Alan Bilansky

Search, Reading, and the Rise of Database

If you are reading this essay on a screen, it might look like any other essay also on your screen, even if the other was printed a century ago. Current technologies efface many traditional differences between archive and living literature. It is easy to view digital text as a technical and ideological construction, but in this way it resembles older technologies. As part of this distributed system, a scholar, teacher, critic, writer, or book is a node in a network.[1] By altering the network, any new form of textuality offers the possibility of reengineering writing, scholarship and readership.

For Derrida, the drive to archive is a manifestation of the death impulse, because archive must be unmoving and unchanging. Archive promises return to the originary moment, the source of ground truth. This requires texts to be frozen in form and isolated in place: "There is no archive without a place of consignation, without a technique of repetition, and without a certain exteriority. No archive without outside." (Derrida, 1998, pp. 11-12). Texts in the archive, in this view, are dead. They do not change or grow. In contrast, we can speak of living literature, the literature that circulates freely among the living, often in the form of printed books. The canon, we have come to understand, is by its nature alive, growing and changing. It thrives as it circulates in the discussions of living scholars and other readers.

The subject of this essay is digital textuality (often referred to using database as a collective noun),[2] with the focus here on the effects of primary source databases on reading and writing practices in English-language literary studies. In contemporary discussions of digital texts, grizzly metaphors of bodies, alive, dead and otherwise, are remarkably common. These metaphors are live, or—as Haraway describes our technology—rather lively. New information resources are generally adopted in an evolving disciplinary context wherein users and resources influence one another. Over the past decade, literary scholars have come to rely on a complex and growing ecosystem of digital collections of primary texts—vast and searchable. While literary scholars have been quick to celebrate this development and to make good use of it in their work, little systematic analysis of changes in reading and scholarly practice has yet been presented.[3]

A decade ago, Sean Latham asked "How might the digitization of the archive change our critical reading practices?" and hypothesized that digitization could energize scholarly practices by "extracting familiar texts from anthologies and critical editions and resituating them in a strange and often productively alienating context" and allowing "scholars [to] engage in a re-energized cultural studies no longer limited by the material impositions of print" (2004, pp. 412, 420). Ten years later, I believe we now have sufficient evidence to argue that one seemingly minor technical innovation enables a substantially altered way of reading. Search, I argue, is not a trivial change. It has been transformative in ways Latham could not have predicted, spurring significantly increased numbers of citations of primary texts and a new set of reading habits, including sophisticated conceptual mapping of specific digitized sources and self-awareness about the need for a variety of reading techniques.

This paper aims to characterize these changes in scholarly practice. I argue that database represents a third form of textuality, distinct from archive and living literature, and that a new way of reading has developed around it. This article relies on four threads of evidence: First, I survey the scholarly discourse around digitization, examining the ways scholarly readers have discussed database as a category worth attention, necessary to the profession and perhaps transformative of it. Second, I offer a history of the development of significant collections of digital texts based in part on interviews with participants in those developments. The functionality of making text searchable, this history reveals, was not a step that came automatically or easily. Third, I have conducted quantitative analysis of citations of primary texts in two prominent journals, *American Literature* and *English Literary History*, to point to a clear trend in the literature: a substantial increase in citations of primary sources coinciding with the rise of database.   Fourth, interviews with authors of articles from the later sample of articles offer a thick description of scholars' relationship to database, a description which suggests the development of a new way of reading.

Database and Undead Literature

Database is a contested term marking contested terrain. As a collective noun it is used both to indicate a genre or category of thought and its more workaday denotation of digital text collections. Most famously, Manovich (2001) abstracts database and narrative as two basic aesthetics or ways of knowing. Liu (2008) also sees the same two drives competing in the early modern world, only to renew their contest in the present. For Liu, narrative represents the desire to be inside history while database drives to abstract from history; thus he links New Historicist methodology with database. Hayles (2007) identifies database with large digitized corpora of texts made available to algorithmic reading and argues that text mining requires both database

and narrative: database (with algorithm) to find patterns that we could not; narrative to allow us to understand them (p. 1605).

Another important oppositional pairing is database/archive. Derrida (1998) argues that every archive is at once "revolutionary and traditional" because it institutes (is a place of origin) but also conserves (p. 7). This tension permeates debate over the nature of the relation between database and archive, a discussion exemplified in a 2007 *PMLA* roundtable on "database as genre," consisting of an essay by Ed Folsom, co-founder of the Whitman Archive, with five responses. Folsom opposes archive and database, arguing that archives derive "charm and allure" from their "physicality" while "database facilitates access, immediacy, and the ability to juxtapose items that in real space might be far removed from each other" (p. 1577). Among the respondents, McGann (2007) argued that the achievement of the Whitman project resided not in database but in markup: the use of XML to tag semantic and structural elements in the text following the Text Encoding Initiative (TEI) standard. Database, which powers search, is for McGann crucial to how users will interact with the marked-up text, but a separate category.[4] These distinctions, McGann asserts, help us avoid what he sees as a mistaken vision, derived from Manovich, of "archive as reified knowledge" in contrast to "database as liberated knowledge" (p. 1590). Folsom responds by demonstrating how markup embodies interpretation: "On every page of a manuscript that we transcribe, there are features that we either name as an instance of some category or ignore" (p. 1609). The original text, preserved through digitization for future readers (what he identifies as database), allows for future interpretation. The appeal to archival images as the point of origin raises crucial editorial questions[5]—and questions about the relationship between digital collections (sometimes colloquially called digital archives) and traditional archives.

While I accept McGann's critique of the ideology of "database as liberated knowledge," I want to suggest that database's textuality matches precisely neither that of "dead" texts entombed in archives nor the living (circulating) canon. Following Žižek (2009), I want to emphasize that we call vampires or zombies undead because they are not alive but not exactly dead: they are un-dead, other than dead.[6] Database too belongs in this category. Where archive represents what is fixed, dead and buried, database is mobile and mutable. More vast than the living canon, it contains texts no one alive has read. Yet the possibilities of search render it seemingly more knowable and mobile than archive. Archival texts cannot leave the tomb of the physical archive. Database walks among the living. How it has taken the shape it has, however, is not a story of technological determinism.

From Bibliography to Collection to Database: How We Almost Missed Out on Search

As McGann (2011) has observed, what we think of today as "digital humanities" hasn't always been the province of humanities scholars. Indeed, as McGann points out, "much of the fundamental early work establishing the conditions for online humanities research and publication—OCLC, TEI, Project Muse, JSTOR—was done by librarians, systems administrators, and various kinds of computer technologists" (pp. 183-184). Considering the collaboration among librarians, technologists, and scholars is essential to understanding the history of digitization. Scholars might have justifiable concerns about this collaboration, but interviews with people involved with the early development of these databases indicate that scholars who now find full-text search essential to some extent have librarians to thank for it.

