Models of Digital Documentation:

The 19th-Century Concord Digital Archive

Amy E. Earhart

I wish I could write that I recognized the possibilities of digital scholarship immediately and, with my enlightenment, proceeded to create a project that captured the potential of such scholarship. Instead, the journey to my current digital work has been halting and slow, with many moments of confusion along the way. My mantra, during my early work, was taken from John Unsworth: "If an electronic scholarly project can't fail and doesn't produce new ignorance, then it isn't worth a damn."[1] Ultimately, digital scholarship is in its infancy and digital practitioners are largely self-trained. Missteps and failures necessarily come with experimentation. And, the primary objective of digital work, in my opinion, should be experimentation. The work of digital scholarship is not only about production of the final product, but production of the theoretical and methodological approaches to the digital that we have only just begun to explore. The value of such work is not to be underestimated. Jerome McGann has famously predicted that in "the next fifty years the entirety of our inherited archive of cultural works will have to be reedited within a network of digital storage, access, and dissemination."[2] As our cultural heritage is being digitized at an increasingly rapid rate we are experiencing greater access to materials, but we are also confronted with new problems of use. Scholars will want digital materials to meet our particularized needs. For example, Geoffrey Nunberg has recently described the many problems connected to search capability that stifles scholarly work within Google Books. For the average user, Nunberg notes, Google-based searching is useful, but for the type of work that scholars imagine, "The metadata simply aren't up to it."[3] As

[1] John Unsworth, "Documenting the Reinvention of Text: The Importance of Failure," *Journal of Electronic Publishing* 3 (1997). http://dx.doi.org/10.3998/3336451.0003.201

[2] Jerome McGann, "A Note on the Current State of Humanities Scholarship," *Critical Inquiry* 30 (2004): 410.

[3] Geoffrey Nunberg, "Google's Book Search: A Disaster for Scholars — The Chronicle Review — The Chronicle of Higher Education," *The Chronicle of Higher Education: The Chronicle Review*, 31 August 2009.

Nunberg suggests, scholars must step up and participate in the debate about digital materials if we want to engage with our cultural resources.

As with many digital archives, *The 19th-Century Digital Concord Archive* (*CDA*) started as a website utilizing simple technology and has evolved to a more technologically advanced scholarly site. The *CDA* joins an interdisciplinary team from the Department of English, Texas A&M University; the Digital Humanities Initiative, the College of Liberal Arts, Texas A&M University; the Map and GIS Collections and Services, Texas A&M University Libraries; and the Concord Free Public Library, Concord, Massachusetts, in the development of infrastructures that allow the entities to share metadata easily, develop innovative, visually-based search functions, and make visible and accessible the cultural record of Concord, Massachusetts, in an interactive, free-access digital archive. This project leverages resources and skills across the team to develop a model of interaction between academic, museum-and-library, and community partners, developing multiple ways of displaying information about the town of Concord that will encourage innovative scholarly research. Materials slated for inclusion in the archive include literary texts, historical documents, maps, photographs, census materials, educational minutes, broadsides, physical artifacts, and town records. Concord figures centrally in critical discussions of nineteenth-century literature, philosophy, abolition, women's literature and history, architecture, and government. Scholarly production reflects the importance of this location. Currently, WorldCat lists over 500 books published since 2000 that include Concord in their description. When the search is expanded to include figures that lived or worked in Concord, the numbers grow exponentially. Concord is also an interesting test case for this work as it is a location that helped to define the critical framework of American literature and history. The depth of this small town's historical record proves important to the study of literature, history, government, architecture, philosophy, digital humanities and other fields. By digitizing a broad range of materials we will provide scholars with additional materials to rethink the way in which we conceptualize these fields. Concord is an interesting choice for a digital archive as it bridges the divide between canonical, well-studied figures and unknown figures that flesh out the historical and literary record. Ralph Waldo Emerson, Nathaniel Hawthorne, Louisa May Alcott, Henry David Thoreau, and Bronson Alcott resided in Concord and interacted with those groups less frequently represented by digital archives: free African-Americans, Irish immigrants, the poor, and the criminal class. In addition to the tremendous scholarly interest in Concord, Concord attracts broad general interest as an historical tourism center. The booming tourism trade attracts tens of thousands of visitors a year, many of whom explore Concord virtually before their visit. Given the interest in nineteenth-century Concord, the Concord archive should experience tremendous use and generate a substantive impact.

