The Meaning of the Digital Humanities

ALAN LIU

WHAT IS THE MEANING OF THE DIGITAL HUMANITIES TO THE HUMANITIES?

This question of disciplinary meaning—which I ask from the viewpoint of the humanities generally—is larger than the question of disciplinary identity now preoccupying "DH" itself, as insiders call it. Having reached a critical mass of participants, publications, conferences, grant competitions, institutionalization (centers, programs, and advertised jobs), and general visibility, the field is vigorously forming an identity.[1] Recent debates about whether the digital humanities are a "big tent" (Jockers and Worthey), "who's in and who's out?" (Ramsay), whether "you have to know how to code [or be a builder]" (Ramsay, "On Building"), the need for "more hack, less yack" (Cecire, "When Digital Humanities"; Koh), and "who you calling untheoretical?" (Bauer) witness a dialectics of inclusion and exclusion not unlike that of past emergent fields.[2] An ethnographer of the field, indeed, might take a page from Claude Lévi-Strauss and chart the current digital humanities as something like a grid of affiliations and differences between neighboring tribes. Exaggerating the differences somewhat, as when a tribe boasts its uniqueness, we can thus say that the digital humanities—much of which affiliates with older humanities disciplines such as literature, history, classics, and the languages; with the remediation of older media such as books and libraries; and ultimately with the value of the old itself (history, archives, the curatorial mission)—are not the tribe of "new media studies," under the sway of the design, visual, and media arts; Continental theory; cultural criticism; and the avant-garde new.[3] Similarly, despite significant trends toward networked and multimodal work spanning social, visual, aural, and haptic media, much of the digital humanities focuses on documents and texts in a way that distinguishes the field's work from digital research in media studies, communication studies, information studies, and sociology. And the digital humanities are exploring new repertoires of interpretive or expressive "algorithmic criticism" (the "second wave" of the digital humanities proclaimed

ALAN LIU, professor in the English department at the University of California, Santa Barbara, is the author of *Wordsworth: The Sense of History* (Stanford UP, 1989), *The Laws of Cool: Knowledge Work and the Culture of Information* (U of Chicago P, 2004), and *Local Transcendence: Essays on Postmodern Historicism and the Database* (U of Chicago P, 2008). He started *Voice of the Shuttle*, a Web site for humanities research, in 1994. Recent projects he has directed include the University of California Transliteracies Project, on online reading, and *RoSE (Research-Oriented Social Environment)*, a software project. Liu is a coleader of the *4Humanities* advocacy initiative.

PMLA 128.2 (2013), published by the Modern Language Association of America

in "The Digital Humanities Manifesto 2.0" [3]) in a way that makes the field not even its earlier self, "humanities computing," alleged to have had narrower technical and service-oriented aims.[4] Recently, the digital humanities' limited engagement with identity and social-justice issues has also been seen to be a differentiating trait—for example, by the vibrant #transformDH collective, which worries that the digital humanities (unlike some areas of new media studies) are dominantly not concerned with race, gender, alternative sexualities, or disability.[5]

Of course, there are overlaps of people and methods between tribes—to the point that, taking another page from Lévi-Strauss, we would want to mention trickster figures embodying both sides of digital humanities' differences. One such trickster is the field of book history. Prolific in its revisionary studies of matrial texts, ephemera, marginalia, social reading, publishing history, and so on, book history now partly parallels post-McLuhan media studies. Peter Stallybrass argues, for instance, that "the codex and the printed book were the indexical computers that Christianity adopted" as its "technology of discontinuity" (74, 73). Book history therefore overlaps on its new-media-studies side with media archaeology (especially in the German lineage well represented in this context by Cornelia Vismann's *Files: Law and Media Technology*) and on its digital humanities side with approaches to digital texts rooted in revisionary textual editing, bibliography as the sociology of texts, and the materiality of the digital.[6] A similar trickster is science and technology studies (STS), which are curiously underrepresented in both new media studies and the digital humanities even as they are clearly relevant in a way represented by such scholars as N. Katherine Hayles and Tim Lenoir. So, too, historical sociology—which now applies social-network analysis to document corpora—is a trickster splitting the difference between the social sciences and the digital humanities, the latter of which also increasingly use social-network analysis to study plays, novels, literary reception, history, and so on.[7] And of the communities of people who overlap, trickster-like, between the digital humanities and new media studies, the HASTAC collaboratory—which has had enormous success tapping into the energies especially of graduate students ("HASTAC Scholars")—is perhaps the most coyote.

Yet even if we were to complete our hypothetical ethnographer's chart, it would not adequately explain the digital humanities. We would be leaving unexplained the relation of the digital humanities to the humanities generally. My thesis is that an understanding of the digital humanities can only rise to the level of an explanation if we see that the underlying issue is the disciplinary identity not of the digital humanities but of the humanities themselves. For the humanities, the digital humanities exceed (though they include) the functional role of instrument or service, the pioneer role of innovator, the ensemble role of an "additional field," and even such faux-political roles assigned to new fields as challenger, reformer, and (less positively) fifth column. This is because the digital humanities also have a symbolic role. In both their promise and their threat, the digital humanities serve as a shadow play for a future form of the humanities that wishes to include what contemporary society values about the digital without losing its soul to other domains of knowledge work that have gone digital to stake their claim to that society. Or, precisely because the digital humanities are both functional and symbolic, a better metaphor would be something like the register in a computer's central processor unit, where values stored in memory are loaded for rapid shuffling, manipulation, and testing—in this case, to try out new humanistic disciplinary identities evolved for today's broader contention of knowledges and knowledge workers.