While contemporary digital collections are frequently treated as radically different from physical archives, their histories reach far into the analog. Among these, EEBO, ECCO, and

Evans began as attempts at national bibliographies. In the 1930s, entries in each bibliography were fleshed out with microformed images of the texts. Each has its own history driven by economy and cultural politics.[7] Since then, the microform collections have seen regular use by scholars in English, history, and other fields. Digitization (based on images of microform) came in the late 1990s.

To understand the full impact of the change from searching microfilm to digital searching, we must remember that wrestling with microform takes an investment of time, time that must be blocked out in advance. So does working though a printed bibliography or working with a cross-referenced subject index. This manual searching can be as imperfect and incomplete and therefore unreliable as any other manual activity. Once texts are found (and this is not as trivial as it sounds) and displayed in a viewer, they can finally be read.

In graduate school in the 1990s, I worked with the print index commonly called "Evans" (*American Bibliography: A Chronological Dictionary of All Books, Pamphlets, and Periodicals Published in the United States of America from the genesis of printing in 1639 down to and including the year 1820* by Charles Evans). First, you had to know that a given publication existed and was worth looking at. This for the most part was gleaned informally from mentions by professors, footnotes, advice from your advisor, and so on. Works were in chronological order only, and the microform was keyed to numbers unique to the print index. Moving from a research question to a relevant text was a deliberate, deductive process of funneling down, as if reading the library stacks through a microscope. This method did not allow for many surprises. The important point here is not that microform was difficult to use. The most difficult part of using a pre-digital collection like Evans was not library time. It was knowing where to look.

We might call this step in the research process selection, or access, or identifying texts, but it is the single most important scholarly activity altered by the rise of database. Interviews with participants in two important digitization efforts reveal that the current accessibility of texts in digital collections is the result primarily of an innovation, full-text searching, that came late (although not quite an afterthought), at least in the case of Women Writers Online and the Text Creation Partnership's contribution to EEBO.

The Women Writers Project was very important in the development of the markup of archival documents. The initial grant application to the NEH called for the creation of a "textbase" (a term that never caught on). Typesetting texts was always part of the plan, but in the early 1990s the World Wide Web was hardly a dream. The transcriptions produced were intended to be definitive and to allow for various forms of text mining of the corpus. Simple string search is of course a form of text mining, but search *as a means of discovering and selecting text*, if considered at all in the beginning, would come much later. Through the 1990s, WWP texts were distributed in paper form. Access to the texts for scholars meant a process of knowing that a given text existed, then knowing of the WWP and its coverage. Then, a scholar would have to find a listing of the titles that had been encoded and were available—from a flier at a conference or a posting to a listserv, or by requesting a list through the mail. Then a request would be sent by email, or letter, or over the phone (Renear 2012; Renear and Palmer 2009; Flanders 2013).[8] This is essentially the same process we might follow to request a copy of a non-digitized text from a physical archive somewhere. Now of course search is used to discover texts to read in the collection on the Web (assuming the scholar's institution subscribes).

In the late 1990s, Chadwick-Healey contemplated digitizing collections like EEBO (followed quickly by other publishers such as Gale and Readex and their collections such as

ECCO and Evans, respectively). Originally, they planned on a straightforward duplication of the functionality of microform, with users still accessing texts through the bibliography (Sandler 2011).[9] This would essentially have meant little more than allowing scholars to use the microform without leaving their desks, a change in convenience but not in the nature of the research activity itself. While the choice to include images of the archival source in a digital collection might seem intuitive, in fact it was not a necessary one. The Women Writers Project, when it went online, explicitly chose *not* to associate archival images with marked-up pages. In part, this was a pragmatic choice, but it's worth noting that the developers recognized that producing archival images and transcriptions are two distinct endeavors (Renear 2012).

As they planned EEBO, Chadwick-Healey conducted focus groups with researchers at the University of Michigan's library. Afterwards, librarians advocated going beyond the original plan. From the perspective of the librarian—distinct from that of the then-current microform users—working through the bibliography was a barrier to using the collection. It required specialized knowledge of literary and publishing history of the period, as well as some familiarity with the bibliography concerned. The time required to investigate a single question was not trivial. The reasoning for including full-text versions, rather than only images of the microform, was that full text searching would make the collections much more accessible to a larger group of researchers and students who did not have the specialized knowledge of seasoned scholars but who could pursue interesting research questions if it were somewhat easier. In the case of EEBO, the texts were so old that creating transcripts by optical character recognition (OCR) would be particularly error-prone. When publishers were reluctant to invest in transcriptions, a coalition of libraries—the Text Creation Partnership—was formed to finance the transcriptions to be added to EEBO, ECCO, and Evans, which might otherwise have been only

collections of images linked to a bibliography, perhaps with limited OCR transcripts. This resulted in the non-profit TCP employing (as Price [2014] notes) labor in the global south to keep costs low. Proquest eventually sold EEBO to five times as many libraries as expected (Hogan 2011; Sandler 2011).[10]

At first, the idea of full text search was not considered a necessary component of such a collection. The Women Writers Project demonstrates there is nothing essential about linking digital texts to images of analog archival pages. Although common now, this practice had to be constructed, and e-texts can be and still are created without it. Likewise, there is nothing essential about full-text search. Publishers were so uninterested in search that libraries had to prod and coax them along the path to search, providing the transcriptions (funding them and keeping the costs as low as possible) and marketing for the databases.

There has been much talk of late about new forms of reading, what Franco Moretti calls "distant reading" or what Stephen Ramsay calls "algorithmic" reading, practiced by a few. My subject here is how the rise of database has changed the way virtually *everyone* reads. Reference to a discussion thread in the listserv of SHARP (the Society for the History of Authorship, Reading and Authorship) illustrates where we are now (Gómez-Arostegui 2014). One scholar, Tomas Gómez-Arostegui, posted that he had a problem. He downloads PDFs from Google Book Search, and then later returns to Google Books online to search within each book online. (Why this is necessary may require explanation: the search capability in Google Books is made possible by OCR transcriptions that are not included when downloading the book; the downloads consist of page images without the transcriptions.) The problem Gómez-Arostegui posted to the list was that when he returns to Google Book Search, his old queries no longer return the books he downloaded.

This problem (which list members recognized as a problem) indicates an important shift in our expectations about books. The set of images we acquire when downloading from Google Books is more or less as much information as if we had the printed book in our hands. But reading without search is no longer good enough. Now, we expect our books to help us read them. In response to this problem, list members offered several theories as to why books had disappeared from Google Books, and work-arounds were suggested. Maybe the books were removed because of lacking copyright clearances, maybe Google altered the search algorithms or altered the metadata. One suggestion was to save the URLs instead of relying on repeating the query. Others recommended alternate sources, such as Internet Archive, HathiTrust, university and other websites. Others noted the particular holdings and the quality of the texts in various collections, and exactly what is openly available where (Luey et al. 2014). This extremely sophisticated mapping of web resources has been folded into what it takes to be a professional.

