4

Archive
Katherine D. Harris

Digital, electronic, and hypertextual archives have come to represent online and virtual environments that transcend the traditional repository of material artifacts typically housed in a library or other institution (Price "Electronic" para. 3). Physical archives claim to amass anything that gives evidence of a time that has passed and "is essential for a civilized community" (*OED*). Traditional institutions define an archive as a rare book library, emphasizing the collecting of codex and manuscript representations of writing, authorship, and history. Most rare book and manuscript divisions also collect, preserve, and archive nontraditional forms of printed material and unconventional literary objects. This type of institutional archive is guided by principles of preserving history and an assumption that a *complete* collection will reveal not only that moment but also its beginning, ending, and connection to other moments. Voss and Werner articulate the duality of the archive as both physical space and, now, an imaginative site, both of which are governed by ideological imperatives (i).

Since approximately 1996, the *digital* archive has revised the traditional institutional archive to represent both a democratizing endeavor and a scholarly enterprise (Manoff 9). An archive, if truly liberal in its collecting, represents an attempt to preserve and record multiple metanarratives (Voss and Werner). Curators and special collection directors become editors and architects of digital archives to produce "an amassing of material and a shaping of it" (Price "Electronic" para. 9). However, the digital archive's instability threatens these metanarratives because of its potential for endless accumulation—a contamination.

In *Archive Fever*, Derrida suggests that the moments of archivization are infinite throughout the life of the artifact: "The archivization

produces as much as it records the event" (17). Archiving occurs at the moment that the previous representation is overwritten by a new "saved" document. Traces of the old document exist but cannot be differentiated from the new. At the moment an archivist sits down to actively preserve and store and catalog the objects, the archive is once again contaminated with a process. This, according to Derrida, "produces more archive, and that is why the archive is never closed. It opens out of the future" (68). Literary works become archives not only in their bibliographic and linguistic codes,[1] but also in their social interactions yet to occur. It is the reengagement with the work that adds to an archive and that continues the archiving itself beyond the physical object.

My keyword essay follows a burning desire central to the archive—to return to the origins intermixed with the desire to hold everything at once in the mind's eye. In literature, this of course causes the protagonist to faint, go mad, isolate herself, create alternate realities—all in the name of either escaping or explaining what cannot be known. For example, my Gothic Novel students pointed out that the narrator in a short story, most specifically Lovecraft's, attempts to focus on a few actions in the busy-ness of the world, to focus the reader on what is calculable, knowable, but ultimately *unheimlich*, or the Freudian sense of uncanny.

In much the same way as a narrator, in the digital age we attempt to create archives of a particular moment or medium (The September 11 Digital Archive <http://911digitalarchive.org/>), the entirety of a medium (The Internet Archive <https://archive .org//>), the mutability of language (*The Oxford English Dictionary*), all knowledge (Wikipedia). More than others, the crowdsourced information of Wikipedia attempts to capture knowledge as well as the creation of that knowledge—the history or Talk of each Wikipedia entry unveils an evolving community of supposedly disinterested[2] users who argue, contribute, and create each entry. Wikipedia entries represent that digital version of an archive in the twenty-first century: the archive as a fractured, incalculable moment attempting to hold close all that happens at once in the world. But this concept has become incredibly problematic with the rush of information around us (Harris, "Archive," *The Johns Hopkins Guide to Digital Media*).

Kenneth Price begins my discussion about *archive* by offering a traditional definition of the term (see also surrogate):

Traditionally, an archive has referred to a repository holding material artifacts rather than digital surrogates. An archive in this traditional sense may well be described in finding aids but its materials are rarely, if ever, meticulously edited and annotated as a whole. In an electronic environment, *archive* has gradually come to mean a purposeful collection of digital surrogates. ("Electronic" para. 3)

Later in this article, Price veers into discussing the role of archivist in shaping the archive; his description, though less dramatic, resembles Derrida's. Price's article is in response to the authority of a digital scholarly edition and its editors in the face of traditional print editions. Always, for Price, there is an organizing principle to archiving and, subsequently, editing. However, what we're concerned with in this keyword is inherently the messiness of the archive as it pertains to cultural records, both physical and digital. What gets placed into the archive and by whom becomes part of that record. What's missing, then, becomes equally important. Martha Nell Smith proposes that digital archives are free from the constraints of a traditional print critical edition; more importantly, the contents and architecture of a digital archive can be developed in full view of the public with the intention of incorporating the messiness of humanity.

