# The Differences between Digital Humanities and Digital History

STEPHEN ROBERTSON

*The digital humanities is what digital humanists do. What digital humanists do depends largely on academic discipline but also on level of technical expertise. Each discipline, with varying degrees of intensity, has over the years developed a set of favored methods, tools, and interests that, although shared with other disciplines remains connected to the discipline. The task of the digital humanities, as a trans-curricular practice, is to bring these practitioners into communication with each other, and to cultivate a discourse that captures the shared praxis of bringing technologies of representation, computation, and communication to bear on the work of interpretation that defines the humanities.*

*—Rafael Alvarado, University of Virginia, USA, Day of DH 2011*

Digital humanities as it is currently constituted has not erased the differences between academic disciplines. Digital humanists share a commitment to collaboration, openness, and experimentation; a set of software tools, such as Voyant, Palladio, or Omeka; and a group of venues such as THATCamp. But disciplinary sources, questions, and approaches shape their projects, as well as their choice and use of digital tools.[1] The *Companion to Digital Humanities,* the edited collection that has become a convenient touchstone for the emergence of the field that has succeeded humanities computing, reflects these diverse practices. It contains chapters on work in different disciplines, followed by chapters on "Principles," "Applications," and "Production, Dissemination, Archiving" (Schreibman, Siemens, and Unsworth). However, at the time of the *Companion*'s publication in 2004, work with digital technology and media still received little recognition in academic disciplines, and many practitioners held junior positions or were located in the interstices of institutions (Rockwell). In those circumstances, the attraction of an interdisciplinary digital humanities as a professional home overshadowed the disciplinary differences within it. Digital humanities became a big tent (Terras; Svensson).

Much has changed since 2004. The use of digital technology and media are now recognized, if not always embraced, by most of the major humanities professional organizations. Strands of digital sessions now appear in conference programs for major and more specialized conferences: digital history, for example, can be found in the American context not only at the annual meetings of the American Historical Association (AHA) and Organization of American Historians, but also smaller organizations such as the Southern Historical Association, the Urban History Association, and the International Congress on Medieval Studies.[2] The AHA will soon adopt standards for assessing digital scholarship for promotion and tenure, joining the Modern Language Association and the College Art Association.[3]

requiring expertise in digital technology, some even as the pri-
mary field. Undergraduate and graduate courses, certificates, and masters programs in digital humanities are appearing in a diverse range of institutions. At the same time, in a predictable corollary to this growing acceptance, digital humanities has been the target of attacks in scholarly and more popular publications. In almost all cases, what is labeled digital humanities in those critiques is in fact digital literary studies, effectively casting the big tent as housing only a single discipline.[4]

In this moment, our focus needs to shift back to disciplinary difference within digital humanities.[5] The audiences to which we are now seeking to explain our work are increasingly not only administrators and funding agencies whose broad remits echo the breadth of digital humanities, leaving us to elaborate only the digital. Instead, we are engaged more at the local level, with individual scholars, teachers, students, and groups beyond the academy who rarely see their interests and projects in such broad terms. As a result, to them digital humanities is doubly unintelligible, requiring us to persuade them of the credibility and relevance of both the digital and the interdisciplinary. Their willingness to engage with digital humanities is often further reduced by their association of the term with attacks on digital literary studies. A conversation instead about digital history, for example, allows explorations of the specific technologies and approaches that offer the most possibilities for the sources, questions, and teaching of historians. As Ryan Cordell perceptively notes in regards to teaching digital humanities, a "rapid introduction to as many methods and tools as possible" provides only a "glancing understanding of any aspect," whereas "a more focused introduction to a few tools, methods, and theoretical conversations" within a disciplinary focus provides "fewer but more well-developed skills from which they can build" (see chapter 36 in this book on "How Not to Teach Digital Humanities").

Moving forward, we would be better served by reimagining digital humanities not as single all-encompassing tent but as a house with many rooms, different spaces for disciplines that are not silos but entry points and conduits to central spaces where those from different disciplines working with particular tools and media can gather. Each of the many disciplinary rooms would have a distinctive character, reflecting a particular contribution and orientation to the field.

In this chapter, I explore digital history's use of the Web ar
computational tools as two areas that distinguish the disciplir
within digital humanities, and from digital literary studies ... particular. In singling out digital literary studies as a counterpoint, I seek to disrupt the recent tendency in both scholarly and popular discussion to equate that field with digital humanities. Beyond that historically contingent assertion of difference, I aim to redirect attention to academic disciplines as central to what different digital humanists do in order to make clear what the field offers scholars in those disciplines. A common reaction to the argument I make here has been to point to a handful of projects that do not fit the patterns which I identify, and on that basis, dismiss the importance, if not the existence, of disciplines

preserve of particular disciplines and entirely absent from others. Rather, difference is a spectrum of emphasis, with varying degrees of interest in methods, tools, and values. Looking for projects that do not fit the differences I identify is not the test of the argument; rather, the key is establishing which projects are exceptions and which are exemplars.[6]

The most common use to which digital humanists, including some historians, have put the Web has been the distribution and presentation of material to other scholarly researchers. Such a use of the Web in digital literary studies, for example, is clear in the *Companion to Digital Humanities*; audiences other than scholars barely receive a mention. Matthew Kirschenbaum, in exploring the user interface of the *William Blake Archive,* did respond to complaints that it was "not particularly easy to use" and "chock-full" of "scholarly trappings" by noting that "while we are happy to have users from many different constituencies, the site's primary mission has always been expressly conceived as scholarly research" ("So the Colors Cover the Wires"). Perry Willett's discussion of the audience for electronic texts noted similar tensions and likewise put "general readers" to one side (Electronic Texts). An orientation to a scholarly audience continues to characterize the newer generations of projects, such as the *Shelly-Godwin Archive*.[7] Such uses of the Web to present both primary sources and scholarship were also the focus of the account of digital history offered by William Thomas in the *Companion to Digital Humanities* ("Computing and the Historical Imagination"). Like electronic editions and other TEI projects, digitized archives such as *Valley of the Shadow*, the project that Thomas helped lead, or my own *Digital Harlem*, for that matter, were designed to meet the needs and interests of the scholars who created them.[8] As David Parry notes, placing them online made them publicly available, but did not expand the scope of their audience ("Be Online or Be Irrelevant"). They remained accessible, relevant, and useful primarily to those scholars and their colleagues—and to the incidental audiences of genealogists, who search historical databases for individuals, and readers seeking copies of literary texts.[9] As Sheila Brennan reminds us in her contribution to this collection, your audience is those for who you design a project, not those who can find it (see chapter 32, "Public, First").