The Way We Read Now[11]

How have scholars used these databases, and what effect has this had on their work? Answering this question is complicated. Digital texts still tend to be cited as if accessed in print (Spiro and Segal 2011). This study, however, is interested in how the texts are used, which is even harder to capture. Nevertheless, examination of changes in citation habits and scholars' descriptions of their own writing and reading practices reveals that a new way of reading has accompanied the rise of digital databases in literary studies.

Recent work in information science provides a starting point, but most such work treats scholarship in a broad range of disciplines, effacing distinctive features of reading as practiced by literary and cultural studies scholars. In a provocative bibliometric study of 34 million

articles covering a broad range of disciplines, Evans finds that as more journals and older issues became available online, fewer secondary sources are cited, from fewer journals. Evans (2008) concludes that researchers are reading less. Evans's data are difficult to refute when it comes to citation patterns, but citation behavior does not necessarily parallel reading behavior. Using survey data from 1977 through 2005, Tenopir and King (2008) find that reading patterns differ from citing patterns. In fact, they find that scientists are in fact reading more, even as they cite less. How is this possible?

The way scholars read secondary literature seems to have changed since the growth of large databases of scholarly journals. Scholars employ a broad set of practices to read journals: direct searching, probing, chaining, "netchaining" (a species of chasing citations from one work to another that moves faster and seamlessly because all the texts are on the desktop), scanning, browsing, rereading, reading around, and assessing—often using structural elements like abstracts, conclusions, and pictures to assess without much reading (Bishop 1999; Palmer et al. 2009; Renear and Palmer 2009). All these practices have intensified, until a difference of degree has become a difference in kind. Renear and Palmer (2009) argue that scholars are "scanning, exploring, and getting exposure to more sources," and reading more by engaging in what they paradoxically call "reading avoidance," or "strategic reading," where "the aim is to move rapidly through the literature to assess and exploit content with as little actual reading as possible" (p. 829). Thus, scholars acquire broader exposure, but end up citing less, though in a more focused fashion.

Databases of secondary literature such as the searchable *MLA International Bibliography*, which came substantially earlier than primary source databases, quickly became a required step of scholarship. As Frances Ferguson (2012) recalls

… there was a moment when I was on the editorial board of *Representations* in the 80s,

when the twelve of us sitting around the table looked at each other and said, "This article

is an article that was written using a substantial database, with word search." And that

was the first moment when I was really conscious that you could see that articles had

been affected by the capacity to do searches using specific terms.

It was clear, then, that the search had affected the parameters of the argument.   Likewise, Laura

Stevens (2012), editor of *Tulsa Studies in Women's Literature*, observes that the databases of

secondary literature have affected reviewers' expectations:

It's striking to me that even when – perhaps because – it is so very easy to track down the

relevant scholarship on a given topic, it's very common for our specialist readers to

complain that submissions have not really attended to what has been said already.

The effect of secondary databases, then, has been reciprocal, shaping both how academic readers

use information and how they are expected to read.

In literary and historical studies, the differences between the way primary and secondary

sources are accessed, read, and written about makes the situation different from many other

fields, in which scholarly reading can be equated with reading the work of other scholars. To

examine specifically how historical primary text databases have affected practices in literary

studies, then, I conducted a small-scale examination of citation practices in two literary journals

and interviewed authors of some of the articles that suggested a strong database influence.

Qualitative and quantitative analyses of the contents of two major journals in literary

studies show two clear trends: The average number of primary sources used increases rather

dramatically after the introduction of primary source databases. My analysis covered articles in

sixteen years of *American Literature* and *English Literary History* to enumerate the primary sources cited. These journals were chosen as being long established, respected and highly selective; in addition they represent two major sub-fields of literary study. Two spans of time, 1995–2001 and 2002–2011, were chosen to represent two distinct moments in scholarly practice. The first was before most of these databases were introduced, but after any initial surge in archival research that might be attributed to New Historicism independently of the effect of databases. The second period represents the rise of database.

I only counted articles focused on literature from the beginnings of printing to the end of the nineteenth century, because these were the periods of literary history that seemed likely to have been affected the most by primary source databases.[12] I counted only unique primary sources from the same period.   A qualifying article, then, was one focusing on works, authors or subjects in England or North America between the 1470s and 1900.

Here are my findings:

	Between 1995 – 2001	Between 2002 – 2011	Representing an increase of
English Literary History	3,187 sources in 232 articles for an average of 13.7 primary sources	4,650 sources in 303 articles, for an average of 15.3 primary sources	11%
American Literature	952 sources in 82 articles for an average of 11.6 primary sources	1,642 sources in 104 articles, for an average of 17.3 primary sources	49%

Although this analysis covers only two journals, the results are quite clear: a quantitative escalation in behavior that parallels a clear escalation in discussion of digitization and its importance within the discipline. An eleven percent increase is clearly significant. Forty-nine percent is amazing.[13]

In addition, although I have not quantified it, another trend I noticed was that the diversity of primary sources brought to bear on the analysis of a literary text or topic increases notably. Primary sources cited between 1995 and 2001 tended to be more canonical and more clustered around a single author than did the primary sources cited after 2001. That is to say, in the earlier period, primary sources were more likely to have been drawn from a single author and figures closely associated with that author, while in the later period a more diverse range of primary sources, including more anonymous sources, were used. These changes, added to the significant increase in use of primary sources overall, suggest that the use of primary text databases has substantially altered the practice of literary scholars, enabling them to situate literary studies in a broader cultural context than was previously practical.

A fuller sense of the practices behind these changes emerges from interviews with the authors of a number of the articles in the sample. The interviews demonstrate that use of these databases is close to universal in literary studies over the past decade. Everyone who responded was very clear: the scholars interviewed would have a hard time conceiving of their work without the use of databases. Database is becoming synonymous with doing historical work. While this in itself is hardly surprising, scholars' descriptions of how they use the databases reveal some of the specific characteristics of the modes of reading encouraged by database.

These interviewees complicate an understanding of how search is used. In one sense, reading texts in digital collections parallels what scholars do when reading though archival materials in physical archives. A scholar might open a box of papers (say from a union archive) and scan through expense sheets, letters, monthly reports from organizers, lists of those who signed up for union cards, a menu from an annual banquet, receipts for the purchase of tents, and so on. A general understanding grows, and occasionally something truly exciting is found. Within the realm of database, however, search makes things more efficient and faster, because it is more rapidly focused.

