Laying the Foundation

White, John W., Gilbert, Heather

Published by Purdue University Press

→ For additional information about this book

https://muse.jhu.edu/book/48001

2 | A History of History through the Lens of Our Digital Present, the Traditions That Shape and Constrain Data-Driven Historical Research, and What Librarians Can Do About It[1]

James Baker

INTRODUCTION

Historians have a long and often fraught relationship with numbers. None other than the great French Annalist historian Fernand Braudel acknowledged in 1967 that his methods—temporal and spatial extrapolation of demographic data that enabled him to estimate undocumented population sizes, to grapple with history in the *longue durée*—were controversial. "Historians accustomed to accept only things proved by irrefutable documentation," he wrote, "quite justifiably find these uncertain methods disturbing. Statisticians share neither their misgivings nor their timidity."[2] For although Braudel's historian peers were adept at telling stories across broad sweeps of history, not all were comfortable with statistical representations of past phenomena that seemed divorced from primary sources, that seemed incompatible with the narratives of great men and their institutions whose histories remained in vogue.

Braudel was no prophet, and yet his observations do extrapolate beyond his own temporal surroundings, his very own *histoire événementielle*. Historians today have the opportunity to use long runs of messy textual data, reconstructed models of places and spaces, and tools repurposed from computational and engineering environments to explore past phenomena. For example, by using a process called optical character recognition (OCR), heritage institutions and commercial publishers alike have made millions of pages and billions of words searchable in ways hitherto impossible and unthinkable. This has been an extraordinary boon for scholars. And yet the files created during this process, typically in Extensible Markup

Language (XML) and archival image formats, are never facsimiles of the original source material. Rather their verisimilitude to the traces of the past they seek to capture—the text on a page, the form of that page—can vary wildly depending on a variety of sociotechnological factors. So for all that we librarians do to promote their use and their potential to make a radical intervention in the narratives humanists tell, historians might well find disturbing—and with some justification—the use of these files at scale as a means of exploring past phenomena, just as—by Braudel's reckoning—historians did five decades ago with respect to statistical analysis.

To these concerns we shall return, for simultaneously and largely unperturbed an efflorescence of digital history has taken place. A decade of pioneering work by Tim Hitchcock and Bob Shoemaker on *Old Bailey Online*, *London Lives*, *Connected Histories*, and *Locating London's Past* has brought structured and unstructured humanities data to new audiences, and new audiences to data-driven and computational approaches to historical problems.[3] In turn, this has driven unprecedented and unexpected use of the accounts of trials at the Old Bailey criminal court, source material hitherto appreciated primarily by a small group of social historians working on early-modern crime and punishment in the London and its environs. In areas where data is harder to capture or is less voluminous, historians have undertaken their own data generative work. Here the *Dirty Books* project stands out—research that used a densitometer to study traces of human interaction with the bottom right-hand corners of medieval prayer books and by doing so approach an understanding of the use of those prayer books.[4] People, things, and experiences are also at the heart of the *Virtual Paul's Cross Project*.[5] Here modeling of sound and space re-creates a lost past experience—the experience of hearing an early modern sermon at St. Paul's Cross, an outdoor space beside medieval St. Paul's Cathedral that was lost during the Great Fire of London in 1666. The model has empowered historians to infer fresh insights about how sermons would have been delivered in the unamplified and noisy environment: the imposing aural impact on the model of the bell at St. Paul's that tolled at fifteen-minute intervals suggests that preachers such as John Donne timed their sermons around the bell, perhaps reaching climatic moments just as the bell was set to chime. Historians of the contemporary world, by contrast, have no shortage of data, and those historians whose research has addressed periods

after 1996, after the public deployment of the World Wide Web, are confronted with vast amounts of web data that are almost too large, too complex, and too unstructured to handle. And yet historians have persevered. Ian Milligan has demonstrated how blending traditional elements of the historian's toolkit—sampling, source analysis, close reading—with computational clustering and networking of data can bring the World Wide Web within the purview of historical research.[6] This work is imperative to the future of historical research (discussed later).

