NOTE: This is an unrefereed author’s version of an article submitted for consideration for publication on February 26, 2014.

The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast

By Lara Putnam

The transnational turn has been happening simultaneously with the digital turn and the implications of this entanglement are profound, although as yet largely undiscussed. The great bulk of methodological discussions of "history in a digital age" have so far centered on approaches that harness computational tools to reveal patterns in large sets of textual or mixed sources—emerging techniques of "text-mining" and "distant reading." But more pervasive shifts brought by the internet age are working a much broader impact on what historians do and how. Only a tiny fraction of us are tackling "big data" with quantitative tools. Vastly more of us use the search functions of Google, Google Books, JSTOR, digitized newspaper databases, Ancestry dot com, and the like as we track down qualitative information on particular topics, people, places, or eras.

1 I am very grateful to Julie Greene, Diana Paton, Christian De Vito, and Laura Edwards for comments on earlier versions of this essay.


Precisely because those digital tools simply scale up and accelerate kinds of information gathering historians already did, their integration into our practice has felt smooth and continuous rather than revolutionary. But increasing reach and speed by multiple orders of magnitude matters profoundly. It makes whole new realms of connection visible. It renders whole new kinds questions askable and then answerable. If we recognize this, we can then also see that the new topography of disciplinary possibility has systematic blind spots and tricky shortcuts. It enables ignorance as well as knowledge. The conjoined transnational and digital turns demand caution as well as celebration.

Analyzing this requires treating as remarkable that which has become, almost overnight, quotidian. The informational landscape within which historians labor has been transformed by two overlapping and accelerating trends. First, beginning in the late 1990s, the time cost of accessing secondary texts sought out by title or topic dropped precipitously, as JSTOR and publishers’ initiatives made swathes of scholarship accessible via the web. Secondly, beginning in the mid 2000s, access to both primary and secondary sources by granular content rather than topical label exploded, as Optical Character Recognition (OCR) processing made full-text searchability the norm and Google Books and newspaper and other digitization projects boomed.

Those linked shifts in the accessibility of qualitative information are a sea change at the core of our collective disciplinary practice. If we have hardly remarked it, it is because the same changes have permeated daily life in the wi-fi-ed world in the same years. How can typing words into a search box—which feels as revolutionary as oatmeal, and indeed accompanies oatmeal at many breakfast tables these days—be a sea change?

But those of us who started graduate school on the other side of the digital revolution can cast our minds back to what information acquisition used to look like, and measure the transformation wrought. Instant access first to topic-identified and now to term-identified sources, uploaded from an increasingly broad swath of the globe, has made sideways glancing outside the boundaries of place-based expertise effortless rather than extraordinary. Text-searchability has made tracking individuals, place names, phrases, titles, institutions, and organizations across hundreds of thousands of publications a viable data visualizations—received few responses in AHA’s survey.” Rwany Sibaja, “Teaching and Learning History in the Digital Age,” March 11, 2011 http://teachinghistory.org/nhec-blog/24526. His conviction that the “digital cameras, scanners, search engines, word processors, and online archive searches” now pervasive cannot have a radical impact on “how we ‘do’ history” seems widely shared. But see Tim Hitchcock’s assertion that since “online text... now forms the basis of most published history,” it is urgent to assess “the critical impact of digitisation on our intellectual praxis.” http://journalofdigitalhumanities.org/1-1/academic-history-writing-and-its-disconnects-by-tim-hitchcock/
method for tracing micro-dynamics of border-crossing processes. Together, these two practices—I'll call them side-glancing and term-fishing—radically change the kind of questions that historians are likely to ask and the kind of stories we are able to tell.

The impact of a digitized past on the knowability of past processes, of whatever scale and in whatever locale, is significant. But the impact on the knowability of supranational or transnational processes is overwhelming. To see why, we need to consider the topography of physical information that long shaped disciplinary possibilities and guided individual scholars’ choices. Once we grasp how radically that topography has changed, we can assess the costs as well as the benefits. We can scope the blind spots of the brave new world of sources at our fingertips, and we can ask what transnational history loses when the real-world friction that research on international topics once demanded is radically reduced.

**The real-world geography of textual sources**

Information in physical form (whether manuscript documents, government publications, scholarly articles, or books) tends to cluster in administrative centers near where it was produced. And, with important exceptions—exceptions shaped predictably by imperial rule—information tends to be produced in the places that information is about. Thus in a world before digitization and web access, historians' implicit geography of informational potential—"Where should I look in order to find out?"—spread outward in steps something like this: 1) My institution’s library; 2) The biggest university library I can reach; 3) Archives and libraries in the relevant imperial center (Rome/Paris/London/Washington D.C.); 4) Archives and libraries Near Where It Happened.

Given that the bulk of historical research was conducted by scholars based in resource-rich settings, disproportionately clustered in North America and Western Europe, utilization of information about most “It’s that had happened elsewhere was shaped by extremely high fixed costs: airplane tickets and house-hunting, say, even if language acquisition wasn’t an issue. Fishing expeditions that defied that predictable geography of informational access were prohibitively expensive. We looked for information in books we knew would have it, or in archives we knew would have it, or in newspapers we knew would have it.

---

Source-anchoring reinforced the nation-state bias that was built into our discipline from the start. The founders of nineteenth-century nation-states had carefully founded national archives so that what in Latin America is called “Historia Patria” (affirmative history of the nation) could be written. We followed the geography those founders had traced whether we wanted to or not. What was the alternative? Even the quickest lateral glance would require a trip to the library; a hunt through a card catalogue for recent publications on, say, the country next door to one’s usual target; a perusal of the footnotes to see what the neighbor’s archives might hold; a separate search for a mailing address for the archive; and a letter of inquiry as to what institutional series, for which years, in what quantity might be available. Did I mention the postage stamp? And all of that comprised no more than the fishing expedition necessary to discover whether a fishing expedition might even be desirable.