Based on experiences described by the scholars interviewed, I group encounters with digital database texts into two categories: background and surprise. As Susan Wolfson (2011) put it, digital databases "make my work more efficient, but also surprise me with what I can find in searches." The first category includes a variety of searching and reading strategies, using searchable databases to skim, to "read around" in a body of discourse, to gain general background, to understand ideological and linguistic trends of a period. Such reading might result in a citation, often citing a few sources as exemplars of larger trends, but in some cases this reading will inform further searching but not result in citations. The second category, the occasional "surprise" or "discovery," identifies previously unknown texts that merit close study.

The notion of garnering general background from databases came up repeatedly. Robert Matz (2011) reported frequent use of EEBO "to read around in early modern writings on a subject," while Edlie Wong (2011) called database a "snapshot" of cultural context. Douglas Guerra portrays that snapshot in linguistic terms, saying that databases "helped me to get a broad look at some of the key terms in my research" offering "the ability to situate some of the language that I'm tracking in a kind of statistical sweep. . . a kind of on-the-fly OED for

American culture and language; a personally directed anthology." Such background research, Guerra (2011) notes, "doesn't always directly make its way into the work, but informs a perspective indirectly."

In terms of periodicals, context includes the ability to examine facsimiles, as described by Blackwood (2011), who says databases provide fuller understanding of context:

> both physical context on the page and the cultural context. . . and to discover interesting juxtapositions (noting, for example, the physical proximity of runaway slave advertisements--as I argue in the article, a sort of "portrait"--to classified advertisements for daguerreotype studios).

> Blaine Greteman (2011; who also says "main gain is efficiency") tells a story of his own

use of database in his scholarship that suggests a much more direct shift in habits of citation:

> I use EEBO tons, and *ELH* (and *RQ*) . . . actually changed their policies because of how much I use the databases (i.e., they used to require full html links, but it got to be absurd because of how many I needed to use).

This narrative is fascinating, both in its sense of the power of database and as a clear example of how one scholarly practice affects another and both center on technology, as scholarly communication itself strains under the weight of digital texts—even as we might see digital texts as practically weightless and most credit them only with greater efficiency.

The use of databases described by scholars is substantially, recursively, and seamlessly integrated with the writing process in a way that archive is not. Unlike what he calls "intensive, non-improvisational stuff (hands on research trips, following bibliographic leads, etc.)," Guerra said database allows more freedom:

Many of these [digital] archives essentially allow you to use the microfiche at home. So

you're accessing the same materials, but because you don't have to PLAN on using

something, you can more easily go where your thought takes you, be more

improvisational with a piece of writing. . . . I definitely found myself in the midst of

writing, struck with an idea of something I had seen, and then, being able to scope out

that source very quickly, was able to get the idea/source into my work.

Although he would not say that database changes the questions he asks,

it does allow you to ask "more" questions. . . . having a database at hand allows you to

come to a question that you hadn't necessarily prepared for and take a shot at addressing

it by getting a quantitative perspective or hard to find information much more quickly (or

to know more readily if that question is too big to get at without more legwork).

It's notable how empowered Guerra and the other scholars were in discussing the use of this

technology. The use of database merges seamlessly with the rest of his research practice.

Database seems to have influenced what scholars read, increasing reading of ephemera and

tracking of cultural trends through a variety of texts.

What of the reading methods employed? Blackwood (2011) theorizes her database

reading as a hybrid:

I often switch between the "distance reading" techniques that a keyword search can

provide and more traditional close reading techniques. For example, I use keyword

searches to get a sense of how a particular topic or issue or word becomes more or less

present in a cross section of popular cultural artifacts (such as newspapers and

magazines). This is the "distant reading" part. But then, I also like to employ extreme

close reading techniques that examine in detail how these topics/issues/words are

employed, or what sorts of contexts they appear in, et cetera.

Here Blackwood uses the term "distant reading" *not* in Moretti's sense: reading at a scale that

makes close reading by humans a practical impossibility (2005; 2009)—reading corpora of texts

without reading. My point is not to quibble with a reappropriation of Morretti's term. Rather,

what's noteworthy is Blackwood's recognition that she is reading differently, doing something

different enough from traditional close reading to deserve its own name. Reading database is

distinct from either close or distant (as in algorithmic) reading, in part because it calls on at least

some version of each. Search results are not enough to serve as evidence on their own. They

suggest trends and help to identify texts to read more closely.

One genuinely new part of reading required by database is the process of developing a

working bibliographic map of the databases relevant to a scholar's interests, understanding the

content, limits, strengths, and weaknesses of specific databases. For example, Wong (2011) and

others pointed to a specific limit of certain databases: publications by and for minority and

subaltern groups are still terribly underrepresented in digital form.

A number of newspapers such as the Jamaica-based Gall's Packet News Letter (for my

current research on turn-of-the-century North American writer of Chinese descent Edith

Maude Eaton also known as Sui Sin Far) are not included in these newspaper databases

[such as American Periodicals Series] even as work continues on digitizing the full run of

the Frederick Douglass' Paper.

Experience makes Wong aware of technical shortcomings: "I've found the keyword

searches on APS (in particular) to be unreliable and incomplete." She concludes, "Generally, I

use the online databases to triangulate information and to test out initial hypotheses—often from

the comfort of my home, and I then conduct a second round of primary research via newspaper microfilm collections." Similarly, two scholars both describe research processes that not only make use of digital resources but adapt to the particular strengths and limitations of those resources:

> I had marked up my copy of *Clarissa* three times, and then I did related searches on the ECCO etexts to see if I could find the same things--or different things of interest. (I did the same with the version in the 18th-c. Fiction database, although this isn't keyed to a particular text in the same way.) I also searched related volumes of texts in the library to answer questions that had arisen, or questions that might arise with diligent reading. I did ECCO searches to try to approximate this process as well. (Valenza 2011)
>
> Some of this has to do with the interface too, what kinds of tools the database offers for manipulation, mark-up, etc. If you just have a search bar, you're just going to search. And if you want to have one of those browsing "a-ha!" moments that you can have in a library, then you need to create your own browse-able space with your search--not always visually encouraged by a hierarchical list of results. (Guerra 2011)

Scholars are very aware of the functional limitations and strengths of particular databases. William Christie (2011) also spoke of designing his own interface. All interviewees spoke of developing sophisticated ways to use databases intelligently, to compensate for weaknesses, and to combine database research with physical archives and print texts. Search strategies are tailored to particular databases and particular research agendas. As Blackwood (2011) says:

> I believe that digital sources and archives have not only changed but actually shaped the way that I research. . . . in a sense, I learned how to research while using a mixture of these newer and older techniques.

