Introduction

The Materials at Hand

ALISON BOOTH AND MIRIAM POSNER

ALISON BOOTH is professor of English and academic director of the Scholars' Lab at the University of Virginia. She directs Collective Biographies of Women, a digital project based on her book, *How to Make It as a Woman: Collective Biographical History from Victoria to the Present* (U of Chicago P, 2004). Booth's other books include *Homes and Haunts: Touring Writers' Shrines and Countries* (Oxford UP, 2016) and *Greatness Engendered: George Eliot and Virginia Woolf* (Cornell UP, 1992).

MIRIAM POSNER is assistant professor of information studies and digital humanities at the University of California, Los Angeles. She has practiced and taught DH for more than a decade and is at work on a book about information in global supply chains. Her work has appeared in *American Quarterly*, *Debates in the Digital Humanities* (U of Minnesota P, 2016), *Moving Image*, *Feminist Media Histories*, *Logic*, and *The New Yorker*.

WHERE IS DIGITAL HUMANITIES IN 2020? AS SOME DH programs head into their third decade, it seems inappropriate to call this a burgeoning or nascent field, still less a hot new thing. And yet these clichés feel inescapable as one attempts to describe a field that has yet to coalesce around a set of methods or assumptions. Indeed, new technologies and aspirations—not to mention caricatures—sprout up so often that it is difficult to speak of unified trends, or to get proponents or critics to be sure they are talking about the same things. This introduction is not the occasion to reproduce existing histories of DH or to adjudicate controversies that occupy leading journals such as *Critical Inquiry* and *New Literary History*. We can agree that DH in 2020 has come far. No one would maintain that DH in 2030 will be what it is today. The contributions to this issue attest to the rich varieties of DH at this time, and some of them address the shape and purpose of this capacious field. Before we introduce the essays that follow, let us offer some views of DH as it is, and a snapshot of some institutional changes at our current moment, based on our experiences leading DH groups and shuttling between libraries and departments of English or information studies. This group portrait might be captioned, some day in the future, "DH, before [midcentury metamorphosis to be named]."

For at least as long as there has been an MLA, there have been arrivals of schools (New Criticism was once threatening) and tools (microfilm was once state-of-the-art), and hence new purviews of research for its members. The ultimately successful arrivals were initially heralded about as warmly as invasions of barbarians. We might say that the invasion of the digital is, in some sense, over. Textual and other humanities scholars all use computers and online search tools, many share discussions with their students in forums

© 2020 ALISON BOOTH AND MIRIAM POSNER

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

or portfolios, most of them access documents or artifacts online, and many at some point in their research apply data processing and visualizations formerly more common in social and natural sciences. Yet all this computer-related activity is not the same thing as the research field we exhibit in this special issue.

So how do we define DH? One short phrase cannot grasp it. *Advanced humanities research that uses and reflects on computational methods or digital tools* is a serviceable definition. But it needs added clauses. Beyond teaching with technology or sharing an essay on a Web site, DH prizes *interdisciplinary collaboration and technical experimentation*. This mode of scholarship is most worthwhile if it also *promotes public engagement and humanistic knowledge and understanding*. It can be curatorial, but it must be suspicious of archives taken at face value. It can offer skills that lead to alternative or academic jobs, but it must also recognize the value of beautiful speculation, remembering that all art is quite useless, as Oscar Wilde proposed. Moreover, we reassert a fact that we believe to be self-evident: that DH must concern itself deeply with race, gender, disability, economic and linguistic access, and other intersecting axes of power embedded in our materials and methods, as demanded by this troubled moment in the world.

Though DH addresses long-standing questions in humanistic disciplines, it does so in new ways that transform the products of research and demand a different kind of training and assessment. Much of the most advanced digital literary scholarship is so technical that it takes years of study to understand the methods underlying it. This of course makes it hard for DH scholars to communicate with peers who have not had such training. Even if a computational literary scholar is exceptionally talented at explaining his or her methods to laypeople, the would-be DH critic still must base any criticism on the initiate's explanation rather than on testing or reproducing the results. Students in the twenty-first century, supposedly digital natives, will nevertheless find that they need to attend workshops or take tutorials to understand the inside of their computers, the structure of the Internet, or the command line for coding. And it takes more than that to induct anyone, at any stage of education and career, into this discourse, where many of the nouns and verbs are as specialized as they are in the discourse of chemistry.

Social and institutional conditions, as well, differ in DH from what is familiar to a practiced individual researcher in a literature or language department. Midcareer humanities faculty members embarking on a digital project may find it akin to learning a new language or discipline. In addition to learning the skills involved, they may for the first time apply for multiyear, interinstitutional grants that their university skims for overhead, and find themselves collaborating with student researchers or non-tenure-track staff, as scientists do. Managing such a project does not resemble teaching a course or chairing a committee. The results arrive slowly, and then often are harder to fit in the categories of a humanist's CV. While DH may yield a peer-reviewed article, essay collection or journal, or single-author monograph, these are often based on datasets, programming, documentation of method and results, visualizations, and user interfaces, all of which are hard to encompass in one reading for review, not to mention that they remain living artifacts created by many hands, seldom finished or preserved on one publication date.

These observable differences between DH and customary humanities practices have been sources of friction, but this is not a reason for humanists to delegate to STEM fields the effort to foster *interdisciplinary collaboration and technical experimentation that also promotes public engagement*. Collaborative research programs dealing with systematic data are not, in fact, alien to the humanities—think of museums, libraries, the *Oxford English Dictionary*, and classical philology

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

and prosopography, all well into their second century, if not much older. DH projects may resemble quantitative social sciences in some methods and data analysis, but with few exceptions DH researchers avow the element of interpretation and indeterminacy in the materials of culture instead of emulating the fields that seem most pleasing in the eyes of the university's lords. Rather, an admonition heard often in DH circles is that there is no neutral ground just as there is no split between coding and scholarship; human interpretation has a hand in every aspect of planning, coding, implementing, and sharing a project, all the way down (Gallon; Risam; Noble; Johnson).