The national scope of archives and libraries created great economies of scale for nation-specific information—economies of scale that functioned as a distorting subsidy, making it much cheaper to trace processes within a nation’s borders than to determine whether they ended at the nation’s edge. Once you knew a lot about, say, modern Mexico, finding out one more thing about modern Mexico was easy. You owned some of the right books, you knew where others were, you knew which guides indexed scholarly publications for that topic: you might even already have a trip to Mexico City planned. Under these structural circumstances, it made perfect sense for those who sought to study the past to invest heavily in national and regional expertise. The marginal value of extraregional knowledge was slim. An extra day’s reading on nineteenth-century Guadalajara could point toward a new filing cabinet in the Archivo General and a significant intellectual payoff. A single day’s reading about Prussia got a Mexicanist no closer to Berlin.

Thus, we trained graduate students in national or at most regional historiographies; we looked in card catalogues under MEXICO—SOCIAL LIFE AND CUSTOMS for publications we might have missed. Investigating the history of places smaller than the nation was both possible and common. But when seeking to contextualize or compare those local or regional findings it was the national frame that seemed most apt—most vital for debate, most open to collective advance in knowledge—because it really was.

Writing a history of labor migration to and through Caribbean Costa Rica for my dissertation in the late 1990s, I read much anthropological and sociological literature on present-day “transnational migrants,” and noted the similarity of what I read there to the turn-of-the-century lives my sources captured. But “my sources” meant the sources I could track down in person, which meant judicial cases from the Archivo Nacional de Costa Rica, newspapers held at the Biblioteca Nacional de Costa Rica, and the few travellers’ accounts

---

5 Literature I discovered through footnotes and photocopied from bound volumes page by page, if you can believe it.
with descriptions of the relevant province that I managed to stumble across in other people's footnotes or on Widener Library's shelves.

It's like the joke about the drunken partygoer looking for his keys under a lamppost: “Didn’t you lose them on the other side of the street?” “Yeah, but the light’s better here.” Economists tell the joke to encapsulate their dependence on questions for which statistical indicators can be found. But it is equally relevant to the qualitative research the great majority of historians rely on. Before the Internet, it was certainly possible to go off looking for information outside of the lamppost of a given national archive’s index files. Doing so, though, was the equivalent of arming yourself with a penlight and heading off to search for lost keys in the dark. The odds that you were going to stumble upon anything that would make the effort worthwhile were overwhelmingly slight.

So even though I was explicitly looking for evidence of connection and circulation across borders, my research strategies were circumscribed by them. My best bet was to count on the anchoring of documentation to place. I poured over judicial cases from the port city at the nexus of migration, since testimonies occasionally mentioned the international relocations that shaped working lives, and sought out oral history transcripts from places I knew migrants had come from—like Jamaica—or where I knew some had stayed—like Limón, Costa Rica.

This points to a broader truth. Obviously, historical research into transnational processes predates Google Books. Indeed by the mid 2000s attuned commentators were already describing a “transnational turn” accomplished, reflecting a rising wave of publications each year in the making. They had been conceived and largely conducted in a world of grounded information that made certain kinds, but only certain kinds, of transnational research possible. Projects often followed the contours of a single institution that had generated and assembled data from multiple sites. The British Empire at one’s fingertips from the reading room at Kew is the iconic example, but important too were non-state institutions like the Rockefeller Foundation, into whose archives historians were invited in

the 1990s, producing a wealth of studies that illuminated the interactions and exchanges that shaped the development of public health in the Americas.7

In another long-standing model for transnational research in an analog world, a single scholar might build expertise on a small number of sites whose density of connection remained visible, making in-depth knowledge-acquisition a reasonable investment. Thus, in order to understand migrants’ lives in turn of the century Limón, it made sense to become as expert as I could in the histories of both Jamaica and Costa Rica. That investment paid off as knowledge of secondary ports’ trajectories in Jamaica, for instance, allowed me to identify patterns in the parish-specific origins of migrants in Limón that might otherwise have seemed meaningless.8

Densely knit transnational systems, then, could be studied pre-internets, and were. But peripheral vision was prohibitively expensive. Border-crossing movement of people, ideas, or goods that was low-intensity, diffuse, and extra-institutional tended toward invisibility even if it created recurrent patterns or enduring circuits. Those things were in motion far from the lamppost, and the penlight’s beam was tiny.9

The digitized world: An information landscape re-priced

It surely is not purely happenstance that the “transnational turn” has accelerated in the same decade in which suddenly commonplace technology has radically reduced the cost of acquiring information about people, places, and processes outside the borders of one’s


9 It is worth noting, too, that before recent years the most impressive works uncovering transnational processes tended to be written toward the end of long careers, reflecting the painstaking accumulation of expertise and evidence amassed over years of “gathering string” on a topic (one thinks, for instance, of Sidney Mintz’s Sweetness and Power: The Place of Sugar in Modern History (New York: Viking Penguin, 1985)). The current boom in transnational dissertation topics is in this sense a particularly radical shift. That such projects seem feasible reflects our implicit accounting for technologies that reduce the cost of information gathering. One worries, however, that deep understanding of multiple (nation-based) historiographies cannot be gained as quickly as the profession seems to demand of its young—and that the analysis of the information gathered cannot but reflect that lack.
own prior knowledge. In much of what follows I focus on the digitized accessibility of primary sources. But I want to begin by stressing the vital importance of near-instantaneous access to secondary sources. Be it Wikipedia or JSTOR or Google Books or Amazon’s Look Inside, digitized secondary and tertiary sources allow quick eye-balling of the bigger picture or doings next door: a sideways glance that can uncover supranational connections or simultaneous developments that might be worth looking into. Where were those exports going anyway? What was going on there? Who was that guy before he showed up here? Why was that place sending missionaries (or migrants, or movies) over here to begin with?