In "Googling the Victorians," Patrick Leary describes all sorts of digital archives about Victorian literature that were springing up—archives that are not peer-reviewed per se but offer an intriguing and sanguine view of the wealth of nineteenth-century materials. Leary concludes his essay by asserting that whatever does not end up in a digital archive, represented as cyber- and hypertext, will not, in the future, be studied, remembered, valorized, and canonized. Though this statement reflects some hysteria about the loss of the print book, it is also revealing in its recognition that digital representations have become common and widespread, regardless of professional standards. Whatever is not on the web will not be remembered, says Leary (see also memory). Does this mean that the

literary canon will shift to accommodate all of those wild archives and editions? Or does it mean that those mega projects of canonical authors will survive while the disenfranchised and noncanonical literary materials will fall further into obscurity?

Raymond Williams posits that "vulgar misuse" allows for entry into the cultural record (Williams *Keywords* 21), though those in library science object to a normalization of "archive" that moves away from their professional standards for a vault for the records of humanity. But the construction of a digital archive in literary studies conflates literature, digital humanities, history, computer programming, social sciences, and a host of other cross-pollinated disciplines. The archive, more than anything right now in literary studies, demonstrates what Williams calls "networks of usage" (23) with "an emphasis on historical origins [as well as] on the present—present meanings, implications, relationships—as history" (23). Community, radical change, discontinuity, and conflict are all part of the continuum in the creation of meaning according to Williams, seemingly similar to Borges's "Library of Babel" and Derrida's "archive fever." While archivists insist on a conscious choice in the use of *archive* (noun or verb), perhaps as part of a professional tradition, I seek to look at the messiness of the word as a representation of the messiness of our means for porting records across the past, present, and future.

The issue with formal digital archives is where to stop collecting to account for scope, duration, and shelf space. In digital archives, sustainability is key; but the digital archive is vastly more capable of accumulating everything and then allowing its holdings to be subjected to liberal and even promiscuous remixing by its users based on the tools available. The primary argument seems to concern who is controlling the inventorying, organization, tagging, coding of the data in service of an archive (user, curator, editor, architect?). And what digital tools are best employed in sorting the information? Even a tool offers a preliminary critical perspective.

For instance, in "The Master's Tools Will Never Dismantle the Master's House" (1984), Audre Lorde identified a schism in feminism that highlights missing voices, those voices that did not align themselves with patriarchal control, voices that refused to work within the system to gain power. In digital humanities' interactions

with literary, library, and media studies, especially in the construction of databases, digital archives, and repositories, those marginalized voices exist, but they exist outside the scope of the traditional literary canon even still. Amy Earhart and Jamie Skye Bianco both notice this lack in digital representations of historical and literary materials; while Earhart focuses on the lack of diversity and the replication of the standard literary canon in "Can Information Be Unfettered? Race and the New Digital Humanities Canon," Skye Bianco asserts something more provocative about the very infrastructure of digital humanities:

> Boiled down blithely, the theory is in the tool, and we code tools. Clearly this position never refers to Audre Lorde's famous essays on tools nor to "the uses of anger," but it does summon their politics. . . . Tools don't reflect upon their own making, use, or circulation or upon the constraints under which their constitution becomes legible, much less attractive to funding. They certainly cannot account for their circulations and relations, the discourses and epistemic constellations in which they resonate. They cannot take responsibility for the social relations they inflect or control. Nor do they explain why only 10 percent of today's computer science majors are women, a huge drop from 39 percent in 1984, and 87 percent of Wikipedia editors—that would be the first-tier online resource for information after a Google search—are men. Tools may track and compile data around these questions, visualize and configure it through interactive interfaces and porous databases, but what then? What do *we do* with the data? (99)