Complementing all this digital history has been no lack of theory. Bob Nicholson has called for wider acceptance of methods that blend close and distance reading. "Faced with this mountain of print," Nicholson writes, "we have two choices: to continue subjecting tiny fragments of Victorian culture to close reading, or to supplement this approach by exploring a much larger proportion of the archive through 'distant reading.'"[7] Of course, millions of digitized pages scratch only the surface of our physical archives, so historians have been at the forefront of stressing the cultural and political biases of mass digitization[8] and the need to construct rigorous models for sampling digital collections that shift bias away from the digitization process and back to the bias in the chosen category of source material.[9] For all the utopian rhetoric around the democratization of historical research in a digital age, research today remains as littered with barriers as in the predigital age, with novel hierarchies often causing research to be bounded by what is permissible rather than by what is possible.[10] And even where permissions are attained, digital historians have been keen to stress the limitations of what is possible with digital platforms, texts, and tools. Digital scholars have emphasized the need to constantly press colleagues and students to consider what is inside the black boxes of interfaces, data, and software.[11]

These critiques are not, however, the same as warning historians away from the use of digital data derived from past traces. For, as historians trained in source analysis, digital historians know the strengths and weaknesses of their sources. In the case of OCR-derived text whose "accuracy" is questionable, this data is not a poor facsimile of traces of the past, but—like a photograph, illustration, or oral history of a past event[12]—is instead a new category of source with its own affordances, limitations, and relationships to those past traces. Seen is this way, digitization is not routine and mechanized, but creative and performative, a transformation of a physical

thing into a new research object, into derived data, into a data form that can enrich, connect, and reconfigure the original data point, the physical thing itself, the stuff of history all historians seek to explore.[13]

This digital history is gathering critical histories.[14] One recent telling of that history argues for digital history to take better account of the history of computing.[15] Such histories are a sign of maturity, and as practitioners broaden their gaze they see that an urge to historicize their practice chimes with a wave of histories of the digital humanities, computing, and libraries. Notable work ranges from Trevor Muñoz's plea in the present volume for librarians to shape the future of the digital humanities through a grounded reinvestigation of the history of librarianship, to Rens Bod's *A New History of the Humanities*, a text that emphasizes with astonishing breadth a deep history of pattern matching in the humanistic method.[16] Elsewhere, Stéfan Sinclair and Geoffrey Rockwell have emphasized the human contingency and materiality of early work in humanities computing as a provocation for reflecting on the human contingency and materiality of current digital humanities project.[17] In a similar vein, both Melissa Terras and Julianne Nyhan, Andrew Flinn, and Anne Welsh have called for greater understanding of the prehistories and histories of the DH movement.[18] Indeed as Willard McCarty has argued, the digital humanities needs "to begin remembering what our predecessors did and did not do, and the conditions under which they worked, so as to fashion stories for our future."[19] And he has a point, because evidence of forgetting to remember and its consequences abound. For example, in June 2014 the newly formed Cambridge Centre for Digital Knowledge (CCDK) published a mission statement whose ahistorical phasing of digital humanities work, a phasing detached from the rich, diverse roots of DH, provoked the not unreasonable ire of McCarty.[20] Bethany Nowviskie would no doubt see CCDK's statement as evidence that there is little end in sight for the eternal September of the Digital Humanities, especially as the field spreads, institutionalizes, and atomizes.[21]

Taken together this body of reflective work constitutes a growing recognition that histories are vital tools for grappling with the future of digital research in the humanities. The remainder of the present chapter takes this history building a step further, concurring with Bod that histories of the humanities from *the vantage point of digital research* are crucial for future cross-fertilization between the two. I take as my example the discipline of history, a discipline whose source material—as I have described—is now available through

network technology and digital libraries at a previously incomparable scale. At the same time this discipline has failed to reap the full rewards of digital transformations in society and culture. For this situation to change, I suggest that librarians armed with knowledge of how and why this failure has manifested itself, of the historiographical traditions that shape and constrain the ability of historians to undertake and assimilate data-driven approaches to the past, are valuable collaborators in digital history projects, research, and pedagogy.

Of course, it is neither wise nor possible to approach as a whole a discipline as wide ranging in geographical focus, exhaustive in chronological scope, and varied in method as the discipline of history. Instead, this chapter restricts itself to exploring the discipline through those introductory texts many historians will be familiar with from the undergraduate classroom. For doing so through the lens of digital history reveals patterns worthy of close attention by all invested in the present and future of both digital history and digital research in the arts and humanities, not least librarians, in whose domain the stewardship and description of digital resources largely fall.