The impact of such side-glancing—formerly rare, as each glance would have demanded hours or days of effort with no likely return; now quotidian, requiring only nanoseconds to search and minutes to read—is profound. It routinizes peripheral vision that opens us to the possibility of impacts across borders of varying scales and kinds. It allows us to wonder about connections outside of the institutions, investments, and invasions that analog Historia Patria had already recognized. Seeing matters, a lot. Eschewing formal modeling (for reasons good and bad), historians confidently believe that we know causation when we see it. We do not build in systematic checks against omitted variable bias. Instead we read primary sources with care, noting how different things are related from the author’s viewpoint or our own, and then build stories in which those things cause each other. Just seeing new things can transform historians’ arguments more immediately than in a discipline whose evidentiary paradigm imposed a little more ballast.

Transnational history, key proponents explain, is at core about seeing connections across borders and taking seriously both the connections and the borders. “It does not have a unique methodology,” Patricia Clavin suggests, “but is motivated by the desire to highlight the importance of connections and transfers across boundaries at the sub- or suprastate level, the composition of categories, and the character and exploitation of boundaries.”

Struck, Ferris, and Revel likewise describe transnational history as a congeries of approaches that share “the conviction that historical and social processes cannot be apprehended and understood exclusively within customary, delineated spaces or containers, might they be states, nations, empire or regions. Consequently, all of these tools or perspectives stress the importance of the interaction and circulation of ideas, peoples, institutions or technologies across state or national boundaries and thus the entanglement and mutual influence of states, societies or cultures.”


11 Bernhard Struck, Kate Ferris, and Jacques Revel, “Introduction: Space and Scale in Transnational History,” International History Review 33, no. 4 (2011): 573-584; quote 573-74. Similarly, Pierre-Yves Saunier’s definition focuses on the researcher’s openness to
Note that it is not just that which circulates that one needs to see in order to evaluate transnational connection, but also those entities that might be, through it, entangled: states, societies, economies, cultures. Side-glances that reveal where goods/people/ideas/cultural products were coming from or going to, and what was going on over there at the time, generate hypotheses of entanglement we couldn’t generate without the glancing. When glancing becomes faster by many orders of magnitude, and national boundaries no longer distort our range of vision, the number of transnational hypotheses shaping our collective mission is necessarily going to rise.12

But the impact becomes even greater as text-searchability becomes the new norm. Granularity matters, a lot.13 When discovery of primary and secondary sources relied on someone’s indexing, people, places, and ideas that appeared in secondary roles could only be seen if you knew beforehand where to look. Bit players who never got star billing were finding explanation by observing connections across boundaries: “Going transnational is not moving to a different field of study, shifting allegiances and references. Rather, it is something that many historians can do to find a way to respond questions that lay unanswered on their working desks since a while. Maybe, after all one does not decide to do ‘transnational history’, but it is rather the research one is developing that calls for the development of a transnational angle. To explain briefly what it means, I would accept the simple definition that the transnational angle cares for movements and forces that cut across national boundaries. It means goods, it means people, it means ideas, words, capital, might, and institutions. It may be useful to have a more sophisticated definition later, but that will do for now.” Pierre Yves Saunier: Going transnational? News from down under, in: H-Soz-u-Kult, 13.01.2006, <http://hsozkult.geschichte.hu-berlin.de/forum/id=877&type=diskussionen>.

12 Timothy Burke’s account of what students need to learn as they tackle the hermeneutic process of progressively defining a topic and seeking information on it offers an extraordinary breakdown of all the component steps, informed judgments, and return-on-investment estimates that go into what I call here side-glancing. We do all that, one marvels—and with speed and skill built of countless iteration. 
http://blogs.swarthmore.edu/burke/blog/2011/05/09/how-i-talk-about-searching-

13 See discussion penned in the infancy of Google Books, a whopping eight years ago: Gregory Crane, "What Do You Do with a Million Books?" D-Lib Magazine vol. 12 no. 3 (March 2006): http://www.dlib.org/dlib/march06/crane/03crane.html
invisible, no matter how many plays they were in. That threshold of visibility means that not just specific cases of connection, but whole categories of connection are unrecoverable via analog sleuthing. Here again, that which is made possible by the speed, range, and granularity of digital search resonates precisely with what transnational history announces as its particular contribution.

**Micro and macro: The transnational angle and questions of scale**

Although in some summations transnational history is grouped together with global and world history as privileging the study of the big, those who have theorized it most carefully instead argue that what distinguishes it is its attention to multiple scales of observation and multiple geographic scopes: scales and scopes determined empirically, in accordance with the dimensions of the historical processes under study, rather than a priori by political boundaries. Pierre-Yves Saunier suggests “the transnational” should be understood not as “another scale located near the top of the nested scales, but rather a foray that cut through levels and partly shattered their conception as distinct social entities.”