The interdisciplinary encounters at our own institutions should be introduced in the context of the tendencies in the past decade for researchers to be attracted by different disciplinary models. In 2010 Matthew G. Kirschenbaum first published his assessment that for two decades DH had been "more rooted in English departments than any other home" (3; see also Posner). In 2011 the annual DH conference took as its theme the big tent, seeking to unite digital literary scholars, multimedia historians, and corpus linguists under one banner. In this era DH meetings advertised "more hack, less yack" or expressed a brand of reckless optimism as if echoing Silicon Valley. Why not write a book in a week, or build a tool in a week (Croxall et al.; Cohen and Scheinfeldt)? The exciting potential remains palpable at DH conferences, but in the light of the current role of social media in electoral politics and growing critiques of human resources conditions in tech companies, few in DH still aspire to the ethos of the tech industry.

More-recent editions of *Debates in the Digital Humanities* represent a decade in which roots have spread to such disciplines as archaeology and the histories of art and architecture, and to many languages and cultures (Gold and Klein, *Debates* [2016] and *Debates* [2019]). Leading voices increasingly guide DH to distinguish itself from tech giants and to resist the rhetoric of objectivity in standard visualizations unsuited to the fuzzy data or aesthetic materials of the humanities and arts (Drucker). Decades of innovative experimentation have shifted the outlook of humanities research to the extent that digital methods will be integral to any further contributions along many lines of inquiry. How can literary history continue to make broad claims based on the small fraction of canonical publications rather than on inclusive, searchable corpora? Without digital intervention, how can more than a few lucky archivists access Spanish-language newspapers printed in North America, or reconstruct the social practices and the layout of buildings at the slaveholding University of Virginia (Villarroel et al.; Johnston et al.)?

This decade of proliferating DH has also seen fragmentation as areas of specialization diverge. For example, some computational literary scholars have recently begun identifying themselves as "computational humanists," while others have begun referring to their field as "cultural analytics." These labels are presumably intended to distinguish some fields from the larger, more eclectic field of DH, in which an online exhibit, map, video, or XML-tagged edition of a writer's corpus counts as scholarship. Perhaps, also, *cultural analytics* signals businesslike or information-science processing, to distance its adherents from the modes of interpretation and advocacy of many DH projects that align with cultural studies.[1]

Computational humanists, like scholars engaged in cultural analytics, are at pains to emphasize their programming bona fides. Folgert Karsdorp tweeted that skilled programmers are anxious to find a forum that "doesn't exclude people with profound computational skills and knowledge" (@FolgertK). Karsdorp then founded just such a forum, the *Journal of Computational Humanities Research*. One might ask, given that "softer," less computationally intensive DH projects are often associated with feminized forms of labor,

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

whether this move to distinguish "serious" computation from DH is partly an attempt to corral the prestige associated with masculinized professions like engineering. Lauren F. Klein, in the article included here as well as her book, *Data Feminism*, written with Catherine D'Ignazio, has highlighted the perils of attributing objectivity to data, and of rendering feminized labor invisible (D'Ignazio and Klein). Miriam Posner's work with cultural datasets and Alison Booth's work on gender and biography in Collective Biographies of Women (CBW) are unmistakably DH and not computational humanities or cultural analytics. And yet we work, evaluate scholarship, and teach in areas defined by terms like *computational*, *data*, and *analytics*.

We can detect another shift in nomenclature: *data* is perhaps edging out *digital*. Indeed, at many institutions the huge interest in data threatens to reduce the humanities either to curricular add-on at engineering schools—classes on ethics or society—or to undergraduate luxury. Data science initiatives, in various institutional guises, are cropping up everywhere. At the University of California (UC), Berkeley, UC Davis, Columbia University, Harvard University, the University of Michigan, and many other institutions, existing and new computational work—including, sometimes, DH—is being reframed or incentivized as data science; degrees and institutes are expanding into endowed schools of data science, as they were at the University of Virginia (UVA) in 2019. There, Booth wished to explore computational textual analysis as an approach to short biographies printed in books, and so in 2018 and 2019 she worked with Rafael Alvarado to mentor three data science master's students on their capstone projects. They had chosen to work on CBW's Biographical Elements and Structure Schema (BESS) data (controlled values that identify textual elements such as events in each paragraph of about 230 chapter-length biographies of women) because the project had an explicitly feminist, humanities purpose. Their cultural distance from the material was wide, however, because they had grown up in India and had specialized in computer science since they were teenagers. References to the history of women in Britain and the United States lacked context for them. The BESS dataset formed by human editorial decisions was more qualitative and smaller than they were accustomed to, while Booth found it challenging to understand their use of statistical language models. Booth and the students found together that sentiment analysis, charting positive or negative connotation of words, is no proxy for measuring plot in biographies.[2] It was a transdisciplinary experience worth pursuing.