When I first began to sketch out what the *CDA* might become, I was a lecturer. While the position provided a low wage and high teaching load with little chance of advancement, it also allowed the freedom to experiment with a project that might have no measurable value in a tenure decision, yet interested me immensely and had, I thought, real scholarly value. During the ensuing years I effectively retrained myself to work with digital scholarship, something that would have been nearly impossible to do under the pressures of the tenure track. I found little infrastructure to support digital work on my campus, so I went to the experts. I attended a TEI/XML[4] workshop at Brown University given by Julia Flanders and Syd Bauman and the first NINES (networked infrastructure for nineteenth-century electronic scholarship) summer workshop, where I learned much from Jerome McGann, Bethany Nowviskie, Laura Mandell, and a small but dedicated group of scholars working on digital archives. I contacted Ken Price, co-founder of the Whitman Archive and a former professor of mine, to ask for advice. I was lucky that these pioneers were generous to a scholar interested in the field and were available for help and support. My story is not unique. Digital projects are often created by scholars outside the traditional academic power structure who believe strongly in the importance of such work or, at the other extreme, leaders in the field who have used their endowed chairs and full professorships to help alter attitudes toward digital work. If you decide to take on a digital project, people and organizations are there to help. Structures are changing. Universities are putting support for digital work into place, new organizations, such as NINES, are emerging, and digital humanities centers are being created to support the digital work that you imagine. But, a scholar interested in digital work needs to be realistic about how current digital work is valued by the academy. Some changes to tenure and promotion criteria are occurring, but many departments are slow to respond. While groups like the MLA Task Force on Evaluating Scholarship have called for development of "a system of evaluation for collaborative work that is appropriate to research in the humanities and that resolves questions of credit in our discipline as in others," the same task force found that "60% of departments in Carnegie doctorate institutions say referred articles in digital format either 'don't count' for tenure in their departments and institutions or that they have no experience evaluating them."[5] Imagine, then, these departments' response to non-referred online digital scholarship. I say this not to discourage work within this field, but to caution you to be realistic and plan accordingly.

[4] TEI/XML markup of texts is the *de facto* international standard for encoding texts in the humanities.

[5] MLA Task Force, "Selected Findings from the MLA's 2005 Survey of Tenure and Promotion," 2005. http://www.mla.org/pdf/taskforcereportppt.pdf

I was trained at Texas A&M University during the height of the New Historicist movement. My dissertation project mapped shifting constructions of race in nineteenth-century Boston, what I designated the "Architecture of Inequality," through historical texts, literature, physical structures, and landscape. However, the print monograph was not particularly conducive to the type of exploration that I imagined. It was difficult to represent shifting constructions of race, architecture, and literature through a static form of scholarship, which limited the representational possibilities of this fascinating set of documents. As I began to experiment with digital environments I thought that the computer might aid such scholarship as I had imagined. Concord was a smaller environment than Boston in which I could experiment with mapping the texts that I was interested in exploring. And, for a person trained to gather historical texts to set in play with literary texts, the modeled digital archive allowed me not only to position these materials in a print analysis, but to do so in a visual manner. Unlike a digital edition, my archive would not include multiple versions of a text. Instead, it would bring the literary text into direct dialogue with the historical text, the image, the map, the census documents, crime statistics, and more. I began to shape an infrastructure to represent my theoretical take on the Concord landscape, choosing to work with digital maps and GIS data. My choice of infrastructure was a choice made to represent the theoretical underpinnings of my literary scholarship, but the challenge was to bringing the technology into a working relationship with literary theory. As my example demonstrates, those who choose to work with a digital project must give careful thought to matching the archive structure—the selection and arrangement of materials, metadata, and interface—to the theoretical goals of the scholar. An infrastructure that is matched to a scholarly edition project is probably not the right infrastructure for a project that seeks to represent shifts of iconography over time. Ultimately, the archive structure needs to allow the scholar to complete the type of work imagined possible.

The *CDA* interface is designed to address my concern that, while many scholars who work with Concord discuss the importance of location, physical structures and landscape, there remains a limitation in the way that textual materials might be explored through traditional print scholarship. Unfortunately, digital archives have historically replicated much of the print book structure, from presentation of text to user interface. One of the goals of the *CDA* is to reimagine the book-based interface (index, table of contents) in a digital environment. Too often the digital archive is merely a digital repository of a broad number of texts, rather than a carefully constructed set of interpretive data. Given the importance of the location, geography and landscape of Concord, a visual means of addressing the humanities information allows for interesting possibilities and should provide new ways of researching the related areas. Our team has developed initial, simple maps that represent the town site, and we are currently developing ad-

vanced map interfaces that visualize the town over time. Using Google Earth, historical and contemporary maps as well as digitized town reports, census, and literary materials, we are hoping to develop a map and connected timeline that allows users to manipulate time and place as well as sift the materials to locate textual data.

Another important issue that the Concord Digital Archive seeks to explore through its interface structure is the way that transnationalism plays out within the particular literary and historical moments of the town. Current work on the *CDA* suggests that the African and Irish Diasporas reveal themselves in town materials and that interactions between these groups impact the literary production of Concord writers and vice versa. Rather than focusing on the few authors that lived in Concord for most of their lives, the *CDA* materials invite the scholar to see those who immigrate, who traverse national boundaries, and who look outward, out of Concord, Massachusetts and the United States to a broader world. The mapping segment of the project is currently being built to show patterns of movement in Concord by Irish- and African-Americans and the response of Anglo-Concordians to both groups by digitizing place of residence, nationality, race, and socioeconomic factors over time. In other words, while the Concord project does indeed look to one particular element of literary history that has been interpreted as "American," the materials found within the archive challenge this simplistic reading.