The distinction between geographic scope and scale of observation is here crucial, for as certain scholars have underlined, very large processes—large in terms of their geographic extension—can be driven by dynamics that function at a small scale of interaction, and are only visible by reducing our observation to that level. Finding ways to


operationalize this insight has been a hallmark of recent work in Indian Ocean history, so much so that some suggest it can be understood as an "Indian Ocean World method" ripe for export.\(^{17}\)

The relevance of the mass digitization of historians’ previously analog corpus to these analytic aims should be obvious. Text-searchable sources make it possible to trace individual people (or songs, or films, or pamphlets, or phrases), allowing us to observe at the micro-level the processes that generate, in the aggregate, macrolevel flows and connections. As repositories digitize and upload information at increasing levels of granularity, the possibilities of using on-line term-fishing for what historians used to call nominal record linkage expand and expand. Ancestry dot com already offers a single portal to a wide range of U.S. government documents, including census sheets, port records, and draft cards, alongside an ever-expanding set of non-governmental texts: obituaries, city directories, and more. Demand from amateur genealogists is driving Ancestry and others to offer access to ever more countries’ nominal records—900 million new records from 27 new countries in 2014 alone, my inbox informs me—with the potential for fruitful poaching by historians expanding apace.\(^{18}\)

Let me give one example of how digitization makes possible the use of microlevel data, sought through term-fishing, to reconstruct the diffuse microdynamics shaping cultural, economic, social, or political exchange. The weekly Limón Searchlight was published by and for the British West Indian community of Costa Rica in the late 1920s. I consulted it on microfilm in the Biblioteca Nacional in San José in 2008, seeking information

an earlier essay that likewise calls historians to engage with urgent public questions, but avoids confusing scale, scope, and evidentiary paradigm, see Barbara Weinstein, “History Without a Cause? Grand Narratives, World History, and the Postcolonial Dilemma,” International Review of Social History 50 (2005): 71-93.


about music, dance, and youth culture in Limon. One 1931 editorial compared the homegrown Central American Black Stars Combination Company to “Benbow’s Follies” in the course of arguing for racial pride among local musicians.19

Three years later, turning my notes into a chapter about the racial coordinates of music and dance styles in the circum-Caribbean migratory sphere, it occurred to me to wonder who exactly these “Benbow’s follies” were. Google Books allowed me—in the space of three fruitful minutes at my desk, rather than a fruitless day at the library—to find out enough about Afro-American showman William Benbow that I knew I wanted to know more. The brief mentions I found showed me his origins in New Orleans and his connections, over decades of touring on the black-owned “chitlin circuit” of vaudeville theatres, to key figures in the 1920s explosion of jazz. Figuring out what the “chitlin circuit” was and where it fit in the entertainment ecology of Jim Crow-era Afro-America required more side-glances, into scholarship outside my discipline as well as my region. But how had they heard of Benbow in Limón? I turned to the (text-searchable) Kingston Daily Gleaner and Pittsburgh Courier and discovered scores of articles and ads documenting decades of circum-Caribbean tours by Benbow’s troupes, something that none of the U.S. music history references had mentioned.

Reconstructing the story of Benbow’s travels, travails, and influences from this suddenly copious mass—one thing you can count on with showmen seeking sales: they know how to make the papers—took much longer. But since I knew from the first term-fishing glance that there was a significant quantity of material to work with, I could make a rational guess that it was worth investing my time. And it was. I found, for instance, that during one stay in Jamaica the troupe sparked a municipal council debate over whether black performers drew stricter obscenity policing than white, illuminating a very different intersection of cosmopolitan blackness with the local politics of race than Benbow had encountered in Garveyite Limon.

The serendipitous discovery of Benbow’s Caribbean tours could not in itself sustain strong claims about impact or patterns, but it could reinforce patterns other sources revealed. Most importantly, it offered a window onto the microlevel dynamics of how performers and performance came to be part of debates over racial belonging at multiple sites. Together with myriad similar findings, it pushed me to argue that early twentieth century circulation and exchange between different sub-spaces of the African diaspora forged new notions of black commonality and, indeed, helped create the very idea of an “African diaspora.”20


The particular intersection of my own research trajectory with the chronology of technological shifts makes me hyper-aware of how much the shifts matter. I first read through the Limón Searchlight at the Biblioteca Nacional while writing my dissertation in the late 1990s, just as Larry Page and Sergey Brin were first meeting at Stanford. I went back to re-read and transcribe in the mid 2000s, specifically interested in music and dance. And I revisited my transcriptions in 2011, reading them on a laptop with instant access to the now text-searchable Courier and Gleaner. The Limón Searchlight was the same source in the same analog format it had always been, but I was reading it differently because the landscape of digitized information availability had transformed which questions about items within it were efficiently answerable, and therefore worth asking in the first place.

This underlines the synergy between term-fishing and side-glancing. Web access to digitized and text-searchable primary sources can make possible microhistorical reconstruction to answer questions only answerable at the micro scale. But the questions often could not have been generated by micro-level observation. In this case, my questions came from patterned results across disparate sites over multiple decades—shifts in musical taste, the emergence of new rhetorics of race. Conversely those large patterns were necessarily invisible in the time-limited local sources necessary to reveal the microlevel dynamics that drove them. Web access to digitized secondary and primary sources was crucial to tracing those international patterns and thus generated the questions that microhistorical reconstruction could answer.

How widespread are the practices I have described here? If a text-mining historian offers N-grams as evidence, it’s as plain as the graph in front of your face. But to what extent have the books and articles you have recently read relied on digital side-glancing or term-fishing? I have given examples from my own work because in regard to others I simply don’t know. Such practices fall into the realm of invisible method, the black box where by consensus we leave so much of our discipline’s heavy-lifting. The extensive discussion of digitization underway in information sciences journals stands in sharp contrast to the silence on this theme in historians’ publications. Are the librarians and database vendors wrong about their target clients? Am I am the only historian hitting search?