Instead of competing with DH, data science can help increase the institutional endorsement of DH, in terms of faculty appointments or curriculum. The DH certificate is in its first full year of operation at UVA in 2019, with twenty-three MA and PhD students in disciplines from art and architectural history to nursing, along with English, history, religious studies, philosophy, French, Spanish, and Slavic. The Scholars' Lab's Praxis Program, initiated by Bethany Nowviskie and now run by Brandon Walsh, has annually gathered six doctoral fellows to develop a joint project that is always catalyzed by their different areas of expertise, from history to musicology, from ancient to contemporary. Humanities students and faculty members wanting to learn R, a programming language for statistical and data analysis, may work with data specialists in the science library, while engineering students may venture to the Scholars' Lab's Makerspace for collaborative play with open-source Arduino kits or 3-D printing. Some members of the Scholars' Lab call what they do spatial tech rather than DH, because they work with geographic information systems, virtual reality, and cultural heritage informatics (capturing 3-D data from sites and artifacts). Yet in this interactive picture of digital practices, it has been politic for

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

the Scholars' Lab as a whole to affiliate with DH, the better to articulate its ethos within UVA and externally ("Charter"). Working with visiting scholars and frequently hosting groups who are planning their own digital-research center, the Scholars' Lab commonly refers to DH to save time without closing off different models, and the team's blogging often addresses the DH community (Visconti).

At the University of California, Los Angeles (UCLA), Posner has been mystified by the majors of the undergraduate students who enroll in the popular DH minor. Posner expected to see English and history majors, along with art historians, classicists, and other humanities specialists. Instead, she got students from cognitive science, statistics, sociology, communication, and business economics, to name the most common majors in the program. Why would these students from the social and "hard" sciences choose to minor in DH? The students say that, like most human beings, they never set out to become specialists in one narrow discipline; they came to college with big questions about things like human nature, technology, and why the world looks the way it does. They were therefore unnerved to find that college required them to switch their sights from understanding the world to understanding the world through the very specific lens of molecular biology or organic chemistry. DH provides these students with a bridge across disciplinary divides and with some of the tools they need to make sense of the changes technology is working on the world around them. We may find these students' interdisciplinary ambitions surprising, but to them digital scholarship makes perfect sense as a method of understanding their everyday lives.

At a time when the news is full of doom-saying stories about the decline of humanities departments, it is heartening to have the opportunity to introduce so many nonhumanities majors to some of the fundamentals of humanities research and scholarship. Students are often surprised to discover that, for example, Michel-Rolph Trouillot's observations about power and the historical method can be applied to the interpretation of data. Humanistic critiques of Cartesian space take on sudden relevance when students confront the assumptions of mapping software. Theories about whiteness's relation to neutrality lend a new critical frame to students' interpretations of data visualizations. DH will never replace the humanities—no one should have to fiddle with code if they do not want to—but, at least at UCLA, it has provided a window onto the attractions and utility of this kind of scholarship to class after class of undergraduates.

These university students, faculty members, and members of staff realize the rewards of taking a mental and an actual walk across campus to extend their skills and repertoire. We suggest that DH and data science can be mutually instructive and supportive, though their aims, insights, standards, and community practices diverge. DH can offer opportunities for genuine interdisciplinarity in data-driven projects, some of which mine large, state-owned datasets.[3] Some researchers in computer science and in data science want to participate in projects that have historical depth, aesthetic appeal, or impact on social justice. Instead of indulging in heedless specialization, new university data initiatives can, in our experience, offer liberal arts engagement for STEM students. Booth has been directing a three-year lab at UVA with Debjani Ganguly and Jack Chen, called Humanities Informatics (concluding in 2020). The lab brings together faculty members and students working in media studies, urban and environmental planning, philosophy, computer science, and other fields to study both the revolution in information since World War II and the implications of emerging technologies such as facial-recognition software and drones. The term *informatics* had the potential to sideline earlier fortes of DH, such as textual and historical studies. And yet Booth's

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

experience with the lab suggests that the next phases of DH need to be conversant with the fields that are under pressure to produce so-called translational research, and in turn STEM fields might be enhanced by the insights of DH. Humanists are trained to expose institutions and structures as human-made and mutable. Members of the Humanities Informatics lab learned from one another's varied enthusiasms for new technological applications as much as from our informed fears about intended and unforeseen consequences of media and technology of many kinds. DH teams may be more likely than data or computer scientists to engage with historically and geographically diverse subjects, from ancient archeology to Quechua. Training in English literature and foreign language programs and related humanities disciplines predisposes us to ask why, at what cost, and for whom technology transforms noise into signal, into communicable meaning, into responsible action.

Humanists should authorize themselves to join evolving interdisciplinary conversations about data, computation, and machine learning. Today, as corporations, governments, and researchers amass data on an unprecedented scale, these conversations have become urgent. Neural networks, for example, now work in ways beyond the understanding of those who unleash them, and we are learning in real time about the implications of machine learning deployed on the scale of the global population. In response, many digital humanists have trained themselves to be attentive to the methods of data science, computer programming, and conventional data visualization, while at the same time pressing at their boundaries and limitations. Humanists, after all, are experts at describing particularity, variance in the data, and the outliers where the system fails—where, indeed, the informative tends to lurk. All models are wrong, as statisticians commonly say, because they belie what does not fit (So; McCarty). The arts, we could say, make something in honor of what exceeds a model and try to hear the loss—for example, what is lost in the compression of an audio file. Humanists are good at collecting and commemorating what changes and eludes comprehension. They have learned how to recollect and rephrase the previous attempts to answer the question of what it means to be human, and to read the history of technology in its social and political contexts. In the past half century, they have become attuned to the modes of thinking and being that Enlightenment paradigms and empire purposely lost, so they might be good consultants on what homogeneous data science will inevitably discount. Thus, there is reason to hope for an interdisciplinary encounter where DH meets data science.