While digital archives offer the scholar a chance to produce groundbreaking research, there remain structural difficulties in the creation of such scholarship. Digital work is often immeasurably slow to produce, so glacial, in fact, that those working within the field often speak of their never-ending projects. If you wish to publish a book, there is a long history of process in place. In addition, a print project has boundaries that are fairly rigid. Presses limit page numbers, contracts limit time to finished product, print publication is finished and a bound book produced. Not so with the digital. Changing technology, the unbounded length of a project, changes in copyright law, and more can create issues with completion. A spring 2009 *DHQ: Digital Humanities Quarterly* volume addresses the difficulty of demarcating production boundaries within digital projects from a variety of perspectives. Matthew Kirschenbaum asks in his introduction, "What is the measure of 'completeness' in a medium where the prevailing wisdom is to celebrate the incomplete, the open-ended, and the extensible?"[6] Or, as Susan Brown et al. state of their project, "the Orlando Project, a large-scale and long-standing digital humanities undertaking, reveals an arbitrariness, even a fictive-

[6] Matthew Kirschenbaum, "Done: Finishing Projects in the Digital Humanities," *DHQ: Digital Humanities Quarterly* 3 (September 2009). http://digitalhumanities.org/dhq/vol/3/2/000037.html

ness or contradictoriness, to the notion of completion of the project as a whole or even of its major online product."[7] Those interested in the creation of digital work should spend time considering how they might structure their projects in response to the inevitable open-endedness of the digital, whether modeling stages or projects within a larger project, housing their digital work within the more traditional structure of a digital press, such as the University of Virginia's Rotunda Press, or declaring a project ever expanding and rejecting the closure of a traditional scholarly project.

The struggle to define a project within the fluid environs of the digital can be rewarding. I have a far greater understanding of how I define digital humanities and, through the interactive laboratory of the *Concord Digital Archive,* have learned invaluable lessons about production and use of digital materials. As Jerome McGann has repeatedly argued, you must build the archive to learn what you need to know: "Translating paper-based texts into electronic forms entirely alters one's view of the original materials."[8] I have learned that basic notions about what I am creating are actually highly contested. I initially titled my project an archive, believing that it would be a digital repository of materials that functioned much like a physical archive. Out of all the initial decisions I made for the *CDA*, I thought titling the set of materials "archive" would be one of the easiest. When I began the project, the archive form was the norm for digital literary scholarship production, suggesting that Jacques Derrida might have been on to something in his analysis of archive fever. However, recent work on the archive, from digitally engaged analysis, such as work in *DHQ* by Margaret Ezell and Ken Price, to more generic analysis, by scholars including Antoinette Burton, *Archive Stories*, and Carolyn Steedman, *Dust: The Archive and Cultural History,* challenges us to rethink a simple conception of the digital archive.[9] In my own work several crucial factors have exploded the concept of archive, yet I continue to use the original title of the project as the best descriptive term to date. An archive is a repository, yet my archive actually does not reposit most of the materials it represents, even in the form of a surrogate. Instead, the archive acts as a search-and-

[7] Susan Brown et al., "Published Yet Never Done: The Tension Between Projection and Completion in Digital Humanities Research," *ibid.* http://digitalhumanities.org/dhq/vol/3/2/000040.html

[8] Jerome McGann, "Imagining What You Don't Know: The Theoretical Goals of the Rossetti Archive," *Voice, Text, Hypertext: Emerging Practices in Textual Studies* (Seattle: University of Washington Press, 2004), p. 383.

[9] See Margaret J. M. Ezell, "Paraplex: Or, the Functions of Angels in the Archives," *DHQ:Digital Humanities Quarterly*, forthcoming. Kenneth M. Price, "Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name?" *DHQ: Digital Humanities Quarterly*, 3.3, 2009. Antoinette Burton, "Introduction: Archive Fever, Archive Stories," *Archive Stories: Facts, Fictions and the Writing of History* (Durham: Duke University Press, 2005), pp. 1–24. Carolyn Steedman, *Dust: The Archive and Cultural History* New Brunswick: Rutgers University Press, 2002.

display space for materials located on disparate servers. In some ways, then, the archive is a display catalog, rather than a rare-books room. And, how does one understand the physical archive that underpins the *CDA*, the Concord Free Public Library's Special Collections? The Concord Free Public Library has an extensive physical archive collection from which it chooses selected items for digitization. Out of that set the *CDA* chooses materials of which to provide metadata and to connect to our search function. Does this make the *CDA* a collection? A presentation? Archive as a term is far more contested within the malleable digital environment.

Digital scholarship is tough and challenging, but the most rewarding scholarly work I have undertaken. However, you must go into the project with your eyes wide open. What infrastructure is available at your institution? Could you develop links to other digital humanities projects or resources, such as NINES or the TEI consortium? Can you standardize your materials to best take advantage of work that is already completed in digital humanities? There is a growing body of best practices for scholars interested in creating digital materials and groups to support the work. Generally, those working in the field are helpful and generous, supporting scholars in this new and exciting work. I encourage you to think about contributing to the field, and I invite you to visit and explore *The 19th-Century Concord Digital Archive* at http://www.digitalconcord.org.