5
6 ▽ .sound {
7       display:inline;
8       border-bottom:blue 1px dotted;
9   }
10
11 ▽ .sound:hover {
12      border:blue 1px solid;
13  }
14
15 ▽ .speech {
16      display:inline;
17      border-bottom:green 1px dotted;
18  }
19
20 ▽ .speech:hover {
21      background:green;
22  }
23
24 ▽ .thought {
25      display:inline;
26      border-bottom:red 1px dotted;
27  }
28
29 ▽ .thought:hover {
30      background:red;
31  }
```

Figure 18: The contents of custom_interp.css

This file contains the custom CSS rules that TEI Boilerplate uses to render figures of sound,

speech, and thought that are generated by the custom.xsl transformation stylesheet.

        Each of the above files—both in XTF and TEI Boilerplate—fulfills a particular role in

the representation of the edition. The XML files encompass the full content of the work,

including the transcription and custom coding, and are the source upon which all other attendant

files act. The XSL serves as an intermediate step, transforming the XML into markup language

interpretable by web browsers. Such "final" HTML pages thus generated are created on demand by a server and are further supplemented by CSS rules for on-screen presentation.

While XTF and TEI Boilerplate are not the only content management systems designed to render TEI-encoded documents, they exemplify two philosophies that result in transformed documents. I also chose them instead of others as they have both been employed by projects that have directly influenced this very project. Finally, while the edition exists as a transformed and digitally rendered document, it does not exist in a virtual vacuum devoid of context. This document seeks to situate and justify not only the creation of this edition, but the digitization and encoding approach as a legitimate means of textual analysis for the humanities as a whole.

CHAPTER SEVEN: THE PROJECT'S PRESENT AND FUTURE

"The form that electronic editions will typically take, though it has not yet settled down, will almost certainly have to respect the logic of their new environment. An electronic counterpart of the printed scholarly edition will presumably therefore take advantage of the visualization capacities of the medium . . . It will preferably offer an automatic method of collating and displaying the variant readings. Unlike the electronic editions of the 1990s and 2000s, it will probably open its files up to collaborative interpretation, including perhaps by readers"

(Eggert 116)

Digital scholarship is an ever-expanding field, made ever more accessible by a widening recognition and advancement of technologies designed specifically to augment the traditional roles of humanist scholars. Historians, poets, librarians, classicists, linguists, literary critics, and more stand to benefit from the possibilities offered by customized markup, the basics of which I have demonstrated with the digital representation of Beresford's translation.

The digital edition that I have produced is a relatively simple visualization made possible by an intricate interplay of texts, technologies, representations, and interpretations spanning more than 2,000 years. The digital markup that renders the edition on a computer screen is not possible without the oral text, the origin stories of Rome as a city, from which the *Aeneid* was born during the Early Roman Empire. Now, the text has become a part of the greater text of the World Wide Web, whose elaborate allusions, references, and expanding interconnectedness lend an entire new complexity to Barthes' idea of the text as a network.

What differentiates my own edition from others discussed herein is the fact that interpretive markup has been included and displayed within my visualization platforms. While editions like the Mark Twain Project or the Perseus Project include editorial notes as associated links from within the digitally rendered text, they refrain from including editorial intervention as *part of* the displayed product. The Twain digital edition allows readers to switch between a plain view for reading and an "interactive" view for viewing editorial markup, but neither view quite

matches the approach of my edition, wherein readers can see as *part of the text* my editorial markup next to an unmarked facsimile rendition of the same portion of the text. Moreover, should readers want to view *only* the facsimile or interpreted versions, that option is also available.

It is this latter fact, especially, that leads me to believe that my edition follows the basic expectation of an MLA-CSE edition: the creation of a reliable text. While one frame features an inextricable layer of interpretive coding, the other frame is not only free from said coding, but reflects an attempt to present a text as faithful to the original as possible. Moreover, the high-quality images of each page are accessible as part of the digital text. In its current state, however, I cannot claim full adherence to the MLA-CSE principles. Given the fact that I alone produced the digital edition and the amount of work inherent in it, no systematic proofing of the transcription has taken place.

Reliability also extends to internal consistency; the use of my custom-defined schema ensures that my own XML coding is reliable to itself, and provides certainty in the regularity of my structural markup. While this cannot ensure that I did not introduce typographical errors into the textual transcription, or that I interpreted my chosen poetic figures correctly and accurately in every instance, it does guarantee that straying from my determined figures was not possible. Should I hope to continue the production of this edition, several key issues deserve consideration—not the least of which is a thorough proofing—before any enhancement can occur.