Notwithstanding Thomas's focus on historians using the Web to reach other scholars, since the earliest days of the Web, historians created online projects for audiences rarely addressed by other digital humanists. Many of the early practitioners of digital history were social historians and radical historians committed to democratizing the creation of the past and to collaborating with teachers to enhance and diversify the history taught in high schools and universities. Their pursuit of those goals gave digital history a distinctive emphasis on using the Web to reach classrooms and the wider public. In the American context, the availability of funding through the U.S. Department of Education's Teaching American History program and the NEH Education Programs and Public Programs facilitated these projects.[10] Professional recognition was also forthcoming: beginning in 2004, the American Historical Association's James

consecutive occasions.[11] In addition, the field of public history has given work with audiences beyond the classroom a place within the historical profession for which there is no equivalent in other disciplines in digital humanities—although it should be noted that that place is not in the mainstream of the profession. Public history grew from a movement in the 1970s to apply history to real-world issues, but in the United States it has only relatively recently gained support and recognition within research-oriented institutions and professional organizations.[12] That digital history is recognized within public history is clear from the NCPH's Outstanding Public History Project, which has been awarded to three online projects since 2008.[13]

The work of the Roy Rosenzweig Center for History and New Media[14] (RRCHNM), founded in 1994, provides examples of how digital historians used the Web to reach both audiences in classrooms and among the wider public. Roy Rosenzweig established the center with the mission of using digital technology and media "to democratize the past—to incorporate multiple voices, reach diverse audiences, and encourage popular participation in presenting and preserving the past."[15] RRCHNM's earliest online projects were collaborations with the American Social History Project created with and for teachers. *History Matters: The U.S. Survey Course on the Web*, launched in 1998, included historical sources like those found in digital archives, but annotated by historians and packaged with online forums with leading historians and teachers, syllabi annotated by teachers, projects created by history students, and interviews with distinguished teachers who shared their strategies and techniques for using the historical sources in the classroom.[16] Other projects explored the French Revolution and highlighted the new fields of world history and the history of childhood and youth, and recent history, such as the events of 1989.[17] *Teachinghistory.org* pursued this work on a particularly large scale, building on and disseminating the results of more than 1,000 *Teaching American History* projects, providing both history content and teaching materials and bringing together different communities concerned with history education.[18] The audience for these sites continues to grow: well into its second decade, and lacking design features now expected on the Web, *History Matters* nonetheless attracts increasing numbers of visitors each year: 2.1 million unique visitors and just over 3 million visits in 2014. That same year, *Teachinghistory.org* saw its audience grow to 1.6 million unique visitors and just over 3 million visits.

RRCHNM also used the Web to develop public history projects involving audiences beyond the classroom. Early projects focused on collecting and preserving the past, in collaboration with a variety of communities. *ECHO* fostered communication and dialogue among historians, scientists, engineers, doctors, and technologists in order to develop and disseminate methods to collect and preserve the recent history of science, technology, and industry.[19] The *September 11 Digital Archive*,[20] a collaboration with the American Social History Project at the CUNY Graduate Center, collected digital material related to the attacks both through a web interface and direct collecting. The *Hurricane Digital Memory Bank* further devel-

ships with more than twenty universities, museums, state hu-
manities councils, media outlets, nonprofit relief groups, and grassroots empowerment groups.[21] In the *Bracero History Archive,* a web-accessible repository facilitated many parties working together to build a collection of oral histories with Bracero guest workers and other types of historical evidence of their experiences, and then present it with supporting materials and teaching resources.[22] *Gulag: Many Days, Many Lives* and *Objects of History* involved collaborations with Russian museums and the National Park Service, and with the National Museum of American History, respectively, to build and present exhibits.[23] While historians also use the Web to distribute and present material to other scholars, the practice of placing material online to reach and collaborate with the wider public, and to reach teachers and students in classroom settings, has become established and professionally recognized within history to an extent that distinguishes digital history within digital humanities.

Digital history is distinctive within digital humanities in a second area—its use of computational tools. This distinction is less intense than that in the use of the Web. Historians can be found using mapping, text analysis, and network software, the three types of computational tools favored in the digital humanities, as well as 3D modeling, image analysis, and social media. However, digital historians have turned to digital mapping to a greater extent than other disciplines in the digital humanities, adopting it as their favored computational tool. For example, of the list of digital humanities projects maintained by John Levin and adopted by the GeoHumanities Special Interest Group of ADHO, almost two-thirds are historical projects.[24] It is certainly the case that digital mapping is not as central to history as it is to archaeology, but mapping has long been a part of record keeping the practices of archaeology. The turn to digital mapping in history is not continuous with older practices to the same extent. If maps have appeared in historical scholarship, historians have until recently written about the past with little attention to space, notwithstanding the discipline's definition around nations and cities.

The distinctive prominence of mapping in digital history is a product of both the spur to mapping that the spatial turn provided to historians and of the more limited availability, until recently, of historical material to which text analysis tools could be applied. In the last decade, historians have given renewed attention to space and place as part of spatial turn in the humanities—which began in the 1970s with a focus on perceptions and representations, that, as Karen Haltunen put it, "tended to the metaphorical" and employed the "idiom of borders and boundaries, frontiers and crossroads, centers and margins," and extended in the early 2000s to a concern with "spatial issues more materially" (2).[25] It was this more recent development that provided a spur to mapping; visualization was not as necessary a part of exploring spatial perceptions and language.