◆ ◆ ◆

John Tosh's *The Pursuit of History* is a classic introductory text in the discipline of history. First published in 1984, it has been substantially revised since and is now in its fifth edition. Together with these periodic revisions, Tosh's clarity, concision, and measured evaluation of scholarly trends have contributed to his volume becoming a favorite in the classroom. The history of these multiple editions offers a valuable perspective on the discipline they serve. For even if an analysis of their differences cannot hope to track changes over time in the research trends of all historians, the editions do represent a significant discursive contribution to the evolving process of self-definition and self-identification within the profession.

Of course "digital history" was unknown when Tosh originally wrote *The Pursuit of History*. "History and computing" on the other hand was an established, if minor, subfield and both Tosh's first and second editions reflect this in the index. Published in 1984 and 1991 respectively, these editions include three entries for "computers," all of which correspond to a chapter on quantitative methods entitled "History by Numbers." Here Tosh argues that the growth of computing in the discipline of history prior to the 1980s can be attributed to two factors: a desire to study more than histories of great men that turned historians to different sources, many of which needed counting;

and the relative affordability from the 1960s onward of computers, which experienced cost reductions that may have kept computers out of reach of individuals but not of research-focused history departments, many of which were able to afford computers, justify their purchase, and acquire prestige from investment in them. This interplay between computing and historical research meant that "both the kind of data it [the computer] could handle and the operations it could carry out were rapidly diversified."[22] Though unattributed, Tosh may well have been referring here to early concordance work with historical texts, the history and significance of which is currently enjoying a renaissance.[23] Nevertheless, the prevailing context for computation in both editions of *The Pursuit of History* is numerical work and statistical analysis, with the computer being a labor-saving, operational, and research management device yoked to numbers. Thus, Tosh sees fit to both emphasize the importance of statistical work to the profession—whether enabled by computational resources or not—and to add a considered note of caution. "Statistics," he writes, "may serve to reveal or clarify a particular tendency; but how we interpret that tendency—the significance we attach to it and the causes we adduce for it—is a matter for seasoned historical judgement, in which the historian trained exclusively in quantitative methods would be woefully deficient."[24] Familiar as it should sound, the argument is worth stressing: past phenomena are not revealed by numbers or by computation, but by the historian's interpretation of those numbers and that computation.

By the fifth edition of *The Pursuit of History* (published in 2010, over a decade after the fourth), the historical profession had changed profoundly. Comparative, postcolonial, and global history had emerged out of the ashes of conflict between macroanalytical social historians and microanalytical cultural historians and the rebuttal of postmodernist critique became a keen focus for work on the historical method.[25] In response to these changes, a range of novel approaches to historical phenomena featured prominently in the fifth edition of *The Pursuit of History*. Whole chapters discussed historian's qualitative research into gender, race, and colonialism. By contrast, a mere two and a half pages were reserved for discussion of quantitative history, statistics, computation, and the implications of macroanalytical work.

Seen from the vantage point of digital history, this is a striking and troubling transformation. For just as historians began to harness the infinite archive, just as digital history was gaining momentum, just as the

digital object libraries that had spent over a decade creating and collecting were beginning to be more widely used by humanities researchers as more than finding aids, just as interfaces—scholarly or otherwise—revealed the unimaginable breadth and volume of sources at the historian's disposal, and just a year before the Alliance of Digital Humanities Organisations invited all on the fringes into its "Big Tent," a key textbook in the discipline of history relegated quantitative history and the skills associated with it— both mathematical and conceptual—to marginal status.[26] In doing so and at a time when computational devices of various forms had become ubiquitous tools in the creation of the historian's work, *The Pursuit of History* removed from its index all references to "computers."