The question of the meaning of the digital humanities best opens such an argument to

view because it registers both a specific problem in the digital humanities and the larger crisis of the meaningfulness of today's humanities.

Meaning is clearly a metavalue and also metaproblem for the digital humanities. To unpack this meaning problem, I will spotlight a recent work of digital literary scholarship by two beginning scholars that is state-of-the-art and representative of major trends in the digital humanities—a tactic that has the additional advantage of providing outsiders to the field with an end-to-end look at an example of research by digital humanists. The work is Ryan Heuser and Long Le-Khac's *A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method* (2012), the fourth of the influential digital pamphlets issued by the Stanford Literary Lab.[8] Heuser and Le-Khac report on their innovations in the methods of "distant reading" and text mining that are the signatures of the lab (where they worked with Matthew L. Jockers, Franco Moretti, and others), and they do so with a methodological self-awareness that puts the meaning problem front and center. They reflect near their opening:

> [W]hat is the *meaning* of changes in word usage frequencies? What do we do with such data? With much current research drawing on word frequencies and other quantifiable aspects of culture, these are big questions. We can see now that the greatest challenge of developing digital humanities methods may not be how to cull data from humanistic objects, but how to analyze that data in *meaningfully* interpretable ways. (4; my emphases)

And they add in their concluding section:

> The general methodological problem of the digital humanities can be bluntly stated: How do we get from numbers to *meaning*? The objects being tracked, the evidence collected, the ways they're analyzed—all of these are quantitative. How to move from this kind of evidence and object to qualitative arguments and insights about humanistic subjects—culture, literature, art, etc.—is not clear. (46; my emphasis)

Throughout, Heuser and Le-Khac give *meaning* a gravity that indicates that in the digital humanities the meaning problem has roughly the same weight as the "saving the phenomena" problem in the philosophy of science—the problem, that is, of relating empirically observed phenomena to explainable, theorizable, or predictable ("saved") phenomena. In essence, Heuser and Le-Khac are out to "save the data" by making it meaningful.[9] The form of data they wish to save is quantitative.

Of course, no single example can typify all aspects of the booming field of digital humanities. So before we examine the specifics of the meaning problem in this instance, it will be useful to make four observations that relate Heuser and Le-Khac's research to other work in the digital humanities. The meaning problem may not be as central everywhere, but its frame of analysis is convertible to other frames so that we can see that many parts of the field link up to the same cluster of issues. The following is a kind of conversion table for relating Heuser and Le-Khac's work to other digital humanities areas.

First, we note that Heuser and Le-Khac select their research material from already digitized texts (2,958 British novels from 1785 to 1895, all from *The Internet Archive* except for 250 from Chadwyck-Healey).[10] This means that their work belongs not in the orbit of digitizing, text encoding, publishing, or archiving (activities characteristic of many projects in digital editing, collecting, and curating) but in that of processing and analyzing already built digital repositories (in the manner of projects in text analysis, social-network analysis, visualization, spatial history, etc.). Still, the digital humanities are young enough that these two broad modes of work are not fully specialized. On the one hand, leading text-encoding and digital-

archiving projects find it necessary to create their own analytic, processing, and visualization tools to present materials. And, on the other hand, text-analysis, visualization, and other processing projects often have to go to great lengths to select, clean, and prepare preexistent digital materials as a usable corpus.

Second, while Heuser and Le-Khac formulate their meaning problem in terms of quantification, that is not the only possible formulation. Quantification is indeed a key digital humanities issue, especially when linked to sheer quantity in the wake of recent "big data" funding competitions in the field. But digital humanists also have nonquantitative ways of putting the meaning problem. One is the idea of models (and modeling processes) as richly developed, for example, by Willard McCarty. Models reveal meaning (recognized in patterns, trends, forms) only by reducing the dimensions and features of meaning. Diagrammatic models, especially the visualizations proliferating in the digital humanities (including Heuser and Le-Khac's essay), are a case in point, since large discourse networks (visualized through such tools as *Gephi*) are comprehensible when their scope or detail is kept low but otherwise grow into beautifully mystifying galaxies of nodes and links. Another case in point is the kind of textual model (instantiated in document-type definitions [DTDs] and XML schemas) that enables text encoding. A well-known debate in the digital humanities thus concerns whether the principle of an ordered hierarchy of content objects (OHCO) underlying such models makes texts machine-readable only by disallowing the full range of what Jerome McGann calls human-readable "overlapping" and "recursive structures" ("Position Statement").

Third, even if we concentrate on quantification as the key meaning problem, it makes a difference which disciplinary branch of the digital humanities we are dealing with, since digital history and digital literary studies, to take two major branches, arrive at quantification through different routes and at different stages. Digital history emerged in a discipline that had already forcefully experienced numerical method—for example, the quantitative side of the *Annales* school and cliometrics. Perhaps as a result, digital history is farther along than most digital humanities branches in molding quantitative work to a related heritage of *Annales* method: spatial and cartographic history. The state of the art in digital history, as it were, is Fernand Braudel plus satellite mapping (Seefeldt and Thomas; Owens).

Fourth, Heuser and Le-Khac's research required a combination of skills in programming and interpretation, thus addressing the "do you have to be a builder?" question that has recently bedeviled the digital humanities by answering, in effect, "you have to be both a builder and an interpreter." Less interesting than this question