At the same time, mapping tools became more accessib
and useful to historians seeking to extend the spatial turn to th
material. The first computational mapping tools, GIS softwar,

plex, and cumbersome and operate only on computers running Windows. Moreover, this software was designed for precise, quantitative data that could be parsed in highly structured tabular databases and cartographic maps that emphasized generalization. It fit uneasily with the sources that most historians use, which are characterized by ambiguity, uncertainty, uniqueness, and with the discipline's move away from statistical analysis after the 1970s. A small group of historians did adopt GIS software and continue to work within its limitations, under the umbrella of the Social Science History Association and its journal, *Social Science History,* and largely disconnected from digital history. In the last decade, the development of the geospatial web has made mapping available and accessible to digital historians, with web mash-ups allowing maps to be created and iterated with unprecedented ease. Web mapping platforms such as Google Maps, MapQuest, CartoDB, Palladio, and Neatline do not have the same quantitative orientation as GIS software, allowing a diversity of spatial data to be visualized.26 More advanced and customized web mapping is also now possible using open-source tools such Leaflet, GeoServer, d3.js, and OpenLayers. With these tools, historians explored a range of questions and topics: urban development, housing, ghettos, moviegoing, language use, sound, industries ranging from iron to cutlery, events such as bombing, military recruitment, the slave trade, slave revolts, and gaining freedom from slavery, the spread of disease, witchcraft accusations, feminist activism, and the circulation of correspondence.27 These tools are obviously available to all digital humanists, and other disciplines have also experienced the spatial turn. So, while those factors help explain the prominence of mapping in digital history, they do not explain its relatively lesser place in the practice of other disciplines within digital humanities.

Digital historians have often used the geospatial web as a platform for presenting historical sources. Maps allow the visual display of information, and for the organization and integration of different sources on the basis of their shared geographic location. The first two winners of the AHA's Roy Rosenzweig Prize for Innovation in Digital History28 were mapping projects built on Google Maps: *Digital Harlem: Everyday Life, 1915–1930* (which I created with collaborators at the University of Sydney) and Bobby Allen's *Going to the Show*. *Digital Harlem* offered a visualization of everyday life through maps of locations, events and individuals' lives, using information extracted from legal records, newspapers, and other published and archival texts.29 *Going to the Show* mapped the locations of movie venues in North Carolina between 1896 and 1922, combined with an archive of related material.30 More recent examples include: *Mapping the Republic of Letters*,31 which traces the circulation of correspondence among enlightenment intellectuals using prototype browser-based tools now being developed as an open-source platform, Palladio; *The Roaring Twenties*, which displays the location of sounds and noise complaints in 1920s New York City, initially using Google Maps, before shifting to MapQuest; and *Photogrammar,* which located 170,000 pho-

Historians have also used web maps that display historical data as part of their analysis of those sources. The interactive, iterative features of digital maps make them research tools, a means of discovering as well as displaying knowledge. Not only are mapped sources placed in their geographic contexts, but selections of those sources can be mapped, different layers of sources can be juxtaposed, and the scale can be zoomed from the level of individual buildings out to neighborhoods, cities, and regions. Viewing these maps can reveal spatial patterns not evident from reading the texts, relationships that facilitate comparisons and prompt questions. The answers to those questions are not on the maps, but in the sources from which they are derived. So mapping a range of sources using *Digital Harlem*,[33] for example, highlighted a more extensive and expansive white presence in the neighborhood, leading to questions about the nature of a range of interracial encounters, when they occurred, and the extent to which life in this black neighborhood could be lived apart from whites (Robertson, White, and Garton). Similarly, *Visualizing Emancipation* mapped the position of Union troops and a range of emancipation events drawn from *Official Records of the War of the Rebellion* and other sources to trace the spread of emancipation during the Civil War and explore the relationship between the presence of troops and enslaved men and women's escape from bondage (Ayers and Nesbit).[34]

Less often, historians have created maps to illustrate arguments or present the answer to questions, forgoing some interactivity and ability to query the data in order to highlight a particular analysis. *Slave Revolt in Jamaica, 1760–1761*, for example, offers an animated narrative map of a slave revolt built using Leaflet. That map allows the user to move within the timeline, but in no other way to affect the map. *Mapping Occupation*[35] presented a spatial narrative that coupled a text with a series of maps to explore the Union Army's occupation of the Reconstruction South. Icons on some of maps can be clicked to reveal the underlying data and the scale can be adjusted, but otherwise the maps are limited to illustrating a particular argument.[36]

The relative prominence of mapping in digital history is also a product of historians' limited use of the textual analysis tools. It is not the case that historians have been unable to see ways to use text mining and topic modeling in their research. To the contrary, historians were among the first digital humanists to experiment with such tools. Dan Cohen[37] made early use of data from Google books, and he went on to be part of the *With Criminal Intent*[38] project that data-mined the Old Bailey trials (*Searching for the Victorians*; *Data Mining with Criminal Intent*). Sharon Block[39] and Rob Nelson employed topic modeling. Block explored the eighteenth-century *Pennsylvania Gazette* ("Doing More with Digitization"). Nelson explored the Confederacy's paper of record during the Civil War (*Mining the Dispatch*[40]). But until recently, little work has built on those experiments.