Whether he saw their causes as intellectual or social, there were good reasons for Tosh to shift the emphasis of *The Pursuit of History* in the direction he did. Though the 1960s and 1970s had been a fertile, confident, and critical period for quantitative work in history,[27] big picture, quantitative history began to decline in the 1980s when microhistorical, qualitative history began its ascendancy. In an Anglophonic context at least, the "fear of the mathematical" that Willard McCarthy characterizes as a defining feature of late-twentieth-century humanist scholarship was reflected in historians distancing themselves—and by extension their students—from numerical work.[28] That fear coalesced with a fear of scale, of appearing insufficiently close to the archive, of accusations of abstraction, and of lacking specialism and focus.[29]

It is curious that Tosh fails to note the implications for the historical profession of these shifts away from research with numbers and at scale. For extending his logic that "the historian trained exclusively in quantitative methods would be woefully deficient," an historian trained exclusively in qualitative methods, with no grounding in numbers, in computation, would be also "woefully deficient." And this scenario is not hypothetical. It is now a reality born out of the apotheosis of the very approaches given prominence in the fifth edition of *The Pursuit of History*. Given the technology and data historians now have at their disposal, the sort of measured discussions in Tosh's first edition around how to do history at scale and by numbers and around how that work fits into the task of historians at large should be a standard part of the historian's craft, of their training, of their conceptual universe.[30] In the fifth edition of *The Pursuit of History* and in the picture of the profession it paints, they are neither present nor required.[31]

An absence of respect for computational analysis can be observed in other comparable texts. In her robust counter to both naive empiricism and postmodernism, Mary Fulbrook's *Historical Theory* lingers on the intersection between traces of the past and historical narratives but not on the varied character of those traces or the skills needed to handle them (except to say that all traces are valid depending on the question at hand).[32] Another popular textbook, *History: An Introduction to Theory, Method and Practice* by John Marriott and Peter Claus, aims to bridge the gap in praxis and epistemology between studying history at school and in higher education.[33] It demystifies concepts and surveys the field circa 2010, but quantitative approaches and methods do not feature. In *Historiography in the Twentieth Century: From Scientific Objectivity to the Postmodern Challenge*, George Iggers traces the discipline of history's gradual abandonment of macrohistory, grand narratives, and its postwar roots in sociological theory. First published in 1997, his epilogue for the 2005 reprint stresses the need for global history to build on the gains it made in the late 1990s and for a program of synthesis. But Iggers doubts that need will translate into reality—for doing global history requires teams of authors to grapple with problems of global scale and for those authors to willingly "operate on a speculative plane of global history alien to historians who avoid empirical work." The implication is that historians who avoid empirical work are in the majority.[34]

In sum, these textbooks—and many more like them—fail to address the loss of quantitative methods from the historian's toolkit and the implications of this for the profession at large. Only Iggers—in language reminiscent of Braudel—notes the potential adverse consequences of that loss with respect to the strength of the global history project. But even he seems curiously nonplussed—*Historiography in the Twentieth Century* contains no call for action and is far from a manifesto for change.[35]

History in Practice by Ludmilla Jordanova is perhaps singular in the genre for arguing at length in favor of rehabilitating quantitative analysis as a core component of the historian's craft. Published in 2000, her first volume argued that the development of undergraduate curricula by the profession should weigh a fashion for certain approaches—for example, microhistorical, cultural approaches—against an overall sense of the skills historians should have. "Economic history," she wrote, "is particularly vulnerable in this respect."[36] Continuing, she said:

> Economic history (like some other fields) is a fundamental part
> of the discipline, of which every student *ought* to have some
> understanding [. . .] Faced with the choice between courses
> on the history of sport or the history of animals and those on
> economic, political, social or intellectual history, I would hope
> students would be able to see that the latter are likely to be of
> more general use than the former.[37]

Central to the historian's craft here is the understanding of how to negotiate the relationship between big and small history, between macro and micro, between "scientific" and humanistic methods.

In the second edition of *History in Practice*, published in 2006, Jordanova extended this discussion of core skills and tools further, to address how historians could and should respond to novelty in the digital age in light of the professional attributes they wish to preserve. A section entitled "Webs of Affinity" begins by setting the scene: websites offer access to "unimaginable" volumes of historical information; the links between them and the portals that allow researchers to discover them are increasing in sophistication; and many hitherto difficult-to-obtain sources are now at the fingertips of the historian. These factors by themselves, Jordanova argues, "hardly possess the capacity to change radically the ways in which professional historians work."[38] What does possess that power is the manipulation of those websites and the data they contain, and the imagination to see that "unforeseen patterns may emerge which could not have been detected without information technology."[39] Such power requires judicious use and the ability for researchers to utilize these technologies. She concludes that scholars will have to reflect with care on their practice, on how working with data may encourage "fantasies of being able to do truly exhaustive research" or of how our present concerns and uses of technology—say, social networks—may cause an unintended vogue for certain approaches—say, network analysis—in the methods historian use to underpin their explorations of historical phenomenon.[40] Once historians have negotiated the potential and pitfalls of digital technologies, Jordanova continues, they will need both new skills and old skills reapplied. And yet the ability of historians to deliver this is at risk in the siloed and fractured professional landscape that emerged from the cultural turn for,

as Jordanova notes: "It is to be regretted that, like economic history and demographic, history and computing is often seen as a specialist domain dominated by enthusiasts."[41]