Difficulties Along the Way

Change has been one of the few constants since the inception of this project. Besides the unexpected lack of help offered by OCR discussed in the previous chapter, other challenges

presented themselves over the course of this project that forced me to re-evaluate my approach and my underlying assumptions. In particular, the ODD file underwent several revisions over the first few months, with regular modifications, though a more-or-less "final" version was necessary in order for consistent coding to remain feasible. As it happened, the first several months of coding required ongoing re-evaluation of how to define and subsequently code figures, as in the case of syncope. Initially, syncope was one of the allowed attribute values for @wfd:sound; given other possibilities of syllabic manipulation, such as prothesis and aphaeresis, a more comprehensive solution was to group all of these possibilities under the more general grouping of "metaplasm."

Based on the wide realm of possibilities offered by poetic analysis, it is unsurprising that several figures might apply to the same or overlapping streams of text. Most common among my chosen analytic set are syncope and hypermonosyllable—specifically, *every* instance of hypermonosyllable also fits the criteria of syncope, but not always vice versa. For instance, the word "hour" could be pronounced as either one or two syllables, as fits the meter of the line. However, Beresford employs no typographical distinction to indicate which case is single- or double-syllable. On the other hand, "pow'r" is clearly indicated as being syncopated by virtue of the apostrophe, but it is also distinguished as such because it appears as "power" in many instances wherein it spans two syllables. One decision I had to make early, then, was how to code specific words that often occur as hypermonosyllable and *not* as syncope in order to give precedence to the fewer examples of the former. It is possible—and occasionally manifest—for more than one single figure to be coded in the same section of text, though such occurrences are less obvious in XTF and TEI Boilerplate.

In a similar vein, the pairs of elision and aphaeresis as well as anaphora and polysyndeton often overlap, offering further categorization and coding challenges. In most cases, if it was only possible to code a single figure, I gave precedence to the one that seemed most important to the understanding of the text at that single point. Ultimately, the selection is interpretive and subjective, just like my coding of each poetic figure in the translation.

Further, due to XML's nesting structure, some creativity had to be applied when tagging figures that occurred over multiple lines. For example, alliteration that began on one line and continued to the next could not simply be opened where it began and closed where it ended as that would violate XML's requirement for well-formedness by interrupting its inherent nesting structure—one significant shortcoming of the selected technology for this project. Thus, other options had to be pursued, including deciding whether to use the @next and @prev attributes—or, what I eventually decided on—the @part attribute to associate a single rhetorical figure comprising multiple elements.

Reflections

As a text, the *Aeneid* has existed for over two millennia and spans all human forms of expression. There is no shortage of performances, translations, or scholarship of it in any medium, whether spoken, handwritten, painted, printed, sculpted, or electronic. Though a poetic analysis in itself is not a unique approach to a poetic text, and a digital encoding is a standard approach for preservation and online representation, my edition is the first to do both at once and subsequently present that poetic analysis for digital consumption. It possesses neither the linguistic depth of the Perseus Project nor the linked frames of Paul the Not-So-Simple and the Mark Twain Project, but it takes elements from each of those and I believe showcases some of what is possible given more time and scope.

While operating in the same vein of some of those projects, my own also offers something more: an interpretive, editorial layer realized through HTML visualization. The TEI projects listed herein might offer a degree of editorial markup in the form of commentary apparatus, but my edition is unique compared to similar ones due to its transformation of a custom element into a display layer on top of the text. Nor is any information from the original text irretrievably deformed by such transformation: should a reader desire only to view the facsimile form of the text, such is available in the same location with the same effort and generated from the same source XML.

Of course, visualization is not limited strictly to HTML. The Myopia poetry visualization tool has been demonstrated by the Poetess Archive, and it appears to be a helpful approach for scholars and other interested individuals by enabling similar poetic visualization to that which I have encoded in my edition through a customized TEI extension. Beyond poetic tropes, the tool analyzes and displays qualitative rhythm and meter, including short and long syllables, based on the encoding of each line in the source TEI file. While the tool is a useful one for the transformation of a TEI document into a text ready for visual analysis, it has only been demonstrated on a very limited scale, and it still requires a more technically specialized client-side framework that prevents it from operating on a webserver. The software requirements also present a higher technical hurdle to surmount. Nevertheless, its schema requirements are suitably exacting in order for the program to display meter and figures appropriately, and they parallel my own; such requirements will be essential if future interchange is desired, whether for online distribution or for offline representation on one's personal computer.