Surely not. An American Historical Association-sponsored survey published in 2010 found three-fourths of historians to be “power users” or “active users” of new technology; of them, nearly 100% reported using library-supported databases (JSTOR and kin), over 95% use online search engines in their research, and over 90% use primary sources accessed online.22 A 2013 report by Ithaka S+R (part of the non-profit organization that includes JSTOR), based on in-depth interviews with three dozen historians, captured a remarkably consistent pattern of digitally-pervaded methodological eclecticism—remarkable because each individual confirmed rarely discussing practices of source discovery or organization with advisors, advisees, or colleagues. “Everything in my field is out of copyright and digitized. It’s all there. I feel like I’m cheating half the time.”23 The report’s authors classify all this under “research practice,” which they differentiate from “digital research methods.” As should be clear by now, I disagree. The digital humanities methods currently drawing attention as such are about counting, or graphing, or mapping. But the mass of historians’ research is about finding, and finding out. That so many of us are now finding and finding out via online access to a digitized knowledge base has significant consequences, regardless of whether we count, graph, or map anything at all.

Disparities reduced and disparities introduced

In sum, the radically reduced time-cost, geographic unanchoring, and heightened granularity of information retrieval has transformed the structural conditions shaping the generation of historical knowledge. The nation-based information-market protectionism imposed by brick-and-mortar archives and libraries has collapsed—at least for scholars with the subscription price of admittance (on which more below). The web-accessible “infinite archive” sheds the particular economies of scale that made nation-specific research distortingly cheap and made side-knowledge a bad investment. The time-cost disparities created by those de facto tariffs have disappeared. This has greatly increased the likelihood that historians will formulate hypotheses about causes or impacts outside the national or regional scope of our initial expertise. And it has enabled new forms of old methods, drawn from microhistory among others, to test those hypotheses.

regarding digitized sources’ new role in the hermeneutic process: Fred Gibbs and Trevor Owens, ”The Hermeneutics of Data and Historical Writing,” in Nawrotzki and Dougherty, Writing History in the Digital Age.


Yet the digital revolution is not inherently egalitarian, universal, or cost-free. To the extent that digital research practice has become the essential-but-unacknowledged handmaiden of the transnational turn, the transnational turn now carries more baggage and follows paths more rutted than we have yet to admit. Of course, not everyone wrote national or subnational history before the digital shift, and not everyone is writing transnational history now. The radical re-pricing of the information landscape does not dictate the choices any given researcher will make. But it exerts a strong underlying pull that is reshaping historical production in the aggregate. As a result, the systematic blind-spots, disparities of access, and particular shortcuts that digitized sources make possible will add up to real trends—and real losses—unless we work actively against them.

Most obviously, the universe of digitized text is anything but representative of the temporal and geographic contours of human life in the past. The nineteenth and early twentieth century Anglophone world has been ground zero of digitization. In part, that reflects disproportions in the historical generation of easily digitizable sources. Production of typescript material exploded with the massification of printing and literacy. That massification happened in some places and not others, generating deep geographic disparities in digitizable source generation then and in digital source availability now.

This pattern from the past is exacerbated by disparities in the present. Digitization projects centered initially in English, secondarily in other Western languages. This, though, is changing rapidly, from Oslo to Buenos Aires to Shanghai.24 One can point to initiatives from the small—efforts to digitize Harvard College Library holdings in Amharic, Berber, Mandinka, Oromo, Somali, Swahili, Tigrigna, and Wolof25—to the unimaginably large. The digitization of the copious Chinese language print production of the last thousand-odd years is underway. Over 1.5 million documents have been digitized and linked to a centralized database already as part of the Qing History Project, with the goal of eventually encompassing the roughly 20 million files from the Qing Dynasty held in repositories in mainland China.26 If language skill acquisition among researchers worldwide shifts in rational ways, this will be the most productive frontier of historical research for the next generation.


25 The Africa’s Sources of Knowledge Digital Library: http://www.ask-dl.fas.harvard.edu/

26 As of this writing the database is of restricted access and not yet text-searchable. Liping Mao and Zhao Ma, “‘Writing History in the Digital Age’: The New Qing History Project and the Digitization of Qing Archives,” History Compass 10, no. 5 (2012): 367–374, 10.1111/j.1478-0542.2012.00841.x
The fact that Anglophone overrepresentation in the digitized world is changing so rapidly is what led me to assert, at the start of this essay, that a sea-change is under way for our discipline as a whole, rather than just for English-speakers studying English-speakers’ pasts. In the pages that follow, I focus on dilemmas other than regional and language-based disparities in source digitization precisely because as those particular disparities lessen—as more and more of the world’s textual heritage comes on-line—the range of histories and historians facing the other dilemmas will grow.

A half century ago, E.H. Carr wrote about historians and fish. “The facts are really not at all like fish on the fishmonger’s slab,” he warned those who claimed a simple empiricism. “They are like fish swimming about in a vast and sometimes inaccessible ocean; and what the historian catches will depend, partly on chance, but mainly on what part of the ocean he chooses to fish in and what tackle he chooses to use—these two factors being, of course, determined by the kind of fish he wants to catch. By and large, the historian will get the kind of facts he wants. History means interpretation.”27 A text-searchable world offers Carr’s descendants space-defying superworms. Toss your line in and if the fact is out there anywhere, it will be on your hook in a nanosecond. Yet history still means interpretation. We are still choosing our bait and our tackle. And when we fish in digitized text, we are fishing in a very particular sea.

Who stands in the digital shadows?