History in Practice stands out among history textbooks as the sole voice that identified and lamented a decline of quantitative skills, latterly computationally enriched, in both the training offered to historians and the historian's craft, a decline this chapter has observed in the publication history of Tosh's *The Pursuit of History*. In the context of the present volume, it seems to me that we—the library community—must both share and expand upon Jordanova's lament. For to do aspects of digital history well, to take full advantage of those sources—be they ledgers, ephemera, books, newspapers, sound recordings, videos, web pages, or personal digital media—that libraries make available to historians as data, as source material that can be manipulated, counted, and prodded by machines working at their behest, that can be queried at scale rather than merely presented in digital forms yoked to print paradigms, the historical profession needs quantitative skills and a critical understanding of the profession's deep and contested relationship with quantitative research. Librarians can be key collaborators who ensure that historians and other humanistic scholars have the ability to do rigorous quantitative research, but, in order for these partnerships to work, it is clear from the before-mentioned textbooks that there is much work to be done.[42] Emerging historians in particular need to know how to count as historians and how to be critical of the role of data and computation in that counting, for should they go on to attempt digital research of a data-driven variety, the quality of their work may depend on their possession—or otherwise—of these once core skills.

If the future of the historical profession itself is not at stake here, then its health as judged by its ability to explore historical phenomena using the best tools and methods for the job certainly is. Dan Cohen and Roy Rosenzweig identified this nearly a decade ago when they called for historians to wake up to the loss taking place of *the* primary historical record of our time—the website.[43] The salience of their concern that historians were not taking the digital age seriously and were ill prepared for research using this category of source has only amplified since. Librarians need to ask urgently whether the historians they work with, many of whom were trained during the apotheosis of cultural microhistorical research, are equipped to

deal with categories of sources such as the archived web. Librarians need to be mindful of whether historians entering undergraduate study in 2016 and graduate programs in 2020 are likely to be capable of exploring the born-digital, data-rich post-1996 world. Librarians also need to understand whether, without major intervention, future historians will be equipped with the skills to tackle vast, technically complex, and enormously rich archives of websites, email, social media traffic, and personal digital media. Slowly we observe that the profession is waking up to this imperative, to the reality of its present, and to how the debates of the past can be of service to its future.[44] In the United Kingdom, nonprint legal deposit powers granted to the British Library have empowered the UK Web Archive to move from selective capture of web publications to annual domain crawls of all ".uk" websites and associated publications.[45] The Institute of Historical Research has taken a leading role in exposing the historical community to this source material, to its affordances, its limitations, its demands of researchers, and its vital role in future historical research. Nonetheless more work remains to be done. For as stewards of digital resources know, a tidal wave of data is not coming—it is here.[46]

Of course it is quite possible the wave might pass by the historical community altogether. Most professional historians living today will never use web archives or personal digital media as research objects. More, but likely far less than a majority, will during their career use digital collections outside of print paradigms and use software tools and algorithms to manipulate data at scale.[47] For these reasons Braudel's struggles may well continue to resonate—many historians may indeed continue to find unpalatable the uncertain methods of a quantitative, at scale, or knowingly imperfect variety. But we should all be concerned if a detachment from data-driven methods crystallizes into uncritical oppositional dogma, not least the many librarians who grapple daily with how to ingest, catalog, describe, and explore such data and how to scale those processes in anticipation of a coming uplift and change in researcher demand. These same librarians are conversant in the challenges of size, technical complexity, and legalities associated with doing research with this data. They have both the will and the skills to effect change, and by complementing these attributes with a perspective on the historical profession as seen through the lens of digital history, of the historiographical traditions that shape and constrain the ability of historians at

large to undertake and assimilate data-driven approaches to the past, these librarians can be valuable collaborators in digital history projects, research, and pedagogy. They can use their contextual knowledge to make the uncertain certain, the unpalatable palatable, and they can work with historians to overcome the profession's timidity toward mathematics, scale, and distance from the archive. Together with historians, these librarians can begin in earnest to exploit in novel and unexpected ways the digital collections that libraries, archives, and museums have spent over two decades managing, securing, and promoting.