A Hand on a New World Geography

In designing and creating the cover image, we allude to the ambitions of Google Books to put the world's printed heritage at our fingertips—a project that evokes digital approaches to text, though it hardly models the practice of DH as a field. In the decades since the pronouncement of this neo-Gutenberg revolution, we have realized some of its potential benefits for humanities research, but we have also seen a groundswell of skepticism about violation of intellectual property rights, corporate monetization of data, and distortion if not destruction of books—the materials at the core of humanities disciplines. And we now reap the consequences of ceding digitization to Google, which appears to have abandoned key parts of the infrastructure of Google Books (Schmidt).

As is so often the case in the era of globalization, much of the work for Google Books was outsourced to the Global South. Millions of volumes in Euro-American research libraries were sent to developing countries to be scanned by laborers with astonishing speed and accuracy. The inevitable errors, an immeasurable fraction of all the scanned pages,

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

seem to peel back the glossy outer layer of Google Books to reveal an enterprise far from omnipotent. As with the outliers in a statistical model or the failures of a theory or software design, so too with these glitches, often coinciding with such disruptions of verbal text as a page header or an illustration. The anomaly is where the researcher should start looking.

Various artists, scholars, and journalists have collected images of botched pages that look as if they are melting, or like tornadoes or exploding buildings, or for all the world like bad photocopies of photographs of fingers and hands at work—glitches collected sometimes in wonder or in satire or critique (A. Wilson; K. Wilson). The hand seems placed on the page like a benediction or ban. Unique, it cannot therefore serve as an emoticon. Still it must be meaningful, indexical of something, and yet it is a communication without a known source or audience. Such images are artifacts of the many coordinated human actions necessary before readers can perceive on their browsers the encoded file that represents a printed page. At the same time, the hand blocks the legibility of that page and seems to warn that the original source, the book, is endangered by this capture from the wild of the library shelves. As Klein insists in this issue, the labor must be visible. Sari Altschuler and David Weimer, also in this issue, challenge the prevalence of the visual, advocating digital recognition of texture and touch. Bibliographers for some time have called attention to the materiality of the document. Against the common supposition that digital media are immaterial and free as a cloud, or the cloud, digital humanists today and in this special issue foreground the materiality and costs of these media. Booth and Posner thought that one of the Google hands would make an appropriate cover image for the material at hand, conveying the feel of the essays that follow.

The Google Books images shared by the HathiTrust Digital Library do not meet the high-resolution specifications for a cover image. Christina Deane, director of digitization services at UVA, suggested that we stage the image, and we saw the advantage of a collaborative piece, a twist on the unintentional portrait of an anonymous hand wearing the telltale tools of the scanning trade, fingerprint prophylactics, or cots. The cover photograph, unlike the genre of image to which it alludes, is a collaboration of named volunteers. The infrastructure of a major university library made it possible, but no money changed hands, although Booth did order a packet of one hundred pink cots from Amazon for pennies apiece. Hanni Nabahe, resident librarian at UVA, fluent in Spanish and English, was the first person we asked to be the model for the hand, and she was willing to collaborate in a representation intended as a critique of the legacy of colonialism in the global production of textual data. With the expert collaboration of UVA's librarian for English, Sherri Brown, Booth explored volumes of early geographies in special collections. We sought a text in a language other than English, in keeping with Élika Ortega's view of the cross-cultural potential of DH projects, that would touch on the early modern worldview documented in James Jaehoon Lee and Joshua Beckelhimer's essay, in which race is differentiated by climate. After viewing enormous folio maps and a black-letter history of the conquest of Mexico, Booth chose the convenient resource for explorers, Gerónimo Girava's *Dos libros de cosmographia* (*Two Books of Cosmography*), printed in Milan in 1556. Having selected some possible pages, Nabahe, Brown, and Booth met in the studio with the freelance photographer and digital production technical lead Eze Amos. Amos and Nabahe devised many images. We began to favor the blur, both of Nabahe's motion caught by the camera and the pages that would not keep still under such anachronistic attention. Here we flutter the "Tabla de los XXIII Climas" ("table of the

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

twenty-three climates"; 38–39), as if gesturing toward a new cosmography for the changing climate (Nowviskie).

What Is Here and What Is Not

Although we planned and strove for variety, we have been surprised by the eventual shape of this special issue—as is, perhaps, the ideal outcome. We urge readers to prepare to be surprised as well. Guest editors of special issues of *PMLA* organize a convention session and invite two full-length essay contributors; they can also solicit contributors to sections like Theories and Methodologies. Yet the thorough peer review of this journal means that guest editors of a special issue do not control most of the contents. Because of *PMLA*'s commitment to modes of research associated with departments of English and languages, and hence to textual humanities, the pool of submissions was not deep in arts and design, though it encompassed the spatial turn (environmental humanities, mapping, visualization of social networks). Nor did our call bring in essays in the field of rhetoric and composition, although digital scholarship is thriving within this expanding area. Likewise, we saw few examples of digital work on photography, film, or television, despite the growth of interest in these older media in DH at large. Yet some contributors to this issue work in media studies departments and DH research centers or libraries, or are affiliated with information schools or data science programs. This issue's variety materialized through our intervention and our good fortune; the essays include examples of classical and early modern studies; studies of poetry as well as of prose; contributions to early American or contemporary African American studies; postcolonial theory and linguistics; sound, touch, and image studies; and beyond. Below we offer some observations about the aims and themes of the very different contributions gathered in this special issue.