Gazing at the past through the lens of the digitizable makes certain phenomena prominent and others less so, renders certain people vividly visible and others vanishingly less so. Firstly, the pages of the periodical press make up a major portion of the raw material now digitally accessible. That means that topics foregrounded in newspaper debate are disproportionately visible, and that the readers, writers, intellectuals and activists who published papers, pamphlets, and letters to the editor are now enticingly within reach. As noted above, historians tend to attribute causality to what we see, without formal modeling to act as a brake on our tendency to do so. If suddenly it is much easier for us to see circulating prints, globetrotting activists, and globe-circling debates, we are likely to start attributing causal impact to those prints, those activists, and those debates. We are going to have to remember not to mistake the window for the why.28

Meanwhile, we are going to have to work actively so that those systematically less present in printed sources do not fall out of view. Size up the absence. Who wasn’t


28 Having just published a book in which travelling activists and periodical press play a sizeable role, I am especially alive to this tension. I find my evidence persuasive—yet recognize the systematic bias of the evidentiary base I rely on. See Putnam, Radical Moves, Ch. 4.
In some ways, this problem (sometimes shorthanded as that of the “voiceless,” but actually that of the “rarely recorded on their own terms”) is not new. Social historians in the 1970s and 1980s spent long hours compiling data by hand: aggregating state-generated sources to track demographic shifts, labor patterns, and market trends and at least show us the masses en masse. I wish I could believe that a dissertation proposal to spend eighteen months reconstructing basic socio-demographic patterns in a single locale over time would still be funded, because such basic research remains urgently lacking for much of the world, and the utter idiosyncrasy of manuscript forms in which potential data is preserved means that no digital magic wand is in the offing.

To be sure, post-colonial scholars developed in the 1990s techniques for reading “against the grain” of official documentation—discerning popular logic and moral economies through the lens of outsiders’ complaints or condemnation. But the systematic underrepresentation of whole strata of people from the print sources that have become the digitized world is not likely to be countered by such techniques, because their absence is accompanied by a fantastic new presence. Possibilities beckon. Intellectual histories can be written about tranches of society (from suffragist women to non-metropolitan jazz hounds) whose ideas were once only painstakingly accessible and impossible to follow across national boundaries. The bit players can finally seize center stage, and it turns out they have so much to say! The optic of the digitizable world captures history made not from the top down but from the bottom of the top and the top of the bottom.30

Social history and Marxian debate wrested our profession away from the conviction that great men made history. The twinned digital and transnational turns push new models,


as a result of source availability and scholars’ excitement more than systematic evaluation of impact. Did middling and mobile men and women make history, albeit never as they pleased? What of people who didn’t cross borders? What of land access and labor process and capital accumulation? One notes with some trepidation the burgeoning genre of “transnational lives”—not because the stories told are not valuable and true, but because an aggregate account of the past in which these stories crowd out others will be as profoundly distorted in its own way as Historia Patria before us. Again, distinguishing the window from the why will be crucial.31

Decontextualization

I noted above that for historians before the digital age, peripheral vision was prohibitively expensive. But ironically, the same source-based constraints that made geographic peripheral vision unduly expensive made topical peripheral vision artificially cheap. Working with tax data or police correspondence or judicial records in a national archive forced you to read through a lot of evidence of political contention and state formation even when what you really wanted to get at was grain prices or prostitution: and vice versa. Analog exploration of written sources—the longtime bread and butter of our craft—built in multidimensional awareness. As a result (again, because if we see it, we think it matters) our disciplinary practice favored multi-causal explanations. There is a structural reason, in other words, for the caricature social scientists throw around: that historians are the scholars who answer any yes-or-no question with “it’s more complicated than that.”

Text-searchability makes possible radically more decontextualized research. This is glaringly true of data-mining projects that track only the frequency of terms or of associations between terms, a fact that drives the skepticism of many regarding the utility of such analyses. But decontextualization is a feature of digitized source use even when

qualitative rather than quantitative methods are employed. Term-fishing offers instant insight. It also deprives you of experiential awareness of just how rare mentions of your term were, of how other issues crowded your topic out in debates of the day. In sum, it erases the kind of sitzefleisch-based test of statistical significance on which our discipline has implicitly relied. Paging through a newspaper in print, in contrast, makes the competing concerns of that place in that moment inescapable, from popular culture to crises of labor, theology, or high politics. Digitized sources do not preclude contextual browsing—quite the contrary. But they make it possible to bypass it, and life is short, and time-to-degree ticking past.

We need a full accounting of the hidden benefits of the unsheddable contextualization that makes work with analog sources so inefficient. For me the dynamic is most vivid in transcribed oral historical sources, like the “Autobiografías campesinas” project that generated hundreds of submissions of life histories in Costa Rica in the late 1970s, now transcribed and archived at the Universidad Nacional in Heredia, and the “Generations of Jamaican Freedwomen and Freedwomen” project, led by Erna Brodber at the University of the West Indies in the same era, which generated hundreds of transcripts preserved at the Sir Arthur Lewis Library of UWI-Mona. Such sources capture, in a varying and not always knowable mix, the priorities of those who organized the project and the priorities of those who recounted their lives. Either has the possibility to be a useful corrective to the researcher’s own convictions and priorities. Analog reading of such sources forces you to read through an awful lot of information about topics that people other than you found important. Is that a bug or a feature? It depends. Certainly it demands a greater investment of time than current doctoral expectations make feasible—

32 For discussion of the separate problems arising from OCR, on the on hand, and the alternating drought and overabundance generated by full-text searches, see Charles Upchurch, “Full-Text Databases and Historical Research: Cautionary Results from a Ten-Year Study,” Journal of Social History 46, no. 1 (2012): 89-105.


especially if there is a perceived need to cover more than one country to produce a dissertation trendy enough to have a chance of landing a job.