<p style="text-align:center">♦ ♦ ♦</p>

As libraries explore the complex forces that shape and constrain the use by historians of their digital collections as data, context—as with all things history touches—will remain king. For seen in the context this chapter discusses, digital transformations in society and culture offer the historical profession as many continuities as discontinuities—in short, the profession has had these discussions, or at least a version of these discussions, before and outcomes of a tone and character satisfactory to the profession at large were reached. Among these were the reflections advanced by the French Annales School. In 1973 Emmanuel Le Roy Ladurie wrote:

> In history, as elsewhere, what counts is not the machine, but the problem. The machine is only interesting insofar as it allows us to tackle new questions that are original because of their methods, content and especially scale.[48]

His "machine" was the computer, the role of which in historical research was—as his fellow Annalist Braudel had observed less than a decade earlier—under scrutiny. But as Ladurie knew full well, that machine could equally be a map, a calculator, a square ruled notebook, a library catalog, a filing cabinet, or indeed any tool historians have profitably used to undertake their craft and to deepen their understanding of past phenomena. As reflexive scholars steeped in these traditions, in a rich and critical continuum of historical research and method, digital historians know that better history results from methods that see not the novelty of a tool, but the new questions that can be asked of sources with the tool in their hands. When that reflexivity is mainstreamed, the digital resources libraries steward and

curate will be best exploited. To achieve that mainstreaming and for the current efflorescence of digital history to be sustained, an efflorescence library professionals are—as the present volume demonstrates—benefiting from and are collaborating in, the historiographical traditions that shape and constrain data-driven historical research should be emphasized, disseminated, and fostered. By taking into account not only the traditions and perspectives but also the histories and controversies of humanities disciplines, while laying the foundations for digital humanities work, library professionals can, I argue, play a crucial role in making this happen.

NOTES

1   I thank Thomas Padilla for his generous comments on an early draft of this chapter.

2   Fernand Braudel, C*apitalism and Material Life, 1400-1800*, trans. Miriam Kochan (London: Weidenfeld and Nicholson, 1973), 6–7.

3   Tim Hitchcock et al., *The Old Bailey Proceedings Online, 1674–1913* 7.0 (March 2012), www.oldbaileyonline.org; Tim Hitchcock et al., *London Lives, 1690–1800* 1.1 (April 24, 2012), www.londonlives.org; Tim Hitchcock et al., *Connected Histories* (2013), www.connectedhistories.org; and Tim Hitchcock et al., *Locating London's Past* 1.0 (December 17, 2011), www.locatinglondon.org.

4   Kathryn Rudy, "Dirty Books: Quantifying Patterns of Use in Medieval Manuscripts Using a Densitometer," *Journal of Historians of Netherlandish Art* 2, No. 1–2 (2010).

5   John Wall et al., *Virtual Paul's Cross Project* (2013), https://vpcp.chass.ncsu.edu.

6   Ian Milligan, "Clustering Search to Navigate a Case Study of the Canadian World Wide Web as a Historical Resource," in *Digital Humanities 2014* (2014), http://dharchive.org/paper/DH2014/Paper-83.xml.

7   Bob Nicholson, "Counting Culture; or, How to Read Victorian Newspapers from a Distance," *Journal of Victorian Culture* 17, No. 2 (2012). By using the phrase "distant reading," Nicholson invokes and reimagines a phrase coined by Franco Moretti to argue for literary analysis using digital tools. See Franco Moretti, *Distant Reading* (London: Verso, 2013).

8   David Armitage and Jo Guldi, "The Return of the Longue Durée: An Anglo-American Perspective," *Annales: Histoire, Sciences Sociales* 69 (2014): 44.

9   Pieter Francois and Ben O'Steen, *The Sample Generator* (2013), http://sample generator.cloudapp.net.

10   Andrew Prescott, "Dennis the Paywall Menace Stalks the Archives," *Digital Riffs* (February 2, 2014), http://digitalriffs.blogspot.co.uk/2014/02/dennis-paywall-menace-stalks-archives.html.