One explanation for the relatively limited use of text analysis tools by historians is a lack of digitized sources. As Dan Cohen noted in 2008, "Unless we can have machines scan, sort, and apply digital techniques to the full texts of documents, we can't

have been digitized and made machine-readable on a far lesser scale than those on which literary studies relies. In part, this reflects the fact that historians rely more on documents that are not published or printed. Making what is digitized accessible to computational tools is a more expensive and time-consuming task in the case of those sources than with published sources. The optical character recognition (OCR) software that extracts text from images currently only works with printed texts, meaning that handwriting must be transcribed. Although the *Proceedings of the Old Bailey*, one of the best-known and most widely used digitized collections, used manual keying as part of its digitization process, transcription is beyond the resources of most institutions and employed largely only through crowdsourcing.[41] Much of the early digitization of historical sources was highly selective, chosen from larger collections to answer specific research questions or to make available well-known or popular documents.[42] Both forms of collection have limited attraction to a wider audience of researchers, especially given the premium the profession places on original research.[43] Contrast the use of the *Valley of the Shadow*[44] and the *Proceedings of the Old Bailey*:[45] the former is a pioneering virtual archive of sources selected by Edward Ayers for a comparative investigation of two communities in American Civil War; the latter is a collection of all the surviving editions of the published accounts of trials that took place at the Old Bailey. Although the *Valley* project produced a path-breaking digital article in the *American Historical Review* and a prize-winning book, and has been widely used in teaching, I could identify only one other scholar who used it for research.[46] As of July 2015, by contrast, 401 publications cite the *Proceedings*.[47]

In addition, many of the historical sources that are digitized are only accessible by search, in databases that do not have an API or other means of accessing the complete data. This is particularly true of the commercial databases that contain digitized newspapers, a published source widely relevant to historical research and less so to other fields in digital humanities. Sixteen of the most important American newspapers and nine of the leading African American newspapers, for example, can only be found in *Proquest Historical Newspapers*. While those databases allow researchers to search text generated by OCR, the results that are returned are only images of the page, containing no machine-readable text. Not only is there no means of bulk export, but the licensing agreements expressly prohibit data mining. Both Proquest and Gale Cengage have announced plans to open up their databases to computational analysis, but it remains unclear at what cost and by what means.[48] Peter Leonard and Lindsay King did obtain access to Proquest's Vogue Archive and their project, *Robots Reading Vogue*, shows what would be possible with access to these collections, combining an Ngram viewer, topic modeling, a frequency analysis for advertising, and a visual analysis of magazine covers. Those visualizations link to articles in *Vogue,* providing an alternative to search as a gateway into the archive—assuming that you are affiliated with an institution that subscribes to the Proquest database.[49]

encouraged historians and other humanities scholars to be simply consumers of digital content, to accept the search interfaces provided by vendors as the means by which to conduct research in digitized sources rather than looking to tools like topic modeling software. Studies of research practices have shown that many humanities scholars rely on search, beginning with Google search to identify sources, and proceeding to full text keyword searching to research within digitized collections (Rutner and Schonfeld; Kemman, Kleppe, and Scagliola; Chassanoff). So naturalized has search become that few recognize that it is not, as Ted Underwood points out, "a finding aid analogous to a card catalog . . . [but] a name for a large family of algorithms that humanists have been using for several decades to test hypotheses and sort documents by relevance to their hypothesis" (65). It can be a powerful method, disrupting the hierarchies and categories of information established in the past. Full-text search examines every publication in a database, not simply those that have become canonical, and every word in a publication, working from the bottom up rather than from the top down—for example, from the journal to the issue to the article, as a researcher in a library would. In the case of newspapers, full-text search checks many words that researchers likely would not have read when browsing, where their focus is on particular sections and headlines, not the advertisements, schedules, and notices that make up the bulk of most papers. In the case of archival collections, search can remove information from the context of the institution that structured the collection; it can, in Tim Hitchcock's words, de-center institutions in favor of individuals ("Digital Searching"). But search is also a limited method. Search struggles to deal with what lies outside a set of results. In returning only the terms one enters, a search filters out any alternative hypotheses. For historians, this poses particular challenges, as the language and ways of organizing knowledge in the past often differ significantly from contemporary terms and patterns of thought. If we use the wrong search terms, we literally misread our sources. And working with interfaces that tell us how many results we found without reference to how many results were possible, it is not always clear just how significant those results might be. In the case of most newspaper databases, search also returns individual stories, removed from their place on the page and in newspaper as a whole. In other words, search radically decontextualizes the results it produces.[50] Text mining and topic modeling can offer a different perspective that addresses those limits and complements search and other approaches to historical research.

Recognition of the value of topic modeling as a complement to search is one of several factors that have led to more use of text analysis by historians. Digitized historical sources open to text mining and topic modeling are also beginning to become more available. In the U.S. context, a key development is the Library of Congress' ongoing *Chronicling America* project.[51] Now approaching 9.5 million pages, this collection of newspapers from across the nation is not just freely available, but is also equipped with an API, allowing researchers to extract large datasets. However, although *Chronicling America* contains a

same date range as *ProQuest Historical Newspapers*, being limited to the years before 1923 that are out of copyright. And the more limited resources of this project means that the OCR-generated text is uncorrected and consequently marred by numerous errors.[52] OCR accuracy is greater in commercial products, but just how much more accurate is impossible to know, as the companies do not release that information, notwithstanding how important it is to assessing the significance of search results.

A small number of historians have already text mined *Chronicling America*. As part of their *Mapping Texts* project, Andrew Torget and Jon Christensen used word counts, named entity extraction, and topic modeling to examine language patterns in Texas newspapers.[53] In a more focused study, Cameron Blevins used named-entity extraction with two of those Texas newspapers to explore what they revealed about the imagined geography of the United States in a period understood as one of integration and incorporation. What he found and reported in the first digital history article to appear in the *Journal of American History* ("Space, Nation, and the Triumph of Region"[54]) was that, at odds with the prevailing view, the newspaper was focused on region, not nation.[55] In *An Epidemiology of Information: Data-mining the 1918 Flu Pandemic*,[56] E. Thomas Ewing and others from Virginia Tech employed topic modeling and segmentation and tone classification on a set of twenty newspapers to explore the transmission of disease-related information (Ewing et al.). *Chronicling America* is also serving as a basis for research projects in literary studies. Ryan Cordell's *Viral Texts* project traces networks reprinting texts in nineteenth-century newspapers.[57] Elizabeth Lorang and others are developing an image classifier to identify poetry in newspapers, in order to explore the magnitude of poetic content in the press (Lorange et al.). And Lauren Klein and her collaborators at Georgia Tech are using abolitionist newspapers as the case study in their project to build *TOME*, a tool to visualize topic-modeling results (Klein).