It is also my hope that others who might have been unaware of the means of digital representation see this edition and recognize what is possible; as a fledgling classicist beginning

my undergraduate career studying Ancient Greek, I could never have imagined such an edition, let alone the fact that I could create it myself with little more than the computer I already employed to play video games and write essays. Given the technologies available to increasingly wider audiences, in smaller and faster forms, it is likely that coming years will witness an increase in the number of scholars not only believing in this approach, but actively pursuing and engaging in it.

From my own experience I know that the simple fact that the technology exists does not make the process "easy"—even for someone who has worked with such technology for as long as I have. But a beautiful trait of interdisciplinarity is that there is an increasing will and an increasing number of people with the requisite knowledge and enthusiasm to support these projects.

The Future

> "Unlike a traditional book or set of books, the hypertext need never be 'complete'"
> (McGann, Radiant Textuality 71)

Though I present here a complete digital edition, the potential for this project still looms large. In any edition, complete satisfaction always seems just out of reach. However, given the digital nature of this one, any desired change can be accomplished more easily than in the traditional realm of scholarly editing. The most obvious "next step" for this digital edition is expansion and renovation. That is, knowing that nearly every time I read a passage I can identify another figure to code or a different means of coding from that which I've already identified, there can be no foreseeable end to my own application of markup.

The future potential of the project is open and involves a veritable laundry list of items to reconsider, including the re-evaluation of various groups of figures, the addition of currently

135

excluded figures, improving searchability, and building corresponding English-Latin visualizations within the XTF and TEI Boilerplate interfaces—or even within other platforms that this project does not yet implement. I have contemplated several coding revision possibilities related to poetic figures, including the re-evaluation of anaphora, diacope, symploce, and epistrophe as subsets of a new "repetition" speech group, and simile/metaphor as subsets of a new "analogy" thought group. These coding changes are minor, however, and at the current juncture would serve me little to implement, especially since in my reading thus far I have encountered no textual units in Beresford's translation I would identify as symploce or epistrophe. If, on the other hand, I expand my current project to include more than just Beresford's text as a source, the addition of these omitted figures will be a minimum necessity. Were such the case, the figures I have thus far selected will serve only as a starting point, and I will have to consult Harmon, Abrams, and their like much more exhaustively for a wider range of figures befitting a larger corpus.

Although the XTF CMS allows searching within the full text and metadata of the items it contains, its default implementation does not enable the searching of custom XML tags; that is, it is not currently possible for users to search for specific instances of poetic figures, which is admittedly a drawback considering the goals of this edition.

In addition to my desire for straightforward coding enhancement, other possibilities include the adoption of automated processes. Research in artificial intelligence and computer language recognition might allow for the automatic recognition of specific figures, such as alliteration, anaphora, simile, and metaphor that I otherwise might have overlooked. These fields are well outside my realm of expertise, however, and therefore an opportunity presents itself for

researchers outside of the humanities to become involved, potentially further increasing interdisciplinary collaboration on university campuses.

Regardless of the future scope of the edition, I anticipate corrections to my ODD file for the purposes of TEI conformance and clearer documentation. To the former end, I must at least re-evaluate my customization of the TEI verse <rhyme> element and perhaps migrate its role to my custom <wfd:poetryFigure> element as a value of @wfd:sound. Using this method would facilitate interchange by not forcing a change onto the TEI's pre-defined implementation of the <rhyme> element, and it would better support my desire to create a schema that enables true TEI-conformant documents rather than TEI extensions.

Any of these options is feasible under the auspices of interchange: even should another person wish to take up the project, my customization is straightforward enough that anyone with digital markup proficiency could continue and augment my process. This collaborative possibility speaks directly to Eggert's prediction that newer scholarly editions will allow scrutiny and input by not only experts and project curators, but also by a distributed audience. In the case of the Beresford text, this would allow a diverse crowd of poets, classicists, and those interested in textual preservation to offer their unique perspectives and insights to the ongoing enhancement of the edition.

As acceptance of and engagement with textual representation in the digital mode increases, I am hopeful that similar approaches to my own may take root. This possibility is feasible considering the similarly increasing pervasiveness of digital humanities technologies and the literacy required to use them. Additionally, with the huge numbers of digitized texts and born-digital documents produced by academics, it is becoming more important for institutions to

support digital projects and initiatives, both in terms of access and expertise. My own project illustrates how these various concerns can be taken up by a small group or even a single person.