But being forced to pay attention to other people’s priorities has been a crucial moral motor within the historical discipline. That’s true, for instance, of our need to point out the gendered presumptions of middle-class reformers, or the race-based dreams of abolitionists, in order to make use of the sources they left about tenement life or slavery’s end. It’s all the more true when using autobiographical or oral sources. To take just one example, scholars working with life stories are routinely confronted with subjects’ insistence on talking about domestic violence even when the questions posed do not include it. Intrafamilial violence—from mothers and fathers as well as husbands and lovers—is routinely undercounted in official tallies, in part due to hard choices made by the targets themselves. But that violence is routinely foregrounded by some of these same targets when, absent threat of outside intervention, they are asked about what shaped their lives.

“I had to beat her so she would leave me.”36 “When de baby born, de baby was dark, and it should be brown… well that was the breaking up of dat. She go away and go to Panama.”37 “I turned 17 on a Saturday and the mistress gave me the beating of the century… I begged the Virgin from my heart to protect me and that I would leave before dawn.”38 “I always thought of my children. How he treated me I wouldn’t treat even an animal but there I was, bearing his insults and jealousy, but one day I couldn’t bear any more anguish and I decided that when he returned from his drinking spree, already 15 days, I would no longer be there.”39

The records of human social life now captured in the digitized world tell us so much about so much that we might forget to remember the systematic absences within them. If our accounts of the movement of people, ideas, and things across borders echo the patterned silences of the sources most available to us, certain conditionants of those flows and certain constraints on mobility and voice will be absent: not because no one thought they mattered, but because no one said so in print.


The consequences of becoming a desk discipline

The geographic emplacement of physical information, sketched at the start of this essay, bears the imprint of past structures of power. But unhinging data from place does not erase global disparities. On the contrary, it may break down inefficiencies that had important equity effects.

There are real reasons one might want to tie data to place. The creation of national archives by nation-states-in-the-making sought to facilitate Historia Patria, and it worked. Oral historical initiatives in multiple sites in the 1970s and beyond reflected a related ideological project: the desire to capture the experiences of those who Historia Patria and its political heirs had pushed to the margins. In many cases, including the two noted above, such material still sits on the shelves of university libraries or archives. Of course the tomes could be digitized and uploaded. Increasingly, they will be. But—even setting aside the questions of who pays and who guarantees subsequent access—their conservators have every right to be wary. Why make it possible for scholars from afar to access even more of the raw materials of scholarly knowledge with even less obligation to do some processing locally, with the externalities local processing implies?

Some of those externalities benefit in-country institutions: it helps to be able to point to physical visitors when defending an archival budget. But many others benefit sojourning researchers and the histories they write, whether they realize it or not. To assess these individual benefits and their collective consequences, it’s necessary to recognize the particular weight in our discipline of scholars based in the global North, who enjoy a disproportionate share of research support and have a disproportionate impact on publication and debate.40 They have no monopoly, however, on historical knowledge.

Things happen in archives and national libraries and on the way to them. This experiential friction, the very thing that made international historical research in a non-digitized world inefficient, tends to teach border-crossing researchers things they need to know whether they know it or not.41 When foreign researchers slog away in archives day—

40 See discussion of the “geopolitics of history at world scale,” especially, the gravitational pull of U.S.-based resources, in Saunier, Going transnational? A similar question is raised, and then side-stepped, by Matthew Pratt Guterl, “Comment: The Futures of Transnational History,” American Historical Review (February 2013): 130, 138-139. Meanwhile, on disparities within Northern academe, in particular the dominant position of English as lingua franca for global/transnational histories, see Bartolomé Yun Casalilla, “‘Localism,’ global history and transnational history: A reflection from the historian of early modern Europe,” Historisk Tidskrift 127, no. 4 (2007): 675.

41 Here my conclusions differ from the insightful discussion by Sean Takats, who sees less substantive intellectual loss in the compression of field time than I do (or maybe is graciously trying not to whine about spending less time in Paris than he used to): “The End of (French) History,” Quintessence of Ham [weblog]
after day next to in-country intellectuals, they can be forced to confront the value of locally produced expertise. I say in-country intellectuals rather than “local scholars,” a term that might suggest their knowledge is somehow limited to the local: definitively not the case. In-country intellectuals may have no publishing profile in the venues that graduate students at Northern institutions routinely survey. But many have extraordinary erudition, only partly captured—given resource constraints—in locally printed booklets, theses, and low-circulation journals. Lucky visiting researchers will find themselves schooled by in-country experts, and should count themselves lucky whether that “schooling” comes in the form of kind instruction or intellectual whupping. Being forced to acknowledge one’s ignorance early and often is the gift offered by academic exchange, whether across borders or within them.

Digitization and uploading makes it increasingly possible to do history as a desk discipline, at least for scholars linked to well-funded institutions in the global North. The last clause is crucial. It would be great if the unanchoring of historical texts from sites of preservation had brought a compensatory shift, in which researchers based in Panama or Paraguay now found themselves technologically empowered to write transnational histories from the South responding to questions deemed urgent in their particular contexts, perhaps topics like “How much profit did that U.S.-based company actually extract from investments here?” or “Which politicians from here went north to drum up support for that invasion?” Technology, of course, is not the problem. Digitized document or newspaper collections belong to someone and don’t come cheap. Scholars in resource-poor institutions in the global North confront this same barrier, and work around it as best they can. Some efforts have been made to address issues of international access.42 But at least for the moment, global disparities in access to sources for international or transnational history are profound.

