PART IV][Chapter 18

()

Can Information Be Unfettered?

Race and the New Digital Humanities Canon

()

Amy E. Earhart

n the 1990s, the rallying cry of proponents of the Internet was the democratization of knowledge made possible by the developing technological infrastructure. Lost or excluded texts began to be published on the net, some developed by scholars, others by fans, and still others by libraries and museums. I remember the possibilities that these materials offered for the literary scholar. I could create a website for students that linked the recovered e-text of Harriet Wilson's Our Nig, period images of slaves, and the variety of African American cultural and historical documents found on the then-fledgling Schomburg Research Center website. The seemingly expansive materials for use on the web were far more complete than materials found in print anthologies or other such course materials. For scholars interested in reinserting writers of color into critical discussions, the recovery efforts were a boon. We imagined that the free access to materials on the web would allow those previously cut off from intellectual capital to gain materials and knowledge that might be leveraged to change the social position of people of color. The new space of the Internet would allow those who had been silenced to have a voice. Hypertext theorist Jay David Bolter promoted the freeing power of the web-based environment as a space that encouraged "the abandonment of the ideal of high culture (literature, music, the fine arts) as a unifying force. If there is no single culture, but only a network of interest groups, then there is no single favored literature or music" (233). As the 1990s drew to a close, and the number of digitally recovered texts seemed to grow each day, Bolter's prediction seemed correct. However, a review of digitized materials production and the current treatment of race in the digital canon suggests that Bolter's hopes have not been realized.

I want to focus my discussion by examining a subset of the digital humanities, digital texts.¹ I'm interested in the digital work being produced by those associated with academia and those with strong connections to traditional humanities fields including history, literature, classics, art history, and archeology, among others. My

()

[309

()

focus includes pay-walled scholarly production, such as the excellent *Clotel* project published by Virginia's Rotunda Press, and open-access materials but excludes large-scale digital projects produced by for-profit publishers, such as Gale-Cengage, or nonscholarly produced projects, such as Google Books. I am also most interested in projects that make something. Here I would like to echo Stephen Ramsay's recent argument that "Digital Humanities is about building things." While Ramsay has come under fire for his insistence on the applied nature of digital humanities, the history of digital humanities reveals the centrality of building to the field. In fact, scholars invested in early work on race in digital humanities insisted on building editions and digital texts as an activist intervention in the closed canon. While we should continue to explore tool building, visualization, and data mining as crucial areas within digital humanities, the narrow digital canon should remind us why we cannot stop digital edition work.

()

While those invested in digital text production should continue to flesh out the digital canon, other areas of digital humanities, such as tool building and visualization, should also be invested in the investigation of canon on their work. For example, the Metadata Offer New Knowledge (MONK) project has harnessed materials from Documenting the American South, Early American Fiction, Early English Books Online (EEBO), Eighteenth Century Collections Online (ECCO), Nineteenth-Century Fiction, Shakespeare, and Wright American Fiction 1850-75. While the purpose of MONK is not text recovery but visual analysis, a broad understanding of the literature of this period is only as good as the data from which the analysis draws from. In the case of MONK, a quick search reveals that texts by Sojourner Truth, Sui Sin Far, and Maria Christina Mena-authors of color included in most standard anthologies of American literature-are absent.² Add to this MONK's claim that "For users of public domain materials, MONK provides quite good coverage of 19th century American fiction," and we are reminded that a more direct analysis of the position of race in digital humanities work is necessary (Unsworth and Muller, 2). We shouldn't be surprised at the lack of certain texts used in the MONK project, as the digitized humanities corpora is scant and the project's primary goal was to develop a prototype of data-mining work rather than an inclusive data set. However, the emphasis on "good coverage" should concern digital humanists. As a field where collaborative teams are able to produce better results than that of the lone scholar, it is important to include project participants that can help speak to the importance of cultural criticism. Without careful and systematic analysis of our digital canons, we not only reproduce antiquated understandings of the canon, but we reify them through our technological imprimatur.

Unlike related fields, digital humanities has historically deemphasized theoretical examination of the digital utilizing cultural studies frameworks. Those working within rhetoric, media, and communication, and particularly those working in game studies, have constructed a body of scholarly work that interrogates the theoretical implications of race construction within technology.³ With superb work being

Can Information Be Unfettered? [311

produced by scholars such as Lisa Nakamura, Beth Kolko, and Tara McPherson, these fields have begun the difficult work of theorizing the way in which technology impacts the digital object. In digital humanities, however, we have much theoretical work to do in the selection of materials and application of digital tools to them.