Concluding Thoughts

Ultimately, this project is about the representation of text. As I have asserted, though, both "representation" and even "text" themselves are not necessarily straightforward concepts to define. More fundamental ideas, including mediation and interpretation, must be considered before speaking about the full representation of any textual object. Again, then, this project is about the representation of text—as demonstrated by *the representation of a text*, which itself is only part of a history of many texts collectively called the *Aeneid*. Whether a text is an ordered hierarchy of content objects or a deformative chain of transformation stretching back to an inaugural idea, text cannot exist without representation, mediation, and interpretation.

A text becomes interpreted—for better or worse—when it has been marked up. Markup can serve several purposes and can exist in every medium. It can be oral, whereby a story retold by a secondary source introduces variation to a narrative, or it can be physical, as when an editor corrects a manuscript with a set of correction marks, or it can be digital, when a printer prepares a digital draft for distribution by using a typesetting markup language like TeX to specify font size, color, and justification. In all cases, interpretation results in transformation.

In application, markup serves too many purposes to be meaningful without further specification. As a result, scholars have spent decades honing digital markup tools to represent humanistic works in digital environments like the internet. With structural markup like SGML and its descendants XML and HTML, digital technologies have evolved to the point of being able to represent textual artifacts in a digital space for the continued interpretation and consumption of other interested parties. Though this edition is not the first TEI-encoded edition

of Beresford's translation, it is the first marked up digitally with the explicit intention of integrating textual interpretation as part of the encoding. By merging digitization and textual encoding, I hope to demonstrate that interdisciplinarity is alive and well, and is only becoming more relevant as time advances and digital technologies become ever more accessible.

By creating a new edition of a previous translation of the *Aeneid*, I have effectively contributed a new interpretation to the long-running textual history of the text. Based on the suggestions of editors more experienced than myself, I have made every effort to make careful, meaningful decisions in my text to ensure that what I have contributed is neither superfluous nor destructive. While the goal is to present my own poetic analysis of the text, I want to do so without imposing myself onto the original translation. This conservative editorial philosophy induced me to replicate obvious errors and leave archaic spelling and punctuation as it was originally rendered.

It is my hope that this project proves to some extent boundary-crossing, eye-opening, innovative, and enabling. It is a proof-of-concept that I hope can inspire linguists, translators, poets, historians, and archivists to see something they might not have considered before. While I have my own expectations of where the project might proceed from its current state, it is certainly possible that someone else might take up its cause and drive its interpretation in another direction entirely.

APPENDIX: CURRENTLY ACCESSIBLE PROJECT LOCATIONS AND PLATFORMS

	TEI Boilerplate	*eXtensible Text Framework 3.1*
Personal Server	http://wdorner.com/teibp/content/beresford_teibp.xml	http://wdorner.com:8080/xtf/view?docId=aeneid/beresford.xml;query=;brand=wfd
University Server	http://tandtprojects.cah.ucf.edu/~wi545036/teibp/content/beresford_teibp.xml	http://students.cah.ucf.edu/xtf_wfd/view?docId=aeneid/beresford.xml;query=;brand=wfd

REFERENCES

Abrams, M. H., and Geoffrey Harpham. *A Glossary of Literary Terms*. Eleventh Edition. Boston: Wadsworth Cengage Learning, 2014. Print.

Alpers, Svetlana. "Describe or Narrate? A Problem in Realistic Representation." *New Literary History* 8.1 (Autumn 1976): 15–41. Web. 2 Aug. 2012.

Anderson, William S. "Five Hundred Years of Rendering the Aeneid in English." *Reading Vergil's Aeneid: An Interpretive Guide*. Ed. Christine Perkell. Norman, OK: U of Oklahoma P, 1999. 285–302. Print.

Applen, J.D., and Rudy McDaniel. *The Rhetorical Nature of XML: Constructing Knowledge in Networked Environments*. New York: Taylor & Francis, 2009. Print.

Arroyo, Sarah J. ": Exploring Rhetorical Convergences in Transmedia Writing." *From A to <A>: Keywords of Markup*. Eds. Bradley Dilger and Jeff Rice. Minneapolis: U of Minnesota P, 2010. Print.

Bakker, Egbert J. "Mimesis as Performance: Rereading Auerbach's First Chapter." *Poetics Today* 20.1 (Spring 1999): 11–26. Web. 31 Mar. 2011.