Distant reading has a presence among the full-length essays, not least in articulated alternatives to it. In the first of several essays challenging gender formations, Klein's "Dimensions of Scale: Invisible Labor, Editorial Work, and the Future of Quantitative Literary Studies" counters the idea that the point of large-scale textual analysis is to pull the camera back far enough to discern patterns that would not be evident up close. Instead, such analysis might be a way of drawing our attention to underlying structures that were there in front of us all along—for example, the labor of editors, secretaries, and librarians, which might otherwise go unnoticed. Klein analyzes two sets of abolitionist newspapers, using several computational approaches to draw out the features that define the work of specific women editors. Given the troubled and racist history of quantification, the challenge according to Klein is to "wield the power of quantitative methods to contest these oppressive histories at the same time that we admit what we cannot and, at times, should not know."

The cultural influence of changing technology has long been a motif of research in DH, as it has in contemporary fiction. Alexander Manshel has applied digital methods to engage with a very modern problem: the "lag" between new technologies' popularization and their appearance in novels. Manshel uses computational methods to test the popular sense that authors at the end of the twentieth century hesitated to incorporate new technologies, such as cell phones and the Internet, into their novels' plots. The lag is always with us, Manshel observes, but it does not remain constant over time: the gap between technologies' adoption and their uptake in novels varies, most significantly diverging in popular and literary fiction at the end of the twentieth century.

Like several pieces in the Theories and Methodologies clusters in this issue, Richard Jean So and Edwin Roland's "Race and Dis-

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

tant Reading" addresses the fraught difference of race. Much as Susan Brown cautions us that metadata can reinforce gender binaries, So and Roland recognize that computational analysis can easily descend into racist stereotyping. They offer an approach whereby assumptions about race—in terms of literary authorship—might be illuminated by distant reading. They take as their test bed two corpora of 220 novels each, one by white writers (sorted by genre) and one by black writers, using "reflexive" modeling to show that textual features are consistent or highly predictive that a text belongs in its respective corpus. They find that features associated with the black or the white corpus show variance, however, when found in the opposite corpus. The greater variance in the corpus of black novels—the greater "variance in how novels by black authors use white features"—means that novels by black authors are five times more likely than those by white authors to be misclassified by machine reading. Instead of seeking a true bedrock of racial difference in writing, as this summary might suggest, So and Roland try to reconcile "distant reading and the critique of race."

Whereas some of the essays in this issue, including those by Ted Underwood and Liu, endeavor to rechart the territory of the field while renaming the things revealed in this new landscape, Altschuler and Weimer explicitly offer a manifesto. Like Klein and others, Altschuler and Weimer call for attention to what is too often belittled or ignored: embodied experience, the evidence immediately at our fingertips. Not that Altschuler and Weimer think small, but their manifesto asks us to think by touch, which could extend to "distant touching" of the texture of printing over time. "Texturing the Digital Humanities: A Manifesto" challenges "the assumptions most people make about which senses are involved in reading," and, with examples of digital representations of books for blind and low-vision readers, imagines tools for feeling books that could be available in the near future.

If we can have distant readings of touch, race, and gendered labor, we can certainly learn more about distant reading from those who extensively deploy algorithms on large corpora and cheerfully debate the rationale for such deployment (and share the nuts and bolts on digital repositories like *Github*). Underwood, in "Machine Learning and Human Perspective," seeks to go "beyond hype (and alarm)" about quantitative work with bags of words, acknowledging, for instance, that "distant readers are always working with samples" rather than "the whole library." In a somewhat similar process to that of So and Roland, Underwood has curated samples of novels in various subgenres (according to catalog labeling) and divided the samples into novels by male authors and novels by female authors. Those samples become training sets for machine learning algorithms, which use word patterns to generate predictive models for classifying texts. Underwood's acceptance that perspective underlies genre reception chimes with Brown's argument that categorization is essential to any analysis of data. Beginning with the recognition that gender is a mutable role and that genre categories change with usage, Underwood proposes that it is human readers' choices that define genre as a practice of reception. He is interested in the creative experimental design for parsing digitized corpora that we can place in dialogue with other humanities research on gender and genre. As in a wide range of algorithmic textual analysis, the results may not surprise; having made a pudding from a new and reproducible recipe is more important than the specific proof in it.

Words as tokens in digitized texts underlie the research depicted in the next essay, "Anthropocene and Empire: Discourse Networks of the Human Record," though the perspective is markedly different from the one offered by Underwood's corpus of 94,000 novels from HathiTrust, the novel in English being itself

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

a phenomenon of recent centuries. Instead, Lee and Beckelhimer investigate the origins of the Anthropocene. In Early English Books Online (EEBO) and HathiTrust, various texts offer clues to explain a shift in the way that English speakers thought about the globe around 1610, a date that researchers have identified as key to the increase of the amount of carbon in the atmosphere. Lee and Beckelhimer find that latitude, not hemispheres, defined early modern conceptions of the globe. "Colonial policymakers thought of latitude as a systematic network organizing the global distribution of life and used it as a central tenet of their colonial enterprise." Under the division of the world into "climates" and "zones," race was conceived in terms of human attributes shaped by climate and was "transplantable." "The malleability of the humoral body was exploited as a central tactic of England's colonialist regime," write Lee and Beckelhimer.