Given the ubiquity of database in literary scholarship, it makes sense that learning database is itself now an essential part of the acculturation of scholars and the learning of their trade. It is common to speak of undergraduates as having grown up with instant messaging and Flickr and Facebook and Twitter. While older scholars have been quick to make good use of digital resources, we might also find it useful in thinking about institutional history to recognize as a particular sort of "digital natives" a generation of scholars who came into their own as professionals with the help of EEBO and Google Books.

Digital text has also affected the reading of undergraduates. Hayles (2010), among others, appreciates the reading strategies undergraduate students have developed in digital environments. The problem she begins with is that many educators want to teach close reading, but find students incapable of "deep attention" because of habits of reading acquired from digital media. She argues that alternate ways of reading digital texts have value and scaffold or supplement close reading. She suggests that we should appreciate and make use of these new ways of reading, instead of just trying to bring students into the close-reading fold:

> Given the increase in digital reading, obvious sites for new kinds of reading techniques, pedagogical strategies, and initiatives are the interactions between digital and print literacies. Literary studies has been slow to address these possibilities, however, because it continues to view close reading of print texts as the field's essence. . . (Hayles 2010, p. 65)

Undergraduate use of database is also interesting, in particular as student work is repeatedly referred to as "professional-quality" or "publishable" (Wolfson 2011). Describing a course in American literature, Hanlon (2005) reports that searchable databases allowed him to lead undergraduates to make historical discoveries that shed light on their reading. He does this by

directing them to text-searchable databases. For example, after reading Crane's "The Blue Hotel," "one student got the idea of searching for documents written during the 90s containing the word 'Swede.' . . . the decade before Crane wrote the story was a time of great fretfulness over…Swedish immigrants" (Hanlon 2005, p. 100). Hanlon finds the activity transformative, noting that undergraduates who "couldn't tell me who was president before Clinton at the beginning of the semester, by the end of the semester are teaching me three different ways tableaux vivant were believed to improve women's health during the fin de siècle" (p. 101). The ability to search acts as scaffolding to enable this sort of historical research with minimum barriers to entry.

A similar emphasis on the originality of student work enabled by database is apparent in Valenza's project in which she asked her students to find Gothic Bluebooks in ECCO.

> Then we talk about the features of these books: common words in titles, common words in stories, common prices printed on the covers, common lengths of the books, fixed number of years during which they were published, and students combine these factors to do searches to turn up documents. It's a fun game to play together or as homework, and then I have each student choose one that he or she found to write about in a paper. Most of what you'll find on the topic of Gothic Bluebooks or individual abridgments of Gothic novels in Wikipedia is written by . . . . students who were once in my classes. (Valenza 2011)

Online access to these texts through databases truly marks a qualitative difference. Searches can be completed so much more quickly, it hardly feels like the same activity. Perfect or even imperfect transcripts of pages in these databases are indexed for searching. Even in cases where

OCR gets some words wrong, word search is valuable, with substantially good results obtained quickly.

The interviews clarify that, whatever the technical importance of search, it has made a revolutionary difference in the lives of these scholars. The ease of searching and accessing these texts means that a task takes so much less time, it can be performed iteratively. Scholars can now test multiple hypotheses in minutes, where before they would have to block out time in the basement of the library to answer a single question. Experimentation and risk-taking, following long shots, is easier now. Professors can find texts to use in the classroom, say, an hour beforehand. They can assign students to do original research.

In 2004, Latham claimed digitization would create a difference of kind in literary studies, not merely one of degree: "We make a serious error … when we treat these digital technologies as mere labor-saving devices" because "the results produced by search engines thus constitute not simply a new kind of index, but a new kind of textuality" which he reads as a text in itself (Latham 2004, pp. 416, 418). Latham was predicting new ways of reading enabled by search; a decade later, as I have argued, digital textuality has transformed literary study, and we can now see in sharper focus the nature of this new form of textuality.

While it might seem intuitive that search makes reading efficient, the kinds of reading enabled by digital search have quickly increased in sophistication. Latham serves as an example of where these techniques began. He reads the results of a search for one word, "imperialism" in one journal, *The New Age*, between 1907 and 1922, accessed through the Modernist Journals Project database. Surveying the results in 123 of a possible 130 issues, he concludes,

Gradually, the discontinuous list of hits generates a series of historical coordinates by which the discourses of imperialism can be mapped, coordinates which simultaneously intersect other maps of gender and class. (2004, p. 423)

While an archeology of one term has value, the power of search as a reading technique increases as searches are used to reveal relationships among terms, as when Robert Hume argues in 2007 that a single search relating two terms could form the basis of a book, with a single query in ECCO: to "find 'money' within ten words of 'value' in all eighteenth-century books" (p. 13).[14]

Search started as a convenience, a way to remove barriers to access. But it has become something else. As Underwood (2014) has argued, scholars using search are doing text mining, obscured topic modeling, even as some of those same scholars might regard "text mining" as abstruse and alien—and the mechanics of search could be more transparent for users to understand, reflect on, and alter. Although search can seem like direct interaction with a text, it is actually a sort of mediation. This sort of access, arguably, becomes a new sort of textuality. Databased texts are encountered in a manner distinct from living literature or the dead literature of archive. Physical archives are overtly mediated: there are finding aids, bibliographies, gloves, archivists who want to help or to guard. Even a box of documents that no living person has read will have a label. Yet archives can feel unmediated, taking us back in time, as Derrida says, to the moment when they lived. Because of the explosion in publishing in the 19th century, no database can claim comprehensiveness for that century. This results in a reliance on random selection and serendipity, as the vast size of the databases and the mysterious mechanisms of the search function yield surprise and also a sort of faith in metonymy, where any text uncovered is taken to be possibly representative of a wider discourse of the time.[15]

Database is more than archive mediated. It is archive animated, back from the dead, made to serve. It answers questions, it changes shape to suit our needs, it's here to help. Rather than the chaotic zombies of Romero's movies, these are the original zombies of voodoo, corpses magically reanimated to serve the will of the witch doctor, or Proquest.

In Jonathan Freedman's contribution to the discussion of the Whitman Archive, he leapt from the Whitman Archive to Google to findfound database revivifying scholarship:

> no one doing research would want to forgo the amazing search capacities that Google puts literally at one's fingertips: hours spent at the library or, at best, searching concordances now telescope into microseconds; the Boolean ability to link heterogeneous subjects and find once-occulted connections and interconnections makes scholarship invigoratingly fun. (Freedman 2007, p. 1597)

It's worth noting this comes in an essay that is profoundly worried about digitization.