11   Melodee Beals, "Record How You Search, Not Just What You Find: Thoughtfully Constructed Search Terms Greatly Enhance the Reliability of Digital Research," *LSE Impact of Social Sciences* (June 10, 2013); Andrew Hobbs, "The Deleterious Dominance of the Times in Nineteenth-Century Scholarship," *Journal of Victorian Culture* 18, No. 4 (2013); Daniel J. Cohen and Roy Rosenzweig, *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web* (Philadelphia: University of Pennsylvania Press, 2006). See also the decision of *Connected Histories* and *Welsh Newspapers Online* (2013), welshnewspapers.llgc.org.uk, to expose the OCR text within their interfaces.

12   See Peter Burke, *Eyewitnessing: The Uses of Images as Historical Evidence* (Ithaca, NY: Cornell University Press, 2001).

13   Steven E. Jones, *The Emergence of the Digital Humanities* (New York: Routledge, 2014).

14   The most significant of these remains Cohen and Rosenzweig, *Digital History*.

15   William J. Turkel, Shezan Muhammedi, and Mary Beth Start, "Grounding Digital History in the History of Computing," *IEEE Annals of the History of Computing* (2014): 72.

16   Rens Bod, *A New History of the Humanities: The Search for Principles and Patterns from Antiquity to the Present* (Oxford, UK: Oxford University Press, 2013).

17   Stéfan Sinclair and Geoffrey Rockwell, "Towards an Archaeology of Text Analysis Tools," *Digital Humanities 2014* (2014), http://dharchive.org/paper/DH2014/Paper-778.xml.

18   Melissa Terras, "A Decade in Digital Humanities," *Melissa Terras' Blog*, May 27, 2014, www.melissaterras.blogspot.co.uk/2014/05/inaugural-lecture-decade-in-digital.html; Julianne Nyhan, Andrew Flinn, and Anne Welsh, "Oral History and the Hidden Histories Project: Towards Histories of Computing in the Humanities," *Literary and Linguistic Computing* 30, No. 1 (2013): 71–85. http://dx.doi.org/10.1093/llc/fqt044.

19   Willard McCarty, "Getting There from Here. Remembering the Future of Digital Humanities (Roberto Busa Award Lecture 2013)," *Literary and Linguistic Computing* 29, No. 3 (2014): 283–306. http://dx.doi.org/10.1093/llc/fqu022.

20 Willard McCarty, "Digital Knowing, Not Digital Knowledge," *Humanist* 28, No. 140 (June 23, 2014), lists.digitalhumanities.org/pipermail/humanist/2014 -June/012095.html.

21 Bethany Nowviskie, "Eternal September of the Digital Humanities," *Bethany Nowviskie* (blog), October 15, 2010, www.nowviskie.org/2010/eternal -september-of-the-digital-humanities.

22 John Tosh, *The Pursuit of History: Aims, Methods and New Directions in the Study of Modern History*, 2nd ed. (London; New York: Longman, 1991): 153.

23 Sinclair and Rockwell, "Towards an Archaeology of Text Analysis Tools."

24 John Tosh, *The Pursuit of History: Aims, Methods and New Directions in the Study of Modern History*, 1st ed. (London; New York: Longman, 1984): 197.

25 For the emergence of microhistory, see David Armitage and Jo Guldi, "The Return of the Longue Durée" and *The History Manifesto* (Cambridge, UK: Cambridge University Press, 2014). For the preoccupation with rebuttals of postmodern critique, see Richard Evans, *In Defence of History* (London: Granta, 1997) and Mary Fulbrook, *Historical Theory* (London; New York: Routledge, 2002).

26 "Big Tent Digital Humanities" was the theme of *Digital Humanities 2011*, held in Stanford, California.

27 Jacob M. Price, "Recent Quantitative Work in History: A Survey of the Main Trends," *History and Theory* 9 (1969): 11, 13.

28 McCarty, "Getting There from Here," 7. Criticism of cliometrics (the application of econometric models to past phenomena) was another influential force that contributed to the distancing of historical practice from numerical work. Lawrence Stone's emphasis on the narrative and particular (as opposed to analytical and statistical) nature of historical research in "The Revival of Narrative: Reflections on a New Old History," *Past & Present* 85 (1979): 3–24, is a classic essay in the genre.