In keeping with the essays in this issue that call for gender, racial, or sensory specificity in well-designed digital research, Liu offers a unique take on the widespread advocacy of diversity among digital humanists today. Moving beyond the metaphor of the big tent, which he notes is a holdover from a discourse of variety prevalent in the golden age of gospel meetings and circuses, Liu offers "Toward a Diversity Stack: Digital Humanities and Diversity as Technical Problem" as a set of new directions in the light of world literature and cultural differences, arguing that high-dimension mathematical spaces can replace the big tent. Not unlike Ortega's call for media-cultural hybridity in the Digital Ecologies with a Difference cluster in this issue, Liu's essay sets out layers in a "diversity stack," with sections on multilingual DH, multimedia DH, representative DH corpora, the DH chronotope, and the DH theory of identity. In detail and in spirit, the technological applications Liu recommends would, we feel, support the goal laid out by Ortega (who cites Liu in her essay)—to unite media hybridity with cultural hybridity in the contact zones of new media in the world. Liu, for instance, foresees image-recognition software applied to the arts worldwide rather than to European art only; extensible ontologies or a "diversity XML"; metadata-analysis tools that assess any collection's contents according to a set of diversity measures and then offer the user paths to find the materials missing from the archive. At the top layer of the stack, where investigators expressly commit to diversity, Liu cites the innovative work of Lee and his collaborators, who use word vectors to analyze the EEBO corpus, a project on the early modern geographic conception of race that certainly models a representative corpus and inclusive chronotope. The sequence of interrelated long essays in this issue, then, can lead a reader through varieties of DH research designs, premises, goals, and envisioned futures.

Digital Ecologies with a Difference, our first cluster in the Theories and Methodologies section, offers three vivid takes on matters of identity that were broached in the long essays and that offer particular perspectives on African American studies, Latin American studies, DH and media, and feminist studies of women writers. "Let's cut to the chase," writes Howard Rambsy II. "African American scholars occupy the margins of this expansive realm known as digital humanities." Rambsy examines some of the reasons for this marginalization, pointing out that African American scholars, already a marginalized group, are often forced to make a choice between engaging in African American literary scholarship and taking part in the DH community. Rambsy looks closely at a number of African American scholars who identify as digital humanists, discussing how they negotiate between these disciplines and how their success might be replicated by other scholars. As noted, Ortega amplifies the call for diversity in the DH field that Liu addresses in his essay, and she anticipates Meredith D. Clark's embrace of multimedia artifacts in the sec-

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

ond cluster. In "Media and Cultural Hybridity in the Digital Humanities," Ortega offers a theoretical framework for understanding DH in its current incarnation, as it painstakingly moves beyond a monolingual Internet and Eurocentric history. The efflorescence of transmedia and interdisciplinary work, she argues, reflects the potential for DH to foster a rich hybridity of cultures. The challenge of combining computation with the analysis of power is also at the heart of Brown's article on how to develop schemas that represent gender. Brown examines the often binary options of categorization in DH metadata. A researcher's decision to classify a person or entity is both ethically fraught and quite necessary if one is to make sense of large amounts of data. Such decisions take on special weight when they relate to axes of power, such as gender and race. Brown outlines the approach she and the Orlando Project and the Canadian Writing Research Collaboratory teams have taken to address the problem of how to classify gender in texts with nuance and respect.

Although the contributions to both clusters speak to one another, we asked contributors to the second cluster, Codes and Transformations, to emphasize varieties of technological interventions. In his book *Cyberformalism*, Daniel Shore laid out a method of discovering linguistic forms such as "X is the new Y" in large textual databases. In his essay here, "The Form of Black Lives Matter," Shore explicitly eschews textual analytics trained on individual words, keeping his eye on what would be discarded: "properties of language—syntax, structure, order, compositionality, categorization, idiomaticity." He parses linguistic forms of the present day: memes that work as both Foucauldian discourses and collective keyword generators. The essay begins with a memoir-like encounter with variants of the hashtag #BlackLivesMatter. He uses Mark Davies's database of words from Web-based newspapers and magazines, *NOW Corpus*, to show permutations of the phrase *X lives matter* that make claims for the value for minoritized groups or that dismiss those claims.

From everyday natural language, we shift to a digital project that artfully reads poetry and sound. The poet Katharine Coles demonstrates how to close read the sounds of a poem with a digital tool that she and another poet, Julie Gonnering Lein, developed with the visualization scientists Miriah Meyer and Nina McCurdy. Whereas Altschuler and Weimer seek to bring texture into DH, Coles and her collaborators strive for innovative visualization to represent another sense, hearing, as they quantify poetic sound. The digital tool RhymeDesign detects transmutations of verbal sound across the words in a poem, which are displayed in the user interface, called Poemage. By using this interface to perform close readings of Emily Dickinson's "He fumbles at your Soul" and other poems, Coles shows how Poemage can train readers to recognize the deformation of expected meter and rhyme and the sound and speed of certain words—to read poems anew. Coles's essay challenges the supposition that *techne* is not consonant with *poiesis*—as do many of this special issue's advocates for technological approaches to the interpretation of forms.

Rachel Sagner Buurma and Jon Shaw's "Slow Metadata" attests to a different kind of annotation, the thorough description of eighteenth-century English novels in the Early Novels Dataset (END) project, which annually convenes a group of students, scholars, and librarians to apply custom-made extensions of machine-readable cataloging (MARC) code to the paratexts of eighteenth-century novels. As in other projects described in this issue, the team pays meticulous attention to the process of capturing data and to the loss of specificity this entails for the unique object of study—here, an eighteenth-century book and a card catalog's record of it, which is in turn compressed in machine-readable form. The participants in END resist speedy equivalence.

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

For instance, they do not accept *Pamela*, the convenient modern title of Samuel Richardson's novel, as a substitute for the original title, which is more than sixty-five words long. A team can do its best to make slow judgment calls using "the nuanced but standardizing languages of bibliographic description, library cataloging standards, MARCXML encoding, and tab-separated-value (TSV) files." Such metadata help open up new collections of paratextual features, such as epigraphs that recur across eighteenth-century books. The collaborative project offers a synopsis of early printed volumes, library cataloging in the last century, and the way today's forms of data are applied in bibliographic theory and poetics.