Conclusion

In 2007, roughly five years after the release of ECCO, eighteenth-century literary historian Robert Hume wrote that "Only a fool would go to a school that did not have ECCO if he or she intended to work in the eighteenth century" (2007, p. 13). Database may represent utopian promise for some, but for Hume, it also defines a new economic reality by setting new terms for competition. He is fully aware of the weaknesses of the interface, and the shortcomings of the Short Title Catalog reproduced in ECCO.[16] However, he could not be any plainer about the value of this database to one area of study:

> A university that does not have ECCO is not a serious player in eighteenth-century British and American studies—in literature or in anything else. Any institution giving

graduate degrees in eighteenth-century subjects reduces itself to below minor-league

status if it does not provide ECCO to its students—and is putting its publishing faculty at

a crippling disadvantage. (2007, p. 13)

Hume's claim that one needs access to ECCO to be a "serious player" is interesting because, in a

pre-digital generation, *he* became a serious player without it. Serious scholarship has changed.

It has been re-engineered. Hume wrote this essay as a mature scholar, already an important

literary historian, and at a point in history when he thinks this is no longer possible for people

entering the field without database.

Echoing Hume's sense of the crucial importance of database, an earlier career scholar,

Sarah Blackwood (2011), notes how her access changed with her first job:

When I learned to use these online archival databases, I was a graduate student at a very

well-funded university. I am now an Assistant Professor at an undergraduate teaching

institution that does not have the funds to subscribe to these archival databases. I am still

working out exactly how to reshape my research once again in response to this reality.

Even though Blackwood's narrative of research turned on a discovery made in physical archives,

she still feels the loss of digital archives. One crucial difference, we should note, between these

comments about database and those from the forum on the Whitman Archive is basic economics:

freely available databases enable a discussion about liberation of texts and readers that becomes

more complicated when discussing expensive proprietary databases, where the access of scholars

depends on their location within the differentiated higher education system.

Similar issues of access were raised in a 2010 discussion on the SHARP list, prompted by

an article by Robert Darnton advocating for the creation of the Digital Public Library of

America.[17] The project as Darnton described it displays a faith in database as a cure for several

ills. Notably, Darnton (2010) explains the value of database in terms of Google, which generally

has come to stand in for the idea of search:

> Google demonstrated the possibility of transforming the intellectual riches of our
>
> libraries. . . into an electronic database that could be tapped by anyone anywhere at any
>
> time. Why not adapt its formula for success to the public good—a digital library
>
> composed of virtually all the books in our greatest research libraries available free of
>
> charge to the entire citizenry, in fact, to everyone in the world?

In the listserv discussion of this article, questions of access were raised demonstrating that for

some scholars, at least, the scholarly digital divide was an urgent concern. List members largely

supported the DPLA, but noted that limited access to proprietary databases was an unaddressed

issue.

In particular, unaffiliated scholars pointed to the current lack of access to the large

databases subscribed to by large university libraries. Among the first was Jerry Blaz (2010):

"What chance does an independent scholar have when the scholar must get the same convenient

access on his computer to journal articles [as] those associated with institutions with research

libraries?" Vivienne Dunstan (2010) said, "I recently finished my PhD, and hope to continue to

convert my thesis into more journal papers and do more independent research. But this lack of

access to essential academic resources is extremely limiting. As soon as I graduated my library

access privileges, including online journals, were stopped." More joined the conversation, as this

issue clearly touched a nerve. As Galey (2012) argues: "the fundamental challenge facing digital

textual scholars goes beyond adapting bibliographic vocabulary and methods to new materials."

As Karen Reeds (2010) pointed out, "The immediate, passionate response of SHARP list

members shows how many independent scholars there are on this list, and how important the

issue of accessibility to libraries and online databases is to good scholarship in our field." She

added, "I'm very conscious of the ways my own erratic access to databases skews the kind of

book history I do." This seems a similar moment to one Brunton (2012) describes, "when we

abruptly become aware of the spheres in which we live, the life-support systems on which we

rely, and the recognition of our edges and the extraordinary forces we must manage to make it

clear who we are, and how we conceive of ourselves" (pp. 51-52).

We could call database a convenience, a time-saver. This is how it is often discussed in

publishers' marketing materials and in some research on information behavior.[18] This ignores

the context in which contemporary scholars work, a context of a strongly felt increased pressure

to publish quickly in order to secure a position and earn tenure.

***

Borrowing the term "hyperreading" from Jame Sosnoski, Hayles describes his examples:

which he defined as "reader-directed, screen-based, computer-assisted reading" . . . .

examples include search queries (as in a Google search), filtering by keywords,

skimming, hyperlinking, "pecking" (pulling out a few items from a longer text), and

fragmenting . . . We may add juxtaposing, as when several open windows allow one to

read across several texts, and scanning, as when one reads rapidly through a blog to

identify items of interest. (Hayles 2010, p. 66)

"Hyperreading" sounds strikingly like "strategic reading." The list of the strategies is

remarkably similar.   It also resembles the sorts of activities described by the scholars I

interviewed: strategies, essentially, to read more in less time. I would call this way of reading,

driven by search, bringing together close reading and history, demonstrating a certain mania to

discover new history, database fever.

We can go further. Not only do we see this new sort of reading. Considering the comments of practicing literary scholars over the last decade, we can see an evolution toward a fully developed new form of reading, from Latham to Hume to the scholars I interviewed. In 2004 Latham's demonstration is to read the texts found by searching for one word in one year's run of one journal. In 2005, Hanlon's undergraduates are using simple text searches to make historical discoveries. In 2007, Hume considers basing a new book on a more complicated query, two words in proximity to each other. By 2011, the writers I interviewed all used complicated search strategies across multiple databases, free and subscription, large scale and small, assuming that many interactions with database are required for a scholarly project. And in 2014, all the scholars who contributed to the SHARP-L thread about disappearing Google books have developed sophisticated maps of digital resources. (And I suspect that many scholars if interviewed in 2015 would also feel obliged to use other digital tools like an *n-gram* viewer.) But this is not taken to be anything special now. It's no more surprising than the fact that Barack Obama can't live without a Blackberry, or that Gerald Ford knew how to answer a telephone, while Abraham Lincoln didn't. Some might worry that something is lost. One could argue that scholars with the knowledge to use the old bibliographies like the STC generally also had broad historical background, and thus encountered the texts organically—getting the history (for lack of a better term) right.   However, the scholars I interviewed were publishing in highly selective peer-reviewed journals. Their articles—and the rest of those published in the issues I reviewed—were well-informed historically and used database as a means to new historical discoveries.

In 1913, W.W. Greg called for a "scientific" or "systematic" bibliography that could account for all aspects of all books. A century later, the enterprise of bibliography may seem

antiquarian, but the present calls urgently for a systematic bibliography of digital texts.   Some have argued that we need a *critical* digital humanities, but what I believe we need immediately is a *descriptive* digital humanities to offer a more systematic understanding of digital textuality. While I have attempted here to sketch out some of the material history, the provenance, distribution, and material form of some large collections of digital texts, as well as some of the ways they are read, this is by no means an exhaustive description. Further, different stories would emerge for other collections, such as 18th-Century Fiction, Google Books, Early American Newspapers, American Periodicals Series, British Periodicals, or the Library of Congress's Historic American Newspapers—and this is just thinking of databases associated with Anglo-American literary study. If other national bibliographies were brought into the conversation, we might find even more radically different stories about preconceptions, provenance, distribution, material form, and most importantly their impact on ways of reading among scholars and students.