Barnard, David, Ron Hayter, Maria Karababa, George Logan, and John McFadden. "SGML-Based Markup for Literary Texts: Two Problems and Some Solutions." *Computers and the Humanities* 22.4 (1988): 265–276. Web. 31 Mar. 2011.

Barthes, Roland. "From Work to Text." *Image-Music-Text*. New York: Hill and Wang, 1977. 155–164. Print.

Bartsch, Shadi, and Jas Elsner. "Introduction: Eight Ways of Looking at an Ekphrasis." *Classical Philology* 102.1 (January 2007): i-vi. Web. 2 Aug. 2012.

Bauman, Syd. "Interchange vs. Interoperability." *Proceedings of Balisage: The Markup Conference 2011 Balisage Series on Markup Technologies*. Montréal, Canada: Balisage, August 2 - 5, 2011. Web.

Beresford, James. *The Aeneid of Virgil, Translated into Blank Verse*. London: J. Johnson, 1794.

Berners-Lee, Tim. "Mark-up: From styling to structuring information." Issues in Informatics, 21 Apr. 1992. Web. 27 Jan. 2015.

Berners-Lee, Tim. "SGML." W3C, 1992. <http://www.w3.org/History/19921103-hypertext/hypertext/WWW/MarkUp/SGML.html>. 5 Dec. 2012.

Berrie, Phill, Paul Eggert, Chris Tiffin, and Graham Barwell. "Authenticating Electronic Editions." *Electronic Textual Editing*. Eds. Lou Burnard, Katherine O'Brien O'Keeffe, and John Unsworth. New York: The Modern Language Association of America, 2006. 269–276. Print.

Bolter, Jay David. *Writing Space: Computers, Hypertext, and the Remediation of Print*. Second Edition. Mahwah, NJ: Lawrence Erlbaum, 2001. Print.

Borgman, Christine L. *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA: MIT Press, 2007. Print.

Brower, Reuben A. "Visual and Verbal Translation of Myth." *A Companion to Vergil's Aeneid and Its Tradition*. Eds. Joseph Farrell and Michael C. J. Putnam. Malden, MA: Wiley-Blackwell, 2010. 270–289. Print.

Burnard, Lou. "What is SGML and How Does It Help?" *Computers and the Humanities* 29 (1995): 41–50. Web. 22 Feb. 2010.

Burnard, Lou, and Syd Bauman, eds. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version 2.6.0. Web. 20 January 2014.

Burnett, Ron. *How Images Think*. Cambridge, MA: MIT Press, 2004. Print.

Busa, Roberto. "Foreword: Perspectives on the Digital Humanities." *A Companion to Digital Humanities*. Eds. Susan Schreibman, Raymond George Siemens, and John Unsworth. Malden, MA: Blackwell, 2004. xvi-xxi. Print.

Buzzetti, Dino. "Digital Editions and Text Processing." *Text Editing, Print and the Digital World*. Eds. Marilyn Deegan and Kathryn Sutherland. Burlington, VT: Ashgate Publishing, 2008. 45–61. Print.

Buzzetti, Dino, and Jerome McGann. "Critical Editing in a Digital Horizon." *Electronic Textual Editing*. Eds. Lou Burnard, Katherine O'Brien O'Keeffe, and John Unsworth. New York: The Modern Language Association of America, 2006. 53–73. Print.

Cavallo, Guglielmo, and Roger Chartier. "Introduction." *A History of Reading in the West*. Eds. Guglielmo Cavallo and Roger Chartier. Amherst: U of Massachusetts P, 1999. 1–36. Print.

"The Charles Brockden Brown Electronic Archive and Scholarly Edition." U of Central Florida, Center for Humanities and Digital Research. <http://brockdenbrown.cah.ucf.edu/>.

Chartier, Roger. "Reading Matter and 'Popular' Reading: From the Renaissance to the Seventeenth Century." *A History of Reading in the West*. Eds. Guglielmo Cavallo and Roger Chartier. Amherst: U of Massachusetts P, 1999. 269–283. Print.

Clark, Dave. "Content Management and the Separation of Presentation and Content." *Technical Communication Quarterly* 17.1 (2008): 35–60. Web. 14 Nov. 2011.

Clark, Elizabeth A. "Texts and Contexts." *History, Theory, Text: Historians and the Linguistic Turn*. Cambridge, MA: Harvard U P, 2004. 130–155. Print.