29 As David Armitage and Jo Guldi argue, these fears were not driven by intellectual rationale but by the job market; Armitage and Guldi, "The Return of the Longue Durée," 10–15. Tosh is silent on this subject, though of course a textbook aimed at undergraduates is perhaps not the venue for ruminations on administration in higher education.

30 A comparable argument is made in Armitage and Guldi, "The Return of the Longue Durée," 21, 39, and *The History Manifesto*.

31 In February 2015, as this manuscript was in the final stages of preparation for publication, a sixth edition of *The Pursuit of History* was published. The

edition includes small sections on the place of digital sources in historical research. These overview passages contain reflections that range from astute criticisms of the digital archives historians now have at their disposal, such as the poor coverage of the digitized record in relation to both the size and thematic coverage of the total historical record, to unfortunate lapses of reason: Tosh's comment that "nothing on the Internet is wholly original, though it is often treated as though it was" is not only nonsense, but also takes no account of the wholly born digital materials that citizens, organizations, and governments create every day and that historians will rely on to re-create our present and our recent past. Crucially, no mention is made of computational approaches to historical research using these sources or of combining the digital with the quantitative methodologies that were dropped by the fifth edition. So, although the coverage of digital sources in the sixth edition is a welcome development, the coverage of digital history remains far from satisfactory.

32   Fulbrook, *Historical Theory*.

33   John Marriott and Peter Claus, *History: An Introduction to Theory, Method, and Practice* (London; New York: Longman, 2011).

34   George Iggers, *Historiography in the Twentieth Century: From Scientific Objectivity to the Postmodern Challenge* (Middletown, CT: Wesleyan University Press, 2005; first published 1997): 155. Price ("Recent Quantitative Work in History," 13) has similar qualms about the preparedness of historians for cooperative working.

35   Note that although Pat Hudson's *History by Numbers: An Introduction to Quantitative Approaches* (New York: Bloomsbury Academic, 2000) provides an excellent introduction to counting as a historian, it is not a generic history textbook and falls outside the remit of this study.

36   Ludmilla Jordanova, *History in Practice*, 1st ed. (London: Arnold, 2000), 202.

37   Ibid., 203.

38   Ludmilla Jordanova, *History in Practice*, 2nd ed. (London: Arnold, 2006): 187.

39   Ibid., 188.

40   For the application of network analysis to historical research, see Claire Lemercier, "Formal Network Methods in History: Why and How?," *HAL Sciences de l'Hamme et de la Société* 2 (August 12, 2011), halshs.archives-ouvertes.fr /halshs-00521527.

41   Jordanova, *History in Practice*, 2nd ed., 188.

42 A useful example of how these approaches can complement one another appears in Jordanova, *History in Practice*, 1st ed., 49–55.

43 Cohen and Rosenzweig, *Digital History*.

44 See, for example, T. M. Kelly, *Teaching History in the Digital Age* (Ann Arbor: University of Michigan Press, 2013); Turkel, Muhammedi, and Start, "Grounding Digital History," 74, and Armitage and Guldi, *The History Manifesto*.

45 See "Non-Print Legal Deposit: FAQs," *The British Library.uk* (2013), www .bl.uk/catalogues/search/non-print_legal_deposit.html.

46 For further examples of the research community responding to this challenge, see the Social Sciences and Humanities Research Council (Canada)–funded "Postwar English-Canadian Youth Cultures: A Digital History, 1945–1990" project led by Ian Milligan (ianmilligan.ca/the-next-project/project-proposal-sshrc, 2013) and the National Endowment of Humanities (United States of America)–funded "Archive What I See Now" project led by Michele C. Weigle (secure grants.neh.gov/publicquery/main.aspx?f=1&gn=HD-51670-13, 2013–2014).

47 Historians do of course use search engines, largely uncritically, to manipulate data at scale every day; see Ted Underwood, "Theorizing Research Practices We Forgot to Theorize Twenty Years Ago," *Representations* 127, No. 1 (2014): 2, 9.

48 Emmanuel Le Roy Ladurie, "L'historien et l'ordinateur," *Le territoire de l'historien* (Paris: Gallimard, 1973): 11.