Beyond work with newspapers, other noteworthy historical text analysis projects include two explorations of the *Digital National Security Archive* that are grappling with an even more fragmentary source and bringing digital methods to the field of diplomatic history, one of the discipline's most methodologically conservative fields. Micki Kaufman's multi-award-winning "'Everything on Paper Will Be Used Against Me': Quantifying Kissinger," employs a variety of text analysis methods to explore Henry Kissinger's memoranda and teleconference transcripts.[58] The *Declassification Engine* project, led by Matthew Connelly and David Madigan, looks at the full range of documents in the archive to explore official secrecy through analysis of redacted text and declassified documents.[59] Michelle Moravec's work using text analysis tools in women's history is important not only for bringing digital history to another field where it has been slow to find a place, women's history, but also as an example of the kind of work an individual scholar can do without a formal technical background or extensive institutional or grant support. Among other projects, Moravec has used corpus linguistics to explore gender in the *History of Women's Suffrage* —Elizabeth

(Under this name she is fully described). Molavec's current project, *The Politics of Women's Culture,* also features text analyses of Women's Liberation Movement periodicals to explore networks within the movement.60

Building on this work, responding to the emerging opportunities, and working to obtain more access and additional digitized sources will require historians to confront their reliance on search as a research method and its limits as an approach to understanding digitized sources. Scholars that are simply searching collections with an API like the *Proceedings of the Old Bailey* or *Chronicling America* as part of their research—and many are, judging by the character of publications citing the *Proceedings of the Old Bailey* and *Chronicling America*—are missing an opportunity to put those search results into context.61 The need to contextualize searches applies even to the use of digitized collections to recover the history of ordinary individuals, an important dimension of what the era of "big data" offers historians (Hitchcock, "Big Data, Small Data"). Recent text analysis projects, together with these imperatives, make it likely that the limited use of text analysis tools that currently distinguishes digital historians will become less marked in the future.

Signs of a trend toward text analysis are a reminder that disciplinary differences in digital humanities are not fixed or immutable. They grow from the interests and questions that characterize a discipline at a given time, and which attract awards and recognition, and from the sources used to explore those topics, and which have been digitized and made machine-readable and available to researchers. All those elements are subject to change. Notwithstanding that fluidity, at any given time disciplinary differences are present. Discussions of digital humanities that do not recognize the different sources, approaches, and questions of humanities disciplines limit the audience they will engage and can even alienate humanities scholars who do not see their field represented. In so doing, they create a barrier to interest and to entry to the field. It does not matter how big the tent is if the entrance is narrow. To engage more scholars in digital humanities requires approaching them in disciplinary terms and showing how the values and practices of digital humanities can help them pursue their disciplinary projects. Discussions with historians, for example, need to begin with digital mapping and to present text analysis in relation to the historical sources accessible to computational tools, and to the other methods, such as search, that scholars have been using to explore those sources. Once scholars can see the value of digital humanities to their discipline, once they have entered not a big tent, but a room devoted to their field within a house shared with other disciplines, then they are in a position to engage with digital humanities as an interdisciplinary field. As Tom Scheinfeldt argues, "Understanding what makes us distinctive will help us better see what in our practices may be of use to our colleagues in other disciplines and to see more clearly what they have to offer us" ("The Dividends of Difference"). This chapter represents an effort to understand the distinctive features of digital history, arguing that historians are different in the way they use the Web and in the computational tools they

disciplines distinctive within the digital humanities.

## Notes

1. For an introduction to the values of the digital humanities, see Spiro, "This Is Why We Fight." For an exploration of debates over the extent to which those values are present in digital humanities, see Kirschenbaum, "What Is "Digital Humanities."

2. Southern History Association, http://sha.uga.edu/2013%20Program%20for%20web.pdf; Urban History Association, http://uha.udayton.edu/2012Conf/UHA_Conference_2012_Complete_Oct_12_2012.pdf; International Congress on Medieval Studies, http://scholarworks.wmich.edu/cgi/viewcontent.cgi?article=1051&context=medieval_cong_archive.

3. Ad Hoc Committee on Professional Evaluation of Digital Scholarship by Historians, "Guidelines for the Professional Evaluation of Digital Scholarship in History," April 2015, http://historians.org/teaching-and-learning/current-projects/committee-on-professional-evaluation-of-digital-scholarship-in-history; MLA, "Guidelines for Evaluating Work in Digital Humanities and Digital Media," 2012, https://www.mla.org/guidelines_evaluation_digital; College Art Association, http://www.collegeart.org/news/2014/10/29/mellon-foundation-awards-grant-to-caa-to-partner-with-sah-on-digital-scholarship-guidelines/.

4. The blog post on which this chapter is based was written in response to "In the Shadow of the Digital Humanities," a special issue of *Differences* 25, no. 1 (2014), and Adam Kirsch's article in the May 2014 issue of *The New Republic,* "Technology Is Taking over English Departments: The False Promise of Digital Humanities." What was striking to me at the time was the almost complete absence of any mention of digital history in discussions purportedly about the field of digital humanities. See Robertson, "The Differences between Digital History and Digital Humanities."

5. For two other recent calls to make this switch, see Scheinfeldt, "The Dividends of Difference," as well as Ryan Cordell's "How Not to Teach Digital Humanities," chapter 36 in this volume.