We need sustained discussion of database as the enabling conditions for future research that will take forms we haven't dreamt of yet.

References

Best, S. and Marcus, S. (2009). Surface Reading: An Introduction. Special Issue: *The Way We Read Now. Representations* 108.1: 1-21.

Bishop, A. P. (1999). Document Structure and Digital Libraries: How Researchers Mobilize Information in Journal Articles. *Information Processing & Management* 35: 255-279.

Blackwood, S. Interviewed by: Bilansky, A. (25 July 2011).

Blaz, J. (2010). "Re: Darnton on the crisis in libraries – a must!" [6 December 2010]. *SHARP-L.* Society for the History of Authorship, Reading, and Printing.

Brockman, W. S. et al. (2001). *Scholarly Work in the Humanities and the Evolving Information Environment*. Washington, D.C.: Digital Library Federation Council on Library and Information Resources.

Brunton, F. (2012). Constitutive Interference: Spam and Online Communities. *Representations* 117.1: 30-58.

Bulger, M. et al. (2011). *Reinventing Research? Information Practices in the Humanities.* London: Research Information Network.

Christie, W. Interviewed by: Bilansky, A. (28 June 2011).

Darnton, R. (2010). The Library: Three Jeremiads. *New York Review of Books* (November 23, 2010) http://www.nybooks.com/articles/archives/2010/dec/23/library-three-jeremiads/.

Derrida, J. (1998). *Archive Fever: A Freudian Impression.* Trans. E. Prenowitz. Chicago: University of Chicago Press.

DPLA. "FAQ: What is the DPLA?" Digital Public Library of America, accessed July 22, 2013,

http://dp.la/info/about/faq/.

Dunstan, V. (2010). "Re: Darnton on the crisis in libraries – a must!" [7 December 2010].

SHARP-L. Society for the History of Authorship, Reading, and Printing.

Evans, J. (2008). Electronic Publication and the Narrowing of Science and Scholarship. *Science*

321: 395-399.

Ferguson, F. Interviewed by: Bilansky, A. (18 March 2012).

Flanders, J. Interviewed by: Bilansky, A. (24 March 2013).

Folsom, E. (2007). Database as Genre: The Epic Transformation of Archives. *PMLA* 122: 1571-

1579.

Foucault, M. (1972). *The Archeology of Knowledge and the Discourse on Language.* Trans.

Smith, A. M. S. New York: Pantheon.

Freedman, J. (2007). Whitman, Database, Information Culture. *PMLA* 122: 1596-1602.

Fyfe, P. (2009). The Random Selection of Victorian New Media. *Victorian Periodicals Review*

42.1: 1-23.

Galey, A. (2012). The Enkindling Reciter: E-Books in the Bibliographical Imagination. *Book*

History 15: 210-247.

Gitelman, L. (2014). Searching and Thinking About Searching JSTOR. *Representations* 127.1:

73-82.

Gómez-Arostegui, H. T. (2014). Google Books disappearing…. [20 January 2014]. *SHARP-L*

[online]. Society for the History of Authorship, Reading, and Printing.

Greg, W. W. (1911-13). What is Bibliography? *Transactions of the Bibliographic Society* 12: 39-53.

Greteman, B. Interviewed by: Bilansky, A. (29 July 2011).

Guerra, D. Interviewed by: Bilansky, A. (1 July 2011).

Hanlon, C. (2005). History on the Cheap: Using the Online Archive to Make Historicists out of Undergrads. *Pedagogy* 5.1: 97-101.

Haraway, D. (2004). A Manifesto for Cyborgs. In Haraway, D. *The Haraway Reader.* New York: Routledge. 7-45.

Hayles, N. K. (2007). Narrative and Database: Natural Symbionts. *PMLA* 122: 1603-1608.

Hayles, N. K. (2010). How We Read: Close, Hyper, Machine. *ADE Bulletin* 150: 62-79.

Hogan, J. Interviewed by: Bilansky, A. (14 June 2011).

Hume, R. (2007). The ECCO Revolution. *The Eighteenth-Century Intelligencer* 21: 9-17.

Kichuk, D. (2007). Metamorphosis: Remediation in *Early English Books Online (EEBO)*. *Literary and Linguistic Computing* 22.3: 291-303.

Latham, S. (2004). *New Age* Scholarship: The Work of Criticism in the Age of Digital Reproduction. *New Literary History* 35: 411-426.

Liu, A. (2008). *Local Transcendence: Essays on Postmodern Historicism and the Database*. Chicago: University of Chicago Press.

Luey, B., Henderson, A., Ward, M., MacLeod, K., Faux, D., Jackson, L., Johnson, D., Peterson, W. S., Gibson, J., Carpenter, K., Riedel, D. and Armbruster, C. Replies to "Google

Books disappearing…." [20-24 January 2014]. *SHARP-L.* Society for the History of

    Authorship, Reading, and Printing.

Mak, B. (2011). From Facsimile to Fact in the Information Age. Paper presented at the annual

    meeting of the Society for the History of Authorship, Reading, and Publishing,

    Washington, DC.

Mak, B. (2014). Archaeology of a Digitization. *Journal of the American Society for Information*

    Science and Technology 65.8: 1515-1526.

Manovich, L. (2001). *The Language of New Media*. Cambridge: MIT Press.

Matz, R. Interviewed by: Bilansky, A. (26 August 2011).

McGann, J. (1992). *A Critique of Modern Textual Criticism*. University of Virginia Press.

McGann, J. (2007). Database, Interface, and Archival Fever. *PMLA* 122: 1588-1592.

McGann, J. (2011). On Creating a Usable Future. *Profession*. 182-195.

MLA Committee on Scholarly Editions. (2011). *Guidelines for Editors of Scholarly Editions*.

    http://www.mla.org/cse_guidelines.

Moretti, F. (2005). *Graphs, Maps, Trees: Abstract Models for Literary History*. New York:

    Verso, 2005.

Moretti, F. (2009). Style, Inc.: Reflections on Seven Thousand Titles (British Novels, 1740-

    1850). *Critical Inquiry* 36: 134-158.

Navas, E. (2012). Modular Complexity and Remix. *Anthrovision* 1.1,

    http://anthrovision.revues.org/324.

Palmer, C. L., Teffeau, L. C., and Pirmann, C. M. (2009). *Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development.* Dublin, OH: OCLC Research. http://www.oclc.org/content/dam/research/publications/library/2009/2009-02.pdf?urlm=162919 (accessed 31 August 2014).