(�)

To understand the current position of race and digital humanities work, we must turn to the emergence of the World Wide Web. As the web began to gain popularity in the 1990s, it was portrayed as an idealized, democratic, and free space. *Time Magazine*'s 1994 story "The Battle for the Soul of the Internet" is indicative of the understanding of the newly popular space, uncolonized and free. The original users of the web, scientists, computer geeks, and hackers, according to the story, were battling against the corporate market intent on invading their open space. Advocates of the free web were interested in three ideas: "1) Access to computers should be unlimited and total; 2) All information should be free; 3) Mistrust authority and promote decentralization," all designed to allow "bubbles" of information to rise from the bottom, sowing "seeds of revolutionary change" ("*Battle for the Soul of the Internet*"). Scholars, too, began to see the net as a space that altered power structures. As Paul Delany notes,

The Internet has thus mutated into an unforeseen and unplanned information space. Its virtues can all be attributed to its collegial political economy: in a word, its openness. Internet's most important features are its relatively small hardware investment, a weak (but not ineffective) central administration, little censorship, and an absence of specifiable "bottom-line" objectives. Its explosive growth in the last few years confirms the dynamism of a collegial cyberspace culture in which millions of users exchange information, collaborate on creative projects, and have their say on any subject they care about. (Childers and Delaney)

()

The revolutionary power of the net was based on the belief that the open digital environment was unpoliced and unregulated, open to all who wanted to participate. The low cost, the "small hardwave investment" that Delany points to, was also crucial, as it allowed scholars to produce their materials more cheaply, increasing the types and numbers of texts available. In his 1996 essay "Principles for Electronics Archives, Scholarly Editions, and Tutorials," Peter Shillingsburg concurs it "eventually will cost less to produce and therefore, one assumes, to purchase a compact disk than it costs to produce and purchase Hans Gabler's edition of *Ulysses*" (25). While, in hindsight, we have rejected this view as too simplistic, it was a common refrain in the early, heady days of digital recovery, and the decentralization and shifting power structures, as related by these statements, became part of the narrative and the mythology that in no small part drove the digital literary studies recovery projects that grew in the 1990s.

For scholars interested in reworking the canon, the web seemed an unfettered space that allowed the scholar direct control of what would be digitized and how

()

it would be presented. Susan Fraiman, director of the Orlando Project, lauded the expansive digital environment, remarking that "What is new in the twenty-first century, however, is that now the guest list of history-making women is electronic—and there are always more seats at the table" (143). The belief in the web as a space in which the canon might be broken was likewise espoused by the editors of *Romantic Circles*: "One of the strengths of Web publishing is that it facilitates—even favors— the production of editions of texts and resources of so-called non-canonical authors and works. This is in part a function of the relative simplicity of HTML (and all of the simpler document-type-descriptions of SGML) and of 'workstation publishing' in general when compared to traditional commercial or academic letterpress production and distribution methods" (Fraistat, Jones, and Stahmer). The ease of publication, identified by the editors, was what allowed the broad range of small-scale recovery projects to explode around the web in this early period. The insistence on the web's ability to build new canons, of an applied approach to the digital tied to a theoretical model, is a hallmark of past and current digital humanities work.

()

The importance of building tools for digital work, common to and controversial in contemporary digital humanities, has deep roots in this early recovery work. The applied/theoretical model voiced by digital humanists also fits historically into the work of cultural studies scholars prior to the mainstream Internet. For example, Jean Fagan Yellin's work on Harriet Jacobs' *Incidents in the Life of a Slave Girl* took various trajectories of inquiry, but her 1987 Harvard University Press edition, the edition that brought the important volume to the center of the African American literary tradition, was an applied piece set within her larger scholarly body of work. In recovering the text and conducting the painstaking research required to do so, Yellin has followed a model familiar to those working with digital projects application driven scholarship. This model of building, grown from the cultural studies recovery tradition, would expand as additional scholars began to explore the digital as a tool by which to recover texts by writers of color.