6. That is not a straightforward task. We lack comprehensive compilations of digital humanities projects. There are numerous lists that highlight small numbers of projects as examples of particular approaches or fields within digital humanities, encouraging the tendency to focus on such projects without a sense of how they reflect what is being done in digital humanities as a whole.

7. Shelly-Godwin Archive, http://shelleygodwinarchive.org/about.

8. *Valley of the Shadow: Two Communities in the American Civil War,* http://valley.lib.virginia.edu; *Digital Harlem: Everyday Life, 1915–1930*, http://digitalharlem.org.

9. On readers as audiences for electronic editions, see Willet. On searching for names as an incidental use of *The Proceedings of the Old Bailey, 1674–1913,* see "Getting Started," http://www.oldbaileyonline.org/static/GettingStarted.jsp.

10. For digital historians work with teachers and students, see Brier, "Where's the Pedagogy?"; Robertson, "CHNM's Histories: Collaboration in Digital History"; Robertson, "CHNM's Histories: Digital History & Teaching History"; and Dorn, "Is (Digital) History More than an Argument about the Past?"

12. For a brief definition of public history, see National Council on Public History (NCPH), "What Is Public History?" http://ncph.org/cms/what-is-public-history/. The two major American historical organizations only adopted guidelines on evaluating public historians for tenure and promotion in 2010. See Working Group on Evaluating Public History Scholarship, "Tenure, Promotion, and the Publicly Engaged Academic Historian: A Report."

13. NCPH Outstanding Public History Project Award, http://ncph.org/cms/awards/public-history-project-award/.

14. http://chnm.org.

15. About RRCHNM, http://chnm.gmu.edu/about/. Grant proposals and reports for RRCHNM projects can be found at *RRCHNM20,* http://20.rrchnm.org.

16. More About History Matters, http://historymatters.gmu.edu/expansion.html.

17. *Liberty, Equality, Fraternity: Exploring the French Revolution,* http://chnm.gmu.edu/revolution/; World History Matters, http://worldhistorymatters.org/; *Children and Youth in History,* http://chnm.gmu.edu/cyh/; *Making the History of 1989,* http://chnm.gmu.edu/1989/.

18. About the Project, http://teachinghistory.org/about.

19. *ECHO: Exploring and Collecting History Online,* http://echo.gmu.edu; grant proposals and reports at *RRCHNM20.*

20. http://911digitalarchive.org.

21. *The September 11 Digital Archive,* 911digitalarchive.org; *Hurricane Digital Memory Bank,* http://hurricanearchive.org.

22. *Bracero History Archive,* http://braceroarchive.org.

23. *Gulag: Many Days, Many Lives,* http://gulaghistory.org; *Object of History,* http://objectofhistory.org.

24. "DH GIS Projects," *Anterotesis,* http://anterotesis.com/wordpress/mapping-resources/dh-gis-projects/; "Humanities GIS Projects," *GeoHumanities,* http://geohumanities.org/gis. Of the 143 projects in the list that involve more than the digitization of maps, 92 are historical projects. I did not include in the count of historical projects five projects related to book publishing that straddle the boundary between literary and historical studies. The list includes all the digital mapping projects that I have been able to identify.

25. For the spatial turn, see also Guldi, "What Is the Spatial Turn?"; Robertson, "Putting Harlem on the Map"; White, "What Is Spatial History."

26. For a more detailed, insightful guide to spatial humanities and its relation to GIS, see Bodenhamer, Corrigan, and Harris, *The Spatial Humanities.*

27. "DH GIS Projects," http://anterotesis.com/wordpress/mapping-resources/dh-gis-projects/.

28. http://www.historians.org/awards-and-grants/past-recipients/roy-rosenzweig-prize-recipients.

29. *Digital Harlem: Everyday Life, 1915–1930,* http://digitalharlem.org.

30. *Going to the Show,* http://docsouth.unc.edu/gtts/.

31. http://republicofletters.stanford.edu/index.html.

33. http://digitalharlem.org.

34. See also *Visualizing Emancipation,* http://dsl.richmond.edu/emancipation/.

35. http://mappingoccupation.org/index.html.

36. *Slave Revolt in Jamaica,* http://revolt.axismaps.com; *Mapping Occupation,* http://mappingoccupation.org.

37. http://www.dancohen.org/2010/10/04/searching-for-the-victorians/.

38. http://criminalintent.org.

39. http://www.common-place.org/vol-06/no-02/tales.

40. http://dsl.richmond.edu/dispatch/pages/home.

41. See "Technical Methods," Old Bailey Online, http://www.oldbaileyonline.org/static/Project.jsp#methods. The largest and longest-running crowdsourced transcription project is Transcribe Bentham, http://blogs.ucl.ac.uk/transcribe-bentham/. In the United States, the Smithsonian has a large transcription project; https://transcription.si.edu. On a smaller scale, *Papers of the War Department,* a project of RRCHNM, uses Scripto, a plugin for Omeka; see http://wardepartmentpapers.org/transcribe.php. That open-source platform allows smaller institutions to undertake crowdsourced transcription.

42. See also the other problems in the British case discussed by Tim Hitchcock in "Digitising British History since 1980."

43. For an example of historians' concern about the completeness of digital collections, see Chassanoff, "Historians and the Use of Primary Sources in the Digital Age."

44. http://valley.lib.virginia.edu.

45. http://www.oldbaileyonline.org/.

46. A search of Google Scholar identified only one article reporting research using the project. Ayers and his colleague Scott Nesbit have made use of the *Valley* in another project: *Visualizing Emancipation*.

47. "Old Bailey Proceedings Online: Citations Bibliography," https://www.zotero.org/groups/old_bailey_proceedings_online_citations_bibliography.