Cohen, Daniel J., and Roy Rosenzweig. *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. Philadelphia: U of Pennsylvania P, 2006. Web.

Coombs, James H., Allen H. Renear, and Steven J. DeRose. "Markup Systems and the Future of Scholarly Text Processing." *Communications of the Association for Computing Machinery* 30.11 (Nov. 1987): 933–947. Web. 15 Jul. 2012.

Crane, Greg. "Classics and the Computer: An End of the History." *A Companion to Digital Humanities*. Eds. Susan Schreibman, Raymond George Siemens, and John Unsworth. Malden, MA: Blackwell, 2004. 46–55. Print.

DeRose, Steven J., David G. Durand, Elli Mylonas, and Allen H. Renear. "What is Text, Really?" *Journal of Computing in Higher Education* 1 (1990): 3–26.

Drucker, Johanna, and Emily McVarish. *Graphic Design History: A Critical Guide*. Upper Saddle River, NJ: Pearson Prentice Hall, 2009. Print.

Durusau, Patrick. "Why and How to Document Your Markup Choices." *Electronic Textual Editing*. Eds. Lou Burnard, Katherine O'Brien O'Keeffe, and John Unsworth. New York: The Modern Language Association of America, 2006. 299–309. Print.

Eggert, Paul. "Apparatus, Text, Interface: How to Read a Printed Critical Edition." *The Cambridge Companion to Textual Scholarship*. Eds. Neil Freistat and Julia Flanders. New York: Cambridge U P, 2013. 97–118. Print.

Finneran, Richard J. *The Literary Text in the Digital Age*. Ann Arbor: U of Michigan P, 1996. Print.

Freistat, Neil, and Julia Flanders, eds. *The Cambridge Companion to Textual Scholarship*. New York: Cambridge U P, 2013. Print.

Freistat, Neil, and Steven Jones. "The Poem and the Network: Editing Poetry Electronically." *Electronic Textual Editing*. Eds. Lou Burnard, Katherine O'Brien O'Keeffe, and John Unsworth. New York: The Modern Language Association of America, 2006. 105–121. Print.

Goldfarb, Charles F. *The SGML Handbook*. Ed. Yuri Rubinsky. Oxford: Clarendon Press, 1990.

Harmon, William. *A Handbook to Literature*. Twelfth Edition. New York: Pearson Longman, 2012. Print.

Hayles, N. Katherine. *Electronic Literature*. Notre Dame: U of Notre Dame, 2008. Print.

Hayles, N. Katherine. *My Mother Was a Computer*. Chicago: The U of Chicago P, 2005. Print.

Haynes, Cynthia. "Afterword: <meta> Casuistic Code." *From A to <A>: Keywords of Markup*. Eds. Bradley Dilger and Jeff Rice. Minneapolis: U of Minnesota P, 2010. Print.

Headrick, Daniel R. *When Information Came of Age: Technologies of Knowledge in the Age of Reason and Revolution, 1700–1850*. Oxford: Oxford U P, 2000. Print.

Heffernan, James A. W. *Museum of Words: The Poetics of Ekphrasis from Homer to Ashbery*. Chicago: U of Chicago P, 1993. Print.

Hockey, Susan. "The History of Humanities Computing." *A Companion to Digital Humanities*. Eds. Susan Schreibman, Raymond George Siemens, and John Unsworth. Malden, MA: Blackwell, 2004. 3–19. Print.

Ide, Nancy M., and C. Michael Sperberg-McQueen. "The TEI: History, Goals, and Future." *Computers and the Humanities* 29.1 (1995): 5–15. Web. 30 Sep. 2010.

"Ivanhoe - a ludic playspace." NINES. <http://www.ivanhoegame.org/>.

Jewell, Andrew. "The Willa Cather Archive." U of Nebraska-Lincoln, Center for Digital Research in the Humanities. <http://cather.unl.edu/>.

Kamrath, Mark L., Philip Barnard, Rudy McDaniel, William Dorner, Kevin Jardaneh, Patricia

    Carlton, and Josejuan Rodriguez. "The Charles Brockden Brown Electronic Archive:

    Mapping Archival Access and Metadata." *Archive Journal* 4 (Spring 2014). Web. 7 May

    2014.

Kay, Michael. *XSLT 2.0 and XPath 2.0 Programmer's Reference*. Fourth Edition. Wrox Press,

    2008. Print.

Kiernan, Kevin. "Digital Facsimiles in Editing." *Electronic Textual Editing*. Eds. Lou Burnard,

    Katherine O'Brien O'Keeffe, and John Unsworth. New York: The Modern Language

    Association of America, 2006. 262–268. Print.