During the 1990s and early 2000s, in many ways the most productive age of digital recovery to date, projects fell into two distinctive scholarly groups. One group was the small-scale project in which scholars worked individually or as small collectives. These projects, including *The Charles Chesnutt Archive, Voices from the Gaps*, and *The Online Archive of Nineteenth-Century U.S. Women's Writings*, were mostly unfunded and produced outside of digital humanities centers or libraries. The second type of project production was produced by e-text centers, digital humanities centers, or libraries and museums. Individual scholarly participation was less central to such work. These projects include the *Emory Women Writers Resource Project*; the various projects produced by the Virginia Center for Digital History, including the *Valley of the Shadow* and *Race and Place: An African-American Community in the Jim Crow South*; and the digital *Schomburg Library of Nineteenth-Century Black Women Writers*.

Can Information Be Unfettered? [313

Of these two types of projects, the dominance of small-scale recovery efforts nurtured by an individual scholar who wanted to bring lost texts to scholarly and public attention is surprising. Simple HTML projects such as *The Charles Chesnutt Archive*, by Stephanie Browner; *Voices from the Gaps*, by Toni McNaron and Carol Miller; and *The Online Archive of Nineteenth-Century U.S. Women's Writings*, by Glynis Carr, were developed without the support of a digital humanities center, technological collaborators, or external funding. Some projects, such as *The Charles Chesnutt Archive*, were positioned as undergraduate teaching and learning tools, with undergraduate student partners in the recovery process.

(�)

Alan Liu's *Voice of the Shuttle* provides a good measure of the huge number of early recovery projects focused on literature and history written by and about people of color. A quick perusal of "The Minority Studies" section, however, reveals that a tremendous number of the projects have become lost. For example, of the six sites listed in "General Resources in Minority Literature," half cannot be located, suggesting that they have been removed or lost. The same trend is found with other projects listed on the site. While only 50 percent of the projects in the "General Resources in Chicano/Latino Literature" section are still online, other areas, such as Asian American literature, have a higher percentage of active projects.⁴ Digital humanists are fond of talking about sustainability as a problem for current and future works, but it is clear that we already have sustained a good deal of loss within the broadly defined digital canon.

()

Cocurrent to the DIY projects were institutional initiatives focused on bringing lost texts to view. Much of this work occurred within e-text centers, such as Rutgers University's Center for Electronic Texts in the Humanities (CETH); fledgling digital humanities centers, such as the University of Virginia's Institute for Advanced Technology in the Humanities (IATH); and museums and libraries, including the New York Public Library. With limited exceptions, a majority of the early projects reinforced canonical bias. Catherine Decker argues that the canon crops up in these projects because of their funding and institutional affiliations: "The reasons for the canonicity of the bulk of the electronic texts available on the web are hardly elusive: most of the large textbases are located at and funded by major universities (or grants to university scholars that include a cut for the university at which the project is situated)." Martha Nell Smith extends this contention and argues that digital humanities developed as a space to which practitioners hoped to flee from the shifts in the profession that arose out of the cultural studies movement. In "The Human Touch: Software of the Highest Order, Revisiting Editing as Interpretation," Smith highlights the digital humanities' retreat into modes of analytics, objective approaches as "safe" alternatives to the messy fluidities found in literary studies. She notes, "It was as if these matters of objective and hard science provided an oasis for folks who did not want to clutter sharp, disciplined, methodical philosophy with considerations of the gender-, race-, and class-determined facts of life . . . Humanities computing

()

seemed to offer a space free from all this messiness and a return to objective questions of representation" (4). If Smith is correct, then we not only have a selection problem in digital humanities, but we have a historical structural problem that might be more difficult to reverse.

()

One only needs to review the current work in digital literary studies to see that we have not escaped the traditional canon by turning to new methods of publication. The proliferation of early projects I have cataloged remain but a trace in the current digital canon. A search of websites referenced by the Modern Language Association (MLA) bibliography reveals almost one thousand individual sites, yet very few of these projects are represented in MLA presentations, leaving many new to the field to assume that there are a small number of digital projects, often focused around a core group of digital humanities practitioners. The perception of limited projects and practitioners is what has driven the recent controversy of the digital humanities star system, highlighted by William Pannapacker's post-MLA 2011 blog post, "Digital Humanities Triumphant?" While I, along with many others, reject Pannapacker's representation of key digital humanities scholars as the cool kids at the high-school lunch table, the perception of the exclusionary world of digital humanities is reinforced by a perception of limited projects. Impacting the perception of digital humanities as exclusive, in both practice and product, is the granting model. Examination of funded projects reveals that the shift toward innovation has focused on technological innovation, not on innovative restructuring of the canon through recovery. The National Endowment of Humanities (NEH) awarded 141Digital Humanities Start-Up Grants from 2007 through 2010. Of those grants, only twenty-nine were focused on diverse communities and sixteen on the preservation or recovery of diverse community texts. It is striking to examine the authors and historical figures individually cited in the list of funding: Shakespeare, Petrach, Melville, Jefferson, David Livingstone,5 and Whitman. While there are grants to support work on indigenous populations, African and African American materials, and Asian American materials, in addition to others, the funding of named great men of history deserve scrutiny and even, perhaps, a specific funding program to encourage recovery efforts (NEH).

There may be many reasons for the lack of attention to noncanonical texts. Margaret Ezell cautions that we have not revised the way in which we understand texts; and because of this elision certain texts, particularly noncanonical texts, are not being digitized. She argues that "while we increasingly have the ability to digitalize any text we please . . . editors do not please to select certain types of material and this is in part because perhaps we are not yet changing some of the basic assumptions about what an 'edition' does, or in Hunter's terms, what is 'appropriate'" ("Editing Early Modern Women's Manuscripts," 107). Additional reasons for exclusion are structural, such as the cost of production. Ken Price has discussed "the strong correlation between the 'significance' on which successful grant writing depends and the traditional canon" (281). Susan Belasco, for example, believes "that the traditional

Can Information Be Unfettered? [315

standards for tenure and promotion are, in fact, more entrenched than ever and worse—more restrictive and un-imaginative than they were for an earlier generation" (333). Or, as Martha Nell Smith has argued, the digital humanities community might be adverse to the expansion of the canon and the work that has been reinserted into the mix. All of these possible explanations deserve critical attention if the digital humanities community wants to promote a broader digital canon.

While a good many of the early small-scale digital projects have been displaced or lost from our current digital canon, a few have managed not only to survive but to thrive. 19: Interdisciplinary Studies in the Nineteenth Century is one such project. Begun as a simple HTML journal, scholars affiliated with the project participated in a Nineteenth-Century Scholarship Online (NINES) summer workshop during which they learned to encode with the international standard of TEI/XML. Once the project was re-marked with TEI, it was brought into the NINES federated collection of nineteenth-century materials, helping expand its user base and take an important step toward long-term sustainability. It provides a positive example of how we might take the institutional structures that have developed in relation to the digital humanities canon and leverage them to support small-scale projects. 19: Interdisciplinary Studies in the Nineteenth Century also provides a pivotal clue that might explain why certain projects are currently excluded from the digital canon and others are not. The archive demonstrates that project value is created by editorial principles, content, and technological infrastructure. Projects like 19 have been revitalized by alignment with an institution, whether a collective, like NINES, a digital humanities center, like the Maryland Institute for Technology in the Humanities (MITH), or a library, like the University of Nebraska. The case of 19 suggests that standards and institution have become a core part of project success and sustainability, crucial to the canonization of digital work. Ken Price alludes to the new canon criteria when he argues that "people ready to embrace high quality work wherever it is found hold in highest regard digital work that features a rigorous editorial process and adheres to international standards (for example, TEI/XML)" (275). If institutional affiliation and technological standards are necessary components to the success of a project, then digital humanists must investigate how we might provide both to DIY scholars. Groups like NINES, which provide workshops and outreach, have modeled potential ways by which to generate such parameters. Additional efforts must follow.

If, indeed, we are beginning to construct a digital canon that weighs content and technological choices equally, then it is crucial for digital humanists to theorize the technological with the same rigor as we theorize the content. Alan Liu has more broadly seen the problem as an absence of cultural criticism, noting that "rarely do we extend the issues involved into the register of society, economics, politics, or culture" into our digital work. If we do not theorize our technological approaches with a mind toward cultural constructions, we will continue to exclude certain materials from digitization. One possible model is found in the partnership of Timothy Powell and the Ojibwe elders, who have created the *Gibagadinamaagoom* archive.

()

Powell has written extensively on the crucial impact of technological application to indigenous cultural materials and argues that current work needs to be revamped so that "digital technology can more accurately and artistically represent the indigenous origins and spiritual story lines of expressive culture on these [the Americas] continents" (Powell and Aitken, 253). Part of Powell's response to cultural inclusion was to construct a partnership that shifted ownership of cultural materials from that of the scholar to the Ojibwe elders, in effect creating a mechanism by which the tribe might control their own cultural materials.

()

My digital project, The Nineteenth-Century Digital Concord Archive, is similarly invested in exploring how to appropriately apply technological standards to shifting constructions of race represented in textual materials. Our current challenge is how we represent varying representations of blackness found in the census in a database. How do we represent, technologically, the identification of the same person as West Indian, Mulatto, or black? Amanda Gailey's recent article, "A Case of Heavy Editing: The Example of Race and Children's Literature in the Gilded Age," reveals the depth of theoretical inquiry that Gailey has invested in applying TEI appropriately to complex texts that bear markers of postbellum racial construction, particularly in the decision to utilize TEI to facilitate searching. Gailey notes that "we will use the <orig> and <reg> combination instead of the <sic> and <cor> combination (meaning sic and "correction"), as the former pair makes no claim about the rightness or wrongness of the readings, only how standardized their spellings are" (136). The TEI tag selection is in keeping with current cultural criticism regarding race and language, which rejects the superiority of standard English. The choice of <orig> and <reg> reveals that Gailey refuses to value one language usage over the other, as would be implied through the choice of the <sic> and <cor> tags. These examples provide a helpful way to imagine the next steps involved in digital humanities work and the treatment of race. While we need to continue to consider how to invigorate a robust digital recovery, we also have a good bit of theoretical work to do in the selection, editing, and technological manipulation of our materials.

One of the powerful things about the early period of digital literary studies is the DIY approach that many scholars embraced, the sheer joy and freedom of bringing important texts to the larger scholarly community. As we move from simple HTML sites to TEI and visualization projects, as we move from individual or small collective projects to larger team projects, from nonbudgeted projects to large, externally funded projects, we see fewer scholars working with digital textual recovery. This should concern digital humanists, and we should accordingly begin to strategize how we might reverse this trend. Small steps are under way. We need to examine the canon that we, as digital humanists, are constructing, a canon that skews toward traditional texts and excludes crucial work by women, people of color, and the GLBTQ community. We need to reinvigorate the spirit of previous scholars who believed that textual recovery was crucial to their work, who saw the digital as a way to enact changes in the canon. If, as Jerome McGann suggests, "the entirety of our cultural

inheritance will be transformed and reedited in digital forms" (72), then we must ensure that our representation of culture does not exclude work by people of color.

Bibliography

()

- "Battle for the Soul of the Internet." *Time Magazine*, July 25, 1994. http://www.time.com/time/magazine/article/0,9171,981132,00.html.
- Belasco, Susan. "The Responsibility Is Ours: The Failure of Infrastructure and the Limits of Scholarship." *Legacy: A Journal of American Women Writers* 26, no. 2 (2009): 329–36.
- Bobley, Brett. "New from ODH: Summary Findings of NEH Digital Humanities Start-Up Grants (2007–2010)." National Endowment for the Humanities, Office of Digital Humanities, September 7, 2010. http://www.neh.gov/ODH/ODHUpdate/tabid/108/ EntryId/144/New-from-ODH-Summary-Findings-of-NEH-Digital-Humanities -Start-Up-Grants-2007-2010.aspx.
- Bolter, J. D. *Writing Space: The Computer, Hypertext, and the History of Writing*. Boston: Houghton Mifflin, 1991.
- Childers, Peter, and Paul Delany. "Introduction: Two Versions of Cyberspace." *Works and Days* 23, no. 4 (1994): 61–78. http://www.sfu.ca/delany/wkndays.htm.
- Decker, Catherine. "Crossing Old Barriers: The WorldWideWeb, Academia, and the Romantic Novel." *Romanticism on the Net* 10 (1998). http://www.erudit.org/revue/ron/1998/v/n10/005794ar.html.
- Ezell, Margaret J. M. "Editing Early Modern Women's Manuscripts: Theory, Electronic Editions, and the Accidental Copy-Text." *Literature Compass* 7, no. 2 (2010): 102–9. http://onlinelibrary.wiley.com/doi/10.1111/j.1741-4113.2009.00682.x/full.
- - ——. Writing Women's Literary History. Baltimore: Johns Hopkins University Press, 1993.
- Fraiman, Susan. "In Search of Our Mother's Gardens–With Help from a New Digital Resource for Literary Scholars." *Modern Philology* 106, no. 1 (2008): 142–48.
- Fraistat, Neil, Steven E. Jones, and Carl Stahmer. "The Canon, the Web, and the Digitization of Romanticism." *Romanticism on the Net* 10 (1998). http://www.erudit.org/ revue/ron/1998/v/n10/005801ar.html
- Gailey, Amanda. "A Case for Heavy Editing: The Example of Race and Children's Literature in the Gilded Age." In *The American Literature Scholar in the Digital Age*, edited by Amy E. Earhart and Andrew Jewell, 125–44. Ann Arbor: University of Michigan Press, 2010.
- Kirschenbaum, Matthew. "What Is Digital Humanities and What?s It Doing In English Departments?" *ADE Bulletin* 150 (2010): 55–61. http://mkirschenbaum.files.word press.com/2011/01/kirschenbaum_ade150.pdf Reprinted in this volume.
- Liu, Alan. "Where Is Cultural Criticism in the Digital Humanities." *Alan Liu*. January 2011. http://liu.english.ucsb.edu/where-is-cultural-criticism-in-the-digital-humanities/.

McGann, Jerome. "Culture and Technology: The Way We Live Now, What Is To Be Done?" *New Literary History* 36, no. 1 (2005): 71–82. http://muse.jhu.edu/journals/new_literary _history/v036/36.1mcgann.html.

- National Endowment for the Humanities. "Digital Humanities Start-up Grants." August 6, 2010. http://www.neh.gov/grants/guidelines/digitalhumanitiesstartup.html.
- Omeka. Roy Rosenzweig Center for History and New Media. George Mason University. http://omeka.org/
- Powell, Timothy B., and Larry P. Aitken. "Encoding Culture: Building a Digital Archive Based on Traditional Ojibwe Teachings." In *The American Literature Scholar in the Digital Age*, edited by Amy E. Earhart and Andrew Jewell, 250–74. Ann Arbor: University of Michigan Press, 2010.
- Price, Ken. "Digital Scholarship, Economics, and the American Literary Canon." *Literature Compass* 6, no. 2 (2009): 274–90. http://onlinelibrary.wiley.com/doi/10.1111/j.1741-4113.2009.00622.x/full.
- Ramsay, Stephen. "Who's in and Who's Out." Stephen Ramsay Blog. January 8, 2011. http://lenz.unl.edu/papers/2011/01/08/whos-in-and-whos-out.html.
- Shillingsburg, Peter L. "Principles for Electronic Archives, Scholarly Editions, and Tutorials." In *The Literary Text in the Digital Age*, edited by Richard J. Finneran, 23–35. Ann Arbor: University of Michigan Press, 1996.
- Smith, Martha Nell. "The Human Touch: Software of the Highest Order: Revisiting Editing as Interpretation." *Textual Cultures* 2, no. 1 (2007): 1–15.

Spencer, Amy. DIY: The Rise of Lo-FI Culture. London: Marion Boyars, 2005.

Unsworth, John, and Martin Muller. *The MONK Project Final Report*. 2009. http://www .monkproject.org/MONKProjectFinalReport.pdf.

Notes

()

1. I use the term "digital text" to include digital edition and digital text. I see these as two distinctive types of textual production. I define a digital edition as a project that emphasizes textual variants using traditional bibliographical methods. A digital text is one version of a text that has been brought digital.

2. For example, all of these authors are included in *The Bedford Anthology of American Literature and The Heath Anthology of American Literature*.

3. I recognize that there is an ongoing discussion of how to view the distinction between media studies and digital humanities. My point is that the larger body of digital humanities work has been less concerned with cultural issues than the work produced by those who self-identify as media, rhetoric, communication, and film studies scholars.

4. I utilized a general Google search to locate the projects. If they were not found through this search, I located the institution or scholar associated with the materials to see if materials had been moved elsewhere.

5. David Livingstone is a Victorian explorer of Africa. The grant was awarded to create an online scholarly edition of his Nwangwe field diary.