48. "Gale Leads to Advance Academic Research by Offering Content for Data Mining and Textual Analysis," news release, November 17, 2014, http://news.cengage.com/library-research/gale-leads-to-advance-academic-research-by-offering-content-for-data-mining-and-textual-analysis/. See also the much more ambivalent response of representatives of Cengage and D. H. Thomson at the seminar on "Mining digital repositories" at the Dutch Koninklijke Bibliotheek in 2014: http://blog.kbresearch.nl/2014/04/13/how-to-maximise-usage-of-digital-collections/.

49. *Robots Reading Vogue,* http://dh.library.yale.edu/projects/vogue/.

50. I have found two other accounts of the impact of search as a historical research method particularly useful: see Mussell, "Doing and Making: History as Digital Practice," and Nicholson, "The Digital Turn."

51. *Chronicling America,* http://chroniclingamerica.loc.gov.

52. The exemplar in this area is *Trove,* which currently makes freely available almost 900 Australian newspapers and over 17 million pages, and benefits from less restrictive copyright law in being able to include content published up to 1954. It provides an API, allows users to tag and

53. "Assessing Language Patterns: A Look at Texas Newspapers, 1829–2008," http://language.mappingtexts.org.

54. http://jah.oxfordjournals.org/cgi/content/full/jau184?ijkey=unucsImiwNrelaF&keytype=ref.

55. See also Blevins, "Mining and Mapping the Production of Space," an accompanying essay to the *Journal of American History* article.

56. http://vtechworks.lib.vt.edu/bitstream/handle/10919/46991/An%20Epidemiology%20of%20Information%20Project%20Research%20Report_Final.pdf?sequence=1.

57. About the Viral Texts Project, http://viraltexts.org.

58. *"Everything on Paper Will Be Used Against Me:" Quantifying Kissinger,* http://blog.quantifyingkissinger.com.

59. *The History Lab,* http://www.history-lab.org.

60. Digital History, http://michellemoravec.com/122-2/.

61. For publications citing *Chronicling America,* see NEH, "National Digital Newspaper Program Impact Study 2004–2014."

## Bibliography

Ayers, Edward L., and Scott Nesbit. "Seeing Emancipation: Scale and Freedom in the American South." *Journal of the Civil War Era* 1, 1 (March 2011): 3–24.

Blevins, Cameron. "Mining and Mapping the Production of Space: A View of the World from Houston," 2014. http://web.stanford.edu/group/spatialhistory/cgi-bin/site/pub.php?id=93.

—. "Space, Nation, and the Triumph of Region: A View of the World from Houston." *Journal of American History* 101, no. 1 (June 2014): 122–147.

Block, Sharon. "Doing More with Digitization: An Introduction to Topic Modeling of Early American Sources." *Common-Place* 6, no. 2 (January 2006). http://www.common-place.org/vol-06/no-02/tales/.

Bodenhamer, David, John Corrigan, and Trevor Harris, eds. *The Spatial Humanities: GIS and the Future of Humanities Scholarship.* Bloomington: Indiana University Press, 2010.

Brier, Stephen. "Where's the Pedagogy? The Role of Teaching and Learning in the Digital Humanities." In *Debates in Digital Humanities,* ed. Matthew K. Gold. Minneapolis: University of Minnesota Press, 2012. http://dhdebates.gc.cuny.edu/debates/text/8.

Chassanoff, Alexandra. "Historians and the Use of Primary Sources in the Digital Age." *The American Archivist* 76, no. 2 (2013): 470–471.

Cohen, Dan. "Searching for the Victorians." October 4, 2010. http://www.dancohen.org/2010/10/04/searching-for-the-victorians/.

Cohen, Dan, Michael Frisch, Patrick Gallagher, Steven Mintz, Kirsten Sword, Amy Murrell Taylor, William G. Thomas III, and William J. Turkel. "Interchange: The Promise of Digital History." *Journal of American History* 95, no. 2 (2008): 442–451.

Cohen, Dan, Tim Hitchcock, and Geoffrey Rockwell, et al. "Data Mining with Criminal Intent: Final White Paper," August 31, 2011. http://criminalintent.org/wp-content/uploads/2011/09/Data-Mining-with-Criminal-Intent-Final1.pdf.

http://quod.lib.umich.edu/d/dh/12230987.0001.001/.

Ewing, E. Thomas, Samah Gad, Bernice L. Hausman, Kathleen Kerr, Bruce Pencek, and Naren Ramakrishnan. *An Epidemiology of Information: Datamining the 1918 Flu Pandemic*. Project Research Report, April 2, 2014. http://vtechworks.lib.vt.edu/bitstream/handle/10919/46991/An%20Epidemiology%20of%20Information%20Project%20Research%20Report_Final.pdf?sequence=1.

Guldi, Jo. "What Is the Spatial Turn?" *Spatial Humanities,* 2011. http://spatial.scholarslab.org/spatial-turn/.

Haltunen, Karen. "Groundwork: American Studies in Place—Presidential Address to the American Studies Association, November 4, 2005." *American Quarterly* 58, no. 1 (March 2006): 1–15.

Hitchcock, Tim. "Big Data, Small Data and Meaning," *Historyonics* (blog), November 9, 2014. http://historyonics.blogspot.com/2014/11/big-data-small-data-and-meaning_9.html.

—. "Digital Searching and the Re-formulation of Knowledge." In *The Virtual Representation of the Past,* ed. Mark Greengrass and Lorna Hughes, 81–90. London, Ashgate, 2008.

—. "Digitising British History since 1980." *Making History: The Changing Face of the Profession in Britain*. Institute for Historical Research, 2008. http://www.history.ac.uk/makinghistory/resources/articles/digitisation_of_history.html.

Kemman, Max, Martijn Kleppe, and Stef Scagliola. "Just Google It." In *Proceedings of the Digital Humanities Congress 2012,* ed. Clare Mills, Michael Pidd, and Esther Ward. Sheffield: HRI Online Publications, 2014. http://www.hrionline.ac.uk/openbook/chapter/dhc2012-kemman.