Kirschenbaum, Matthew G., and Doug Reside. "Tracking the changes: textual scholarship and

    the challenge of the born digital." *The Cambridge Companion to Textual Scholarship in*

    the Digital Age. Eds. Neil Freistat and Julia Flanders. New York: Cambridge U P, 2013.

    257–273. Print.

Lamport, Leslie. "Document Production: Visual or Logical?" *TUGboat* 9.1 (1988): 8–10. Web. 8

    Jul. 2012.

Landow, George P. *Hypertext 3.0*. Third Edition. Baltimore: The Johns Hopkins U P, 2006.

    Print.

Lavagnino, John. "When Not to Use TEI." *Electronic Textual Editing*. Eds. Lou Burnard,

    Katherine O'Brien O'Keeffe, and John Unsworth. New York: The Modern Language

    Association of America, 2006. 334–338. Print.

Lessig, Lawrence. *Free Culture: How Big Media Uses Technology and the Law to Lock Down*

    Culture and Control Creativity. New York: Penguin, 2004. Print.

"Mark Twain Project :: Home." U of California-Berkeley. <http://www.marktwainproject.org/>.

McGann, Jerome. "Radiant Textuality." *Victorian Studies* 39.3 (1 April 1996): 379–390. Web.

    27 Sep. 2014.

McGann, Jerome. *Radiant Textuality: Literature after the World Wide Web*. New York:

    Palgrave, 2001. Print.

Miles, Gary B. "The Aeneid as Foundation Story." *Reading Vergil's Aeneid: An Interpretive*

    Guide. Ed. Christine Perkell. Norman, OK: U of Oklahoma P, 1999. 231–250. Print.

"Guidelines for Editors of Scholarly Editions." Modern Language Association, 29 June 2011.

    http://www.mla.org/resources/documents/rep_scholarly/cse_guidelines. 30 Mar. 2013.

Murray, Janet Horowitz. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. New

    York: Free Press, 1997. Print.

"N I N E S." NINES. <http://www.nines.org/>.

"Obituary—Rev. J. Beresford, M.A." *The Gentleman's Magazine*. May 1841. 548. Web.

Ong, Walter. *Orality and Literacy: The Technologizing of the Word*. London: Methuen, 1982.

    Print.

"The Open Greek and Latin Project." Universität Leipzig, Alexander von Humboldt-

    Lehrstuhl für Digital Humanities. <http://www.dh.uni-leipzig.de/wo/projects/open-greek-

    and-latin-project/>.

"Paul the Not-So-Simple." U of Pittsburgh, Department of Slavic Languages and Literatures.

    <http://paul.obdurodon.org/ohrid/>.

Perkell, Christine. "Editor's Introduction." *Reading Vergil's Aeneid: An Interpretive Guide*. Ed.

    Christine Perkell. Norman, OK: U of Oklahoma P, 1999. 3–28. Print.

"Perseus Digital Library." Tufts U. <http://www.perseus.tufts.edu/hopper/>.

 "The Poetess Archive | About: Introduction to Poetess." Texas A&M University.

    <http://idhmc.tamu.edu/poetess/about/index.html>.

PREMIS Data Dictionary for Preservation Metadata: Version 2.2. July 2012. Web.

Price, Kenneth M. "Electronic Scholarly Editions." *A Companion to Digital Literary Studies*.

    Eds. Susan Schreibman and Raymond George Siemens. Malden, MA: Blackwell, 2008.

    Web.

Putnam, Michael C. J. *Virgil's Epic Designs: Ekphrasis in the Aeneid*. New Haven, CT: Yale U

    P, 1998. Print.

Raabe, Wesley Neil. *Harriet Beecher Stowe's Uncle Tom's Cabin: an Electronic Edition of the*

    National Era Version. Diss. U of Virginia, 2006. Web.

Rahtz, Sebastian. "Storage, Retrieval, and Rendering." *Electronic Textual Editing*. Eds. Lou

    Burnard, Katherine O'Brien O'Keeffe, and John Unsworth. New York: The Modern

    Language Association of America, 2006. 309–333. Print.