From twentieth-century science fiction to nineteenth-century poetry to eighteenth-century novels to early modern texts in varied genres, the historical range of this issue extends with the contribution of the classicists Elton Barker, Anna Foka, and Kyriaki Konstantinidou. Like the END group, they venture into the age-old practice of annotation, transformed in the digital age, much as word searching and linguistics or poetic scansion have been altered with new technology. The software *Recogito* allows researchers to enter annotations easily in The Digital *Periegesis*, a project that aims to produce a digital version of Pausanias's *Periegesis Hellados* (*A Description of Greece*). Their project attends to the highly shaped description, not to be mistaken for a real travelogue, by making meticulous annotations that link place references in digitized documents to "global place authorities, known as gazetteers." Barker, Foka, and Konstantinidou, like Buurma and Shaw, note the slow (as opposed to distant) teamwork entailed in the project. Spatial humanities, far from simply mapping a text, can reveal topographic, chronotopic, and analogic encoding of space in the apparent movements indicated in the original written document.

Reading in and among the contributions in this issue, we have had the feeling of using the zoom function in an interactive geospatial timeline. Clark's "Picture This: The Screenshot's Use in Digital Humanities Scholarship" returns to the now of Shore, Liu, Ortega, Rambsy, and others, considering recent tools and institutions. Clark focuses our perception on a "data-collection tool available to just about anyone with access to a computer screen or a smartphone: the screenshot." Clark plays a variation on the theme of "the receipts" that racially profiled people keep as a precaution against the next interpellation as suspect. Screenshots as receipts not only keep a record of ephemeral online interactions but also aid "qualitative media studies of complex online cultures" and political movements. Clark hints at an aesthetic interpretation of the decontextualized capture of a moment, a way of seeing through the illusion that it offers a view from nowhere. Like an advocate of microcomputing, Clark praises the accessibility and low-tech threshold of "a little-'d' democratizing tool" for studying media today.

Extensible Human Handiwork

Among the insights that experience in DH can yield for researchers accustomed to existing university divisions of labor is that, often, the dataset is the work: the organization, collection, and sharing of structured information constitutes intellectual labor. The essays by Buurma and Shaw; Barker, Foka, and Konstantinidou; and Brown, for example, demonstrate the sensitivity and knowledge required to assemble clean and structured data that do justice to the phenomena they purport to represent. Creating and sharing such datasets may seem unprecedented in the humanities, though bibliographers, folklorists, musicologists, and others have collected and taxonomized in similar modes in the past. We can expect to be challenged in the next few years to recognize and reward such work as scholarship. Recognition and reward might help counteract the tendency to display and ad-

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

mire results as if the data were speaking for themselves on a neutral world stage. In impressionist criticism that prevailed before and after 1900, we can trace a similar tendency to exhibit ineffable wonders, though the objects might be single works of genius rather than massive corpora. We can do better, and many working in DH are benefiting from the intervening generations of rediscovery, theory, and critique that somewhat demystify how and at whose expense cultural production happens.

The materials at hand in the essays in this issue are sometimes as small as a printed title page and sometimes as extensive as 94,000 books, but the researchers recognize the interpretative dimensions of how they have munged their data and how they have recruited algorithms or collaborators to work toward meaningful outcomes. Our remarks here foretell the rich varieties of inquiry in the essays themselves. To the extent that the essays in this issue are united by a common question, it might be this: What is the appropriate relation of computation to literature? Computation here may mean programmatic reading of extensive corpora, communicating perspectives on social media or on a single text using visualization software, or practicing African American studies or feminist theory while engaging in digital research at the same time. No authors take exactly the same view. They are uniform, however, in their insistence on interrogating their own methods. Such is, perhaps, the benefit and also a hazard of a field whose definition changes under our feet: a constant need to demonstrate and explain the validity of the methods we choose.

We have reflected on the conditions of our own DH curriculum and research infrastructure and examined the potential divergence of some streams of DH from computational humanities or data science. The hand on the cover is not conjuring to predict the future or seeking to command a single, static discipline. On the one hand, essayists in this issue may be more likely to attend conferences on African American studies, American literature, media studies, classics, or other fields than to attend those on DH per se. Members of schools of data science, on the other hand, preoccupied with machine learning to be used for corporate search engines or investment funds, will not drop everything to engage with the persistent and groundbreaking questions that animate these authors. Everyone involved in higher education will face a globe in environmental and economic flux and will continue to grapple with consequences of the exploration heralded by Girava's *Cosmographia*—a known world of empire, biological essentialism, patriarchy, and economic inequality as well as technological innovation. The gun was initially regarded by many Europeans as a positive sign of civilization (Satia). We should know by now not to regard every piece of technology as an unquestionable boon.

We seem to learn more every day about the potentials and perils of computation at scale: we embrace Amazon's book recommendations even as we see how AI is being deployed to target society's most vulnerable members. The essays in this issue remind us, if our daily lives are not proof enough, of how vital is the need for students of the humanities to understand advanced computation and to shape how data is captured and used. In that hypothetical future in which this issue represents a kind of screenshot of DH before the changes to come, we may find that our robotic avatars are speaking a different language. May they be endowed with memory and some form of stewardship for the strangeness of pre- and extradigital stuff. The material at hand now may be sufficiently challenging for us.

Notes

1. Designations such as "digital scholarship" can be intended to represent the fact that DH often encompasses

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press

work outside the humanities. In 2019, Temple University launched a graduate certificate in cultural analytics, "sometimes called 'digital humanities'" ("Graduate Certificate").