Kirschenbaum, Matthew G. "'So the Colors Cover the Wires': Interface, Aesthetics, and Usability." In *A Companion to Digital Humanities,* ed. Susan Schreibman, Ray Siemens, John Unsworth. Oxford: Blackwell, 2004. http://www.digitalhumanities.org/companion/.

—. What Is "Digital Humanities" and Why Are They Saying Such Terrible Things about It?" *Differences* 25, 1 (2014): 46–63.

Klein, Lauren. "Talk at Digital Humanities 2014," DH Lab, July 24, 2014. http://dhlab.lmc.gatech.edu/news/talk-at-digital-humanities-2014/.

Lorang, Elizabeth, Leen-Kiat Soh, Maanas Varma Datla, Spencer Kulwicki. "Developing an Image-Based Classifier for Detecting Poetic Content in Historic Newspaper Collections." *D-Lib Magazine* 21, 7/8 (July/August 2015). http://www.dlib.org/dlib/july15/lorang/07lorang.html#n18.

Moravec, Michelle. "'Under this name she is fitly described': A Digital History of Gender in the *History of Woman Suffrage*," March 2015. http://womhist.alexanderstreet.com/moravec-full.html#en7.

Mussell, James. "Doing and Making: History as Digital Practice." In *History in the Digital Age,* ed. Toni Weller, 79–94. London: Routledge, 2013.

National Council on Public History. "What Is Public History?" http://ncph.org/cms/what-is-public-history/.

National Endowment for the Humanities. "National Digital Newspaper Program Impact Study 2004–2014," September 2014. http://www.neh.gov/files/divisions/preservation/ndnp_report_2014_0.pdf.

Nelson, Robert. *Mining the Dispatch*. Digital Scholarship Lab, University of Richmond. http://dsl.richmond.edu/dispatch/pages/home.

http://academhack.outsidethetext.com/home/2010/be-online-or-be-irrelevant/.

Robertson, Stephen. "CHNM's Histories: Collaboration in Digital History." Author's blog, October 14, 2014. http://drstephenrobertson.com/blog-post/chnms-histories-collaboration-in-digital-history/.

—. "CHNM's Histories: Digital History & Teaching History." Author's blog, October 27, 2014. http://drstephenrobertson.com/blog-post/digital-history-teaching-history/.

—. "The Differences between Digital History and Digital Humanities," Author's blog, May 23, 2014. http://drstephenrobertson.com/blog-post/the-differences-between-digital-history-and-digital-humanities/.

—. "Putting Harlem on the Map." In *Writing History in the Digital Age,* ed. Kristen Nawrotzki and Jack Dougherty. Ann Arbor: University of Michigan Press, 2013. http://quod.lib.umich.edu/d/dh/12230987.0001.001/.

Robertson, Stephen, Shane White, and Stephen Garton. "Harlem in Black and White: Mapping Race and Place in the 1920s." *Journal of Urban History* 39, no. 5 (2013): 864–880.

Rockwell, Geoffrey. "Inclusion in the Digital Humanities." *Philosophi.ca,* September 7, 2011. http://www.philosophi.ca/pmwiki.php/Main/InclusionInTheDigitalHumanities.

*RRCHNM20.* Roy Rosenzweig Center for History and Media (RRCHNM) 20th Anniversary Conference, George Mason University, November 14–15, 2014. http://20.rrchnm.org.

Rutner, Jennifer, and Roger Schonfeld. "Supporting the Changing Research Practices of Historians (ITHAKA S+R, 2012)," December 10, 2012. http://www.sr.ithaka.org/sites/default/files/reports/supporting-the-changing-research-practices-of-historians.pdf.

Scheinfeldt, Tom. "The Dividends of Difference: Recognizing Digital Humanities' Diverse Family Tree/s." *Found History,* April 7, 2014. http://foundhistory.org/2014/04/the-dividends-of-difference-recognizing-digital-humanities-diverse-family-trees/.

Schreibman, Susan, Ray Siemens, John Unsworth, eds. *A Companion to Digital Humanities.* Oxford: Blackwell, 2004. http://www.digitalhumanities.org/companion/.

Spiro, Lisa. "'This Is Why We Fight': Defining the Values of the Digital Humanities." In *Debates in Digital Humanities,* ed. Matthew K. Gold. Minneapolis: University of Minnesota Press, 2012. http://dhdebates.gc.cuny.edu/debates/text/13.

Svensson, Patrik. "Beyond the Big Tent." In *Debates in Digital Humanities,* ed. Matthew K. Gold. Minneapolis: University of Minnesota Press, 2012. http://dhdebates.gc.cuny.edu/debates/text/22.

Terras, Melissa. "Peering inside the Big Tent: Digital Humanities and the Crisis of Inclusion." Author's blog, July 26, 2011. http://melissaterras.blogspot.com/2011/07/peering-inside-big-tent-digital.html.

Thomas, William. "Computing and the Historical Imagination." In *A Companion to Digital Humanities,* ed. Susan Schreibman, Ray Siemens, and John Unsworth. Oxford: Blackwell, 2004. http://www.digitalhumanities.org/companion/.

Underwood, Ted. "Theorizing Research Practices We Forgot to Theorize Twenty Years Ago." *Representations* 127, 1 (Summer 2014): 64–72.

Willett, Perry. "Electronic Texts: Audiences and Purposes." In *A Companion to Digital Humanities,* ed. Susan Schreibman, Ray Siemens, and John Unsworth. Oxford: Blackwell, 2004. http://www.digitalhumanities.org/companion/.

Working Group on Evaluating Public History Scholarship. "Tenure, Promotion, and the Publicly Engaged Academic Historian: A Report." AHA Perspectives, September 2010. http://www.historians.org/perspectives/issues/2010/1009/1009new3.cfm.

## API and Atom Feed

Data about this text is available via read-only JSON API endpoints: sentences, annotations, comments, and index keywords.

Comments posted on this text can be followed by subscribing to this text's Atom Feed.