Renear, Allen H. "The Descriptive/Procedural Distinction is Flawed." *Markup Languages:*

    Theory & Practice 2.4 (2001): 411–420. Web. 11 Mar. 2015.

Renear, Allen H., Elli Mylonas, and David Durand. "Refining our notion of what text really is:

    The problem of overlapping hierarchies." 1993. Web. 10 Mar. 2015.

Renear, Allen H., Christopher Phillippe, Pat Lawton, and David Dubin. "An XML Document

    Corresponds to Which FRBR Group 1 Entity?" *Proceedings of Extreme Markup*

    Languages 2003. Eds. B. Tommie Usdin and Steven R. Newcomb. Montreal: 2003.

Rice, Jeff. *Rhetoric of Cool: Composition Studies and New Media*. Carbondale: Southern Illinois

    U P, 2007. Print.

Rommel, Thomas. "Literary Studies." *A Companion to Digital Humanities*. Eds. Susan

    Schreibman, Raymond George Siemens, and John Unsworth. Malden, MA: Blackwell,

    2004. 88–96. Print.

"Rossetti Archive." U of Virginia, Institute for Advanced Technology in the Humanities.

    <http://www.rossettiarchive.org/>.

Rude, Carolyn D., and Angela Eaton. *Technical Editing*. Fifth Edition. Boston: Longman, 2011.

    Print.

De Saussure, Ferdinand. *Course in General Linguistics*. Trans. Wade Baskin. Eds. Charles Bally

    and Albert Sechehaye. New York: McGraw-Hill, 1966. Print.

Smith, Martha Nell. "Electronic Scholarly Editing." *A Companion to Digital Humanities*. Eds.

    Susan Schreibman, Raymond George Siemens, and John Unsworth. Malden, MA:

    Blackwell, 2004. 306–322. Print.

Sperberg-McQueen, C. Michael. "Textual Criticism and the Text Encoding Initiative." *The*

    Literary Text in the Digital Age. Ed. Richard J. Finneran. Ann Arbor: U of Michigan P,

    1996. 37–61. Print.

Storey, H. Wayne, and John A. Walsh. "Petrarchive: An edition of Petrarch's songbook." Indiana

    U, 2013. <http://dcl.slis.indiana.edu/petrarchive/>.

 "TEI: Text Encoding Initiative." Text Encoding Initiative, 12 Mar. 2013. <http://www.tei-

    c.org/index.xml>.

"TLG – Home." U of California-Irvine. <http://stephanus.tlg.uci.edu/>.

Ulmer, Gregory L. *Electronic Monuments*. Minneapolis: U of Minnesota P, 2005. Print.

Underberg, Natalie, and Rudy McDaniel. "Using the extensible markup language in cultural

    analysis and presentation." *Digital Ethnography: Anthropology, Narrative, and New*

    Media. Eds. Natalie Underberg and Elayne Zorn. Austin: U of Texas P, 2013. 48–65.

"Using Dublin Core." Dublin Core Metadata Initiative, 2005.

    <http://dublincore.org/documents/usageguide/>. 14 Feb. 2012.

Vergil. *The Aeneid*. Trans. Sarah Ruden. New Haven, CT: Yale U P, 2008. Print.

Virgil. *The Aeneid*. Trans Robert Fagles. New York: Penguin Books, 2006. Print.

Virgil (Publius Vergilius Maro). *Aeneis*. The Latin Library.

Wagner, William P., Vik Pant, and Ralph Hilken. "Adding XML to the MIS Curriculum:

    Lessons from the Classroom." *Journal of Information Technology Education* 7 (2008):

    35–45. Web. 22 Sep. 2011.

Walsh, John. "TEI Boilerplate." Indiana U. <http://dcl.ils.indiana.edu/teibp/>.

"The Walt Whitman Archive." U of Nebraska-Lincoln, Center for Digital Research in the

    Humanities. <http://www.whitmanarchive.org/>.

Welty, Christopher, and Nancy Ide. "Using the Right Tools: Enhancing Retrieval from Marked-

    up Documents." *Computers and the Humanities* 33.1–2 (1999): 59–84. Web. 8 Jul. 2012.

Wittern, Christian. *Summer Seminar 2005: The World of XML Markup*. Diss. Kyoto U, 2005.

"WWP." Northeastern University. <http://www.wwp.northeastern.edu/>.

"XTF." U of California-Berkeley, California Digital Library. <http://xtf.cdlib.org/>.