Digital photography (and to a certain extent, photocopying before it), by making it possible to gather large amounts of data quickly without processing it on site—“processing” being a technical term meaning, in qualitative research, reading and thinking about it—tends in this same direction. When historians research far from home but don’t stay around

42 The British Library’s Endangered Archives program builds attention to such issues in from the start when funding digitization projects where local resources cannot fund urgent preservation: http://eap.bl.uk/. The Digital Library of the Caribbean is another effort to harness the potential of digitization and web access in a project led by and for in-country scholars, promoting international exchange within the Greater Caribbean first and foremost. http://www.dloc.com/

http://quintessenceofham.org/2011/04/08/the-end-of-french-history/#more-584
http://quintessenceofham.org/2011/04/20/time-shifting-and-historical-research/#more-630
long enough to be inconvenienced, insulted, or instructed, the quality of their analysis suffers. Again, the forced contextualization that traditionally made historical research inefficient looks, on reflection, like a significant contributor to the contours of knowledge production in our discipline. One need not be a self-deluding Luddite (“It’s just not the same as back when I had to travel five days by mule train to get to the archive...”) to argue that something is at risk when the broader world becomes simultaneously more present in North-based scholars’ narratives and less present in North-based scholars’ working lives.43

Literary critic Shalini Puri has recently articulated the value of fieldwork in the humanities, underlining the multifaceted impact of presence. Fieldwork not only “invites us to achieve a textured and embodied knowledge of place,” but offers the irreplaceable contribution of “render[ing] the researcher vulnerable to history. When a researcher reads in a library, nobody is reading her back. When one reads in the field, one is constantly being scripted, being made the object of a countergaze, and is thereby forced to confront not only one’s geographical but also one’s historical location.”44 Classic fieldwork has not been the custom of document-based historians. But for those seeking international knowledge, international presence, until now, has been. So even as Puri calls the humanities to recognize the moral and intellectual value of fieldwork, the technological shifts discussed above all suggest that the practice of history, especially global, international, and transnational history, may trend in the opposite direction. Substantial data collection—the collection of enough data to generate publishable findings—is increasingly possible without venturing into “the field” even in the limited sense of visiting a capital city for its archive.

Those cities have stories to tell. They are palimpsests that bear marks of colonial rule, postcolonial dreams, intraregional migration, good governance and its absence, aggressive Free Trade treaties, and more.45 And they don’t only have stories to tell in some metaphorical sense. They are full of people who insist on talking, on asking questions and offering answers of their own. The historian more often finds herself forced to listen to other people’s concerns when queuing for a crowded bus than scrolling from the comfort of home. The cacophony of contemporary reality, especially in the developing world, confronts historical researchers with the real-world stakes of past processes of global connection: a kind of moral peripheral vision parallel to the topical peripheral vision detailed above.

43 I owe Paul Eiss for the mule train.


45 I wrote the sentence thinking of cities in the former colonies, but of course each item holds true, albeit in different ways, for cities that are former metropoles instead.
A grimmer reading of the coinciding digital and transnational turns, then, would conclude that over the course of the past decade it has become much easier for North-based historians to write about places they have never been and may know very little about. At the conference session that inspired this forum, Melanie Newton pointed to the growing number of histories of France, empire, or modernity that include a part of a chapter on the Haitian Revolution, without engaging in any significant way with Haiti’s established historiography and its driving debates, much less with Haiti’s current realities or scholars’ efforts to address them. The pages on Haiti, Newton suggested, often seem more concerned with signaling transnational cachet or non-Eurocentric bona fides than with interest in any collective effort to build understanding of the past in service to the present.46

How common is this kind of gesture, what we might label “drive-by transnationalism,” today? How common will it be tomorrow? Pierre-Yves Saunier suggested almost a decade ago that we should be relieved that “going transnational is not as easy as it sounds.”47 In a world of massified web access to texts of and about the past, it is getting easier every day. Ensuring that the transnational tales thus enabled are full and fair renderings of our interconnected past will require more self-conscious attention to evidentiary paradigms than our discipline has so far embraced. We should recognize, too, that building deep place-based knowledge is no longer the path of least resistance within our discipline—yet it may remain the path to greatest insight, even or especially for those pursuing the transnational angle. The challenge is to take advantage of the optics and methods digitally enabled while remembering those who stand in the shadows, while making time for contextualization, and while ensuring the academic dialogue that counteracts the ignorance of the privileged.

Conclusion

Technology is not destiny. As researchers, advisors, peer reviewers, and panel members we have choices to make. The digital revolution has made finding things out about distant places and the people, goods, and ideas that moved between them cheaper than ever before. The potential is real. But nothing guarantees that the growth of knowledge brought


47 Pierre Yves Saunier, Going transnational?
by fallen barriers, broader vision, and multiscalar research will not be cancelled out by an increase in superficiality or topical narrowness. If transnational gloss is perceived as the currency with which fellowships, article acceptances, and jobs are acquired, we will encourage scholars to invest just enough time and just enough thought in the most far-flung geography plausible. Is this the best use of our hard-pressed humanities dime? Or should we take the digital dividend and invest it in building friction back in: expecting and rewarding engaged fieldwork, deep learning, and international collaboration designed and guided from South as well as North?

The last point is crucial. Ultimately there is nothing inherently equalizing about the conjoined digital and transnational turns. But one can hope that the information-landscape changes that have lowered the barriers to international research by scholars of the global North will increase their interest in the kinds of scholarly connection that can diffuse insights and resources in multiple directions. Virtual communication and exchange could lead to substantive collaboration. There could even be plane tickets. Putting scholars into new places for extended stays—including but not only from the global South into the global North and vice versa—has a great track record of building depth, contextualization, and geopolitical awareness. Digital research that carries us deeper into real-world connection may indeed create the border-crossing wisdom our border-riven world sorely needs.

---

48 I am grateful to Julie Greene for suggesting reasons for optimism.