2. Some propose sentiment analysis as a distant verbal measure to compare the plots of many novels (Archer and Jockers). In biography, a mourning community or a deathbed scene often signals the protagonist's vocational success.

3. See Ahmed et al. Folklorists, computer scientists, linguists, political scientists, and others have collaborated internationally on big data projects funded by the Digging into Data Challenge ("Winners").

Works Cited

Ahmed, Manan, et al. *Torn Apart / Separados*. Vol. 1, 2018, xpmethod.plaintext.in/torn-apart/volume/1/.

Archer, Jodie, and Matthew L. Jockers. *The Bestseller Code: Anatomy of the Blockbuster Novel*. St. Martin's Press, 2016.

"Charter." *Scholars' Lab*, University of Virginia Library, June 2016, scholarslab.lib.virginia.edu/charter/.

Cohen, Daniel, and Thomas Scheinfeldt, editors. *Hacking the Academy*. U of Michigan P, 2013.

Croxall, Brian, et al. *One Week, One Tool*. Center for History and New Media, George Mason University, 2013, oneweekonetool.org/.

D'Ignazio, Catherine, and Lauren Klein. *Data Feminism*. *MIT Press Open*, Nov. 2018, bookbook.pubpub.org/data-feminism.

Drucker, Johanna. *Graphesis: Visual Forms of Knowledge Production*. Harvard UP, 2014.

@FolgertK (Folgert Karsdorp). "I'm thinking about developing a workshop/event/journal/community for computational research in the humanities that doesn't exclude people with profound computational skills and knowledge. Who's in?" *Twitter*, 16 July 2019, 12:31 p.m., twitter.com/FolgertK/status/1151167545539477504.

Gallon, Kim. "Making a Case for the Black Digital Humanities." Gold and Klein [2016], pp. 42–49.

Girava, Gerónimo. *Dos libros de cosmographia*. Milan, 1556.

Gold, Matthew K., and Lauren F. Klein, editors. *Debates in the Digital Humanities 2016*. U of Minnesota P, 2016, dhdebates.gc.cuny.edu/read/untitled/section/14b686b2-bdda-417f-b603-96ae8fbbfd0f.

———, editors. *Debates in the Digital Humanities 2019*. U of Minnesota P, 2019, dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/0cd11777-7d1b-4f2c-8fdf-4704e827c2c2.

"Graduate Certificate: Cultural Analytics." *Temple University*, bulletin.temple.edu/graduate/scd/klein/cultural-analytics-certificate/#text. Accessed 18 Nov. 2019.

Johnson, Jessica Marie. "Markup Bodies: Black [Life] Studies and Slavery [Death] Studies at the Digital Crossroads." *Social Text*, vol. 36, no. 4, Dec. 2018, pp. 57–79.

Johnston, Andrew, et al. "Pavilion IV Garden." *Narratives of Slavery at Jefferson's University*, Institute for Advanced Technology in the Humanities, 2016, fieldmethods.iath.virginia.edu/16SpARH5610/pavilion_iv_garden.

Kirschenbaum, Matthew G. "What Is Digital Humanities and What's It Doing in English Departments?" *Debates in the Digital Humanities*, edited by Matthew K. Gold, U of Minnesota P, 2012, pp. 3–11.

McCarty, Willard. "Modeling: A Study in Words and Meanings." *A Companion to Digital Humanities*, edited by Susan Schreibman et al., Blackwell Publishing, 2004, www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-3-7.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York UP, 2018.

Nowviskie, Bethany. "Digital Humanities in the Anthropocene." Digital Humanities 2014 Conference, 8–12 July 2014, Lausanne, Switzerland. *Bethany Nowviskie*, 10 July 2014, nowviskie.org/2014/anthropocene/.

Posner, Miriam. "What's Next: The Radical, Unrealized Potential of Digital Humanities." Gold and Klein [2016], pp. 32–41.

Risam, Roopika. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Northwestern UP, 2018.

Satia, Priya. *Empire of Guns: The Violent Making of the Industrial Revolution*. Penguin Books, 2018.

Schmidt, Ben. "How Badly Is Google Books Search Broken, and Why?" *Sapping Attention*, 10 Feb. 2019, sappingattention.blogspot.com/2019/02/how-badly-is-google-books-search-broken.html.

Shore, Daniel. *Cyberformalism: Histories of Linguistic Forms in the Digital Archive*. Johns Hopkins UP, 2018.

So, Richard Jean. "All Models Are Wrong." *PMLA*, vol. 132, no. 3, May 2017, pp. 668–73.

Trouillot, Michel-Rolph. *Silencing the Past: Power and the Production of History*. 2nd ed., Beacon Press, 2015.

Villarroel, Carolina, et al. *Recovery: Recovering the US Hispanic Literary Heritage Blog*. recoveryprojectappblog.wordpress.com/. Accessed 30 Oct. 2019.

Visconti, Amanda. "What about Blogging Keeps Me Blogging?" *Scholars' Lab*, University of Virginia Library, 2 June 2019, scholarslab.lib.virginia.edu/blog/why-blogging/.

Wilson, Andrew Norman. *Scan-Ops. Andrew Norman Wilson*, 2012, www.andrewnormanwilson.com/ScanOps.html.

Wilson, Krissy. *The Art of Google Books*. 2017, theartofgooglebooks.tumblr.com/.

"Winners of Round Four of the Digging into Data Challenge." *Digging into Data Challenge*, 29 Mar. 2017, diggingintodata.org/awards/2016/news/winners-round-four-t-ap-digging-data-challenge.

https://doi.org/10.1632/pmla.2020.135.1.9 Published online by Cambridge University Press