Price, L. (2014). Response. *Representations* 127.1: 93-97.

Reeds, K. (2010). "Re: enough already—Darnton and accessibility." "Re: Darnton on the crisis in libraries – a must!" [8 December 2010]. *SHARP-L.* Society for the History of Authorship, Reading, and Printing.

Renear, A. H. and Palmer, C. L. (2009). Strategic Reading, Ontologies, and the Future of Scientific Publishing. *Science,* 325: 828-832.

Renear, A. Interviewed by: Bilansky, A. (13 August 2012).

Sandler, M. Interviewed by: Bilansky, A. (14 June 2011).

Spiro, L. and Segal, J. (2011). *The Impact of Digital Resources on Humanities Research*. Houston: Rice University, http://library.rice.edu/services/dmc/about/projects/the-impact-of-digital-resources-on-humanities-research. accessed 13 September 2011.

Stallybrass, P. (2007). Against Thinking. *PMLA* 122: 1580-1587.

Stevens, L. Interview by: Bilansky, A. (18 June 2012).

Sukovic, S. (2008). Convergent Flows: Humanities Scholars and Their Interactions with Electronic Texts. *The Library Quarterly* 78.3: 263-284.

Tenopir, C. and King, D. W. (2008). Electronic Journals and Changes in Scholarly Article

    Seeking and Reading Patterns. *D-Lib Magazine* 14:11/12,

    http://www.dlib.org/dlib/november08/tenopir/11tenopir.html.

Underwood, T. (2014). Theorizing Research Practices We Forgot to Theorize Twenty Years

    Ago. *Representations* 127.1: 64-72.

Vaidhyanathan, S. (2010). "Thank You, Judge Chin," *The Chronicle of Higher Education* 24,

    http://chronicle.com/article/Thank-You-Judge-Chin/126888/.

Valenza, R. Interviewed by: Bilansky, A. (29 June 2011).

Wald, P. Interviewed by: Bilansky, A. (6 April 2012).

Wolfson, S. Interviewed by: Bilansky, A. (29 June 2011).

Wong, E. Interviewed by: Bilansky, A. (29 July 2011).

Žižek, S. (2009). *The Parallax View*. MIT Press.

Notes

[1] As Foucault describes it, "the frontiers of a book are never clear-cut. . . . it is a node within a network . . . [that] construct itself only on the basis of a complex field of discourse" (1972, p. 23). Following Haraway, we can say that humanities scholars in their dependence on the technologies of text are all cyborgs (2004).

[2] Throughout this essay, I have followed the use of many of my sources, using *database* collectively when referring to a mode of textuality, but as a count noun when discussing specific digital collections arranged in database form. Further discussion of the theoretical considerations follows below.

[3] Information behavior researchers have examined scholarly database use, but mostly with regard to secondary sources and mostly in the physical sciences or the humanities generally. Questions specific to literary studies have as yet received little attention. For natural sciences, see Palmer et al. (2009) and Renear and Palmer (2009). For the humanities, see Brockman et al. (2010), Bulger at al. (2011) and Sukovic (2008).

[4] Moreover, McGann argues, "no database can function without a user interface, and in the case of cultural materials the interface is an especially crucial element" because it "embeds, implicitly, and explicitly, many kinds of hierarchical and narrativized organizations" (p. 1588). Stallybrass (2007) responds to Folsom focusing on the importance of search.

[5] McGann (1992) has argued, and current MLA (2011) guidelines for editors agree, that choosing copy text is always an interpretive choice.

[6] I am appropriating Žižek's explorations of the undead for different ends. For Žižek the undead justify violence and destruction by embodying life without meaning. In the discursive realm, I argue, the reanimation of the archive seems to add meaning.

[7] See Gitelman (2014, pp. 73-82), Kichuk (2007), Mak (2011), and Mak (2014).

[8] In addition to serving as co-director Renear wrote the NEH application to produce a "textbase." Flanders, involved from the start, is the current director.

[9] Sandler, director of the TCP, is Head of Research Services and Humanities at University Library of The University of Michigan.

[10] Hogan is a developer at ProQuest.

[11] As Best and Marcus (2009) note in introducing *The Way We Read Now,* the title indicates that "now we do things a bit differently than they did back then" (p. 2).   This essay does not take up theoretical questions about symptomatic and surface reading discussed there, but shares an interest in uncovering the present moment, here as conditioned by technological change.

[12] Databases have certainly affected manuscript studies, but it seemed likely that the effects were more pronounced in studies of printed materials; thus I excluded manuscript sources. Twentieth-century materials have seen some digitization projects, but the greater availability of these materials and the great volume of printed material, along with issues of copyright, means that there are no databases of twentieth-century materials approaching comprehensive collections.

[13] I touched on the differences in the change in interviews with the editors of *AL* and *ELH*. Priscilla Wald (2012) of *AL* speculated that the existence of American studies and its interdisciplinary direction might have something to do with the larger increase. Frances Ferguson (2012), then editor of *ELH*, said that to some extent the preference of that journal is

still on single authors and texts. Also, it's worth noting that *ELH* publishes more and shorter articles.

[14] Hume and Latham are among the few who have recognized search results as worth theorizing. Another exception is Navas (2012). Interpreting search results in Flickr and youTube, Navas writes: "Search redefines the way people come to terms with historical developments that are constantly recycled and remixed with the use of new media technology."

[15] Interestingly, nineteenth-century critics had precisely the same reaction to the volume of popular texts published then: texts could be chosen practically at random, probably uniform in quality and representative of common opinion of the moment (Fyfe 2009).

[16] Reproduction of bibliographic errors in EEBO is discussed in Kichuk (2007), Mak (2011) and Mak (2014).

[17] The DPLA launched in April 2013. In its current form it aggregates metadata of unrestricted online collections, offering, in their own words, "a single point of access to millions of items— photographs, manuscripts, books, sounds, moving images, and more—from libraries, archives, and museums around the United States." The project aims to "make the cultural and scientific heritage of humanity available, free of charge, to all. By adhering to the fundamental principle of free and universal access to knowledge, it will promote education in the broadest sense." (DPLA 2013). It is also interesting to note that Darnton finds it easy to model the project on Google, just a few months before the rejection of the Google Book Search settlement moved the DPLA to the forefront. For background, see Vaidhyanathan (2010).

[18] E.g.: "Most of the scholars to whom we spoke describe their research practice as something that occurs in bursts, during breaks, when they are not teaching, or engaged in administrative

duties. This underscores the multiple time demands placed on scholars, who must often be creative in finding time to spend on serious research. It also explains, at least in part, the attractiveness of digital resources which are readily available; and a reluctance on occasion to consult texts that require a trip to a library or archive, especially if that is some distance away." (Bulger et al. 2011, p. 70).