The tools, like markup, by their very nature enact a sort of politics that replicates these archival silences that Lauren Klein discusses in reference to American slavery. By offering a "stable publication environment" and peer review to small-scale digital scholarly editions, the 2012 inaugural issue of the revised *Scholarly Editing*, under the editors Amanda Gailey and Andrew Jewell, attempts to balance the digital offerings of cultural materials beyond canonical authors and figures. But, in all of these scenarios, the relationship with the user is also absent; how do users shape an archive and

how do those tools implemented by users reshape and remix that very same archive?

Kathleen Burnett, borrowing from Deleuze and Guattari's concept of the rhizome, notes in "Toward a Theory of Hypertextual Design" that the archive is less about the artifact and more about the user:

> Each user's path of connection through a database is as valid as any other. New paths can be grafted onto the old, providing fresh alternatives. The map orients the user within the context of the database as a whole, but always from the perspective of the user. In hierarchical systems, the user map generally shows the user's progress, but it does so out of context. A typical search history displays only the user's queries and the system's responses. It does not show the system's path through the database. It does not display rejected terms, only matches. It does not record the user's psychological responses to what the system presents. . . . *The map does not reproduce an unconscious closed in upon itself; it constructs the unconscious.* (25; italics added)

The digital archive, some argue, is the culmination of Don McKenzie's "social text," and the database, and to some extent hyperlinks, allow users to chase down any reference. In essence, the users become ergodic and radial readers. McGann, in *The Textual Condition*, defines radial reading as the activity of reading that regularly transcends its own ocular physical bases, which means that readers leave the page in order to acquire more information about the book (i.e., look up a word in the dictionary or flip to an endnote). This allows the reader to interact with the book, text, story through this acquisition of knowledge. The reader makes and remakes the knowledge produced by the text through this continual knowledge acquisition, yet the reader never actually leaves the text. It stays with her even while she consults other knowledge. This creates a plasticity to the text that is unique according to each reader (119).

But the archive, a metaphor once again, is always and forever contaminated, according to David Greetham in *The Pleasures of Contamination*. An archive is less about the text of a printed word

and can be about all facets of materiality, form, and its subsequent encoding—even the reader herself. Scott Rettberg notes that the act of reading prioritizes the experience over the object itself with this idea of ergodic reading:

> The process of reading any configurative or "ergodic" form of literature invites the reader to first explore the ludic challenges and pleasures of operating and traversing the text in a hyperattentive and experimental fashion before reading more deeply. The reader of Julio Cortazar's *Hopscotch* must decide which of the two recommended reading orders to pursue, and whether or not to consider the chapters which the author labels "expendable." The reader of Milorad Pavic's *Dictionary of the Khazars* must devise a strategy for moving through the cross-referenced web of encyclopedic fragments. The reader of David Markson's *Wittgenstein's Mistress* or *Reader's Block* must straddle between competing desires to attend to the nuggets of trivia of which those two books are largely composed or to concentrate on the leitmotifs which weave them into a tapestry of coherent psychological narrative. In each of these print novels, the reader must first puzzle over the rules of operation of the text itself, negotiate the formal "novelty" of the novel, play with the various pieces, and fiddle with the switches, before arriving at an impression of how the jigsaw puzzle might together [*sic*], how the text-machine may run. Only after this exploratory stage is the type of contemplative or interpretive reading we associate with deep attention possible. (para. 13—emphasis added)

As our understanding of digital interruptions in an otherwise humanistic world expands and becomes both resistant and welcoming, we find that in these and other textual encounters, the definition of *archive* expands as well.

See in this volume: cloud, event, memory, prototype, surrogate

See in Williams: alienation, art, consensus, creative, empirical, history, institution, interest, management, representative, standards, subjective, tradition

Notes

1 The *bibliographic code* is distinguished from the content or the semantic construction of language within a text (linguistic code) by the following elements, as George Bornstein describes: "features of a page layout, book design, ink and paper, and typeface … publisher, print run, price or audience.… [Bibliographic codes] might also include the other contents of the book or periodical in which the work appears, as well as prefaces, notes, or dedications that affect the reception and interpretation of the work" (30, 31). *Linguistic codes* are specifically the words. Also within the book are paratextual elements that do not necessarily fall under the bibliographic or linguistic codes.

2 See Matthew Arnold on disinterestedness in *Essays on Criticism* <http://www.bartleby.com/223/0407.html>.

References

"Archive." *OED* 1959 Chambers's Encycl. I. 570/1.

Bianco, Jamie Skye. "This Digital Humanities Which Is Not One." *Debates in the Digital Humanities*. Ed. Matthew Gold. Minneapolis: University of Minnesota Press, 2012.

Bornstein, George. "How to Read a Page: Modernism and Material Textuality." *Studies in the Literary Imagination* 32:1 (Spring 1999): 29–58.

Burnett, Kathleen. "Toward a Theory of Hypertextual Design." *Postmodern Culture* 3:2 (January 1993): 1–28.

Derrida, Jacques. *Archive Fever: A Freudian Impression*. Trans. Eric Prenowitz. Chicago: University of Chicago Press, 1996.

Earhart, Amy. "Can Information Be Unfettered? Race and the New Digital Humanities Canon." *Debates in the Digital Humanities*. Ed. Matthew Gold. Minneapolis: University of Minnesota Press, 2012. 309–18.

Gailey, Amanda, and Andrew Jewell. Introduction. *Scholarly Editing: The Annual of the Association for Documentary Editing* 33 (2012). <http://www.scholarly editing.org/2012/essays/essay.v33intro.html>.

Greetham, David. *The Pleasures of Contamination: Evidence, Text, and Voice in Textual Studies*. Indiana University Press, 2010.

Harris, Katherine. "Archive." *The Johns Hopkins Guide to Digital Media*. Eds. Marie-Laure Ryan, Lori Emerson, and Benjamin J. Robertson. Baltimore: Johns Hopkins University Press, 2014.

Klein, Lauren F. "The Image of Absence: Archival Silence, Data Visualization, and James Hemings." *American Literature* 85:4 (2013): 661–88.

Leary, Patrick. "Googling the Victorians." *Journal of Victorian Culture* 10:1 (Spring 2005): 72–86.

Lorde, Audre. "The Master's Tools Will Never Dismantle the Master's House." 1984. <http://collectiveliberation.org/wp-content/uploads/2013/01/Lorde _The_Masters_Tools.pdf>.

Manoff, Marlene. "Theories of the Archive from across the Disciplines." *portal: Libraries and the Academy* 4:1 (2004): 9–5.

McGann, Jerome. "How to Read a Book." *The Textual Condition*. Princeton, NJ: Princeton University Press, 1991. 119.

Price, Kenneth. "Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name?" *Digital Humanities Quarterly* 3:3 (Summer 2009) <http: digitalhumanities.org dhq vol 3 3 000053 000053.html>.

———. "Electronic Scholarly Editions." *A Companion to Digital Literary Studies.* Eds. Susan Schreibman and Ray Siemens. Oxford: Blackwell, 2008. <http://www.digitalhumanities.org companionDLS>.

Rettberg, Scott. "Communitizing Electronic Literature." *Digital Humanities Quarterly* 3:2 (Spring 2009). <http: www.digitalhumanities.org/dhq/vol/3/2 000046.html>.

Smith, Martha Nell. "The Human Touch: Software of the Highest Order." *Textual Cultures* 2:1 (Spring 2007): 1–15.

Voss, Paul, and Marta Werner. Introduction. "Towards a Poetics of the Archive." *Studies in the Literary Imagination* 32:1 (1999): i–vii.

Williams, Raymond. *Keywords: A Vocabulary of Culture and Society.* New York: Oxford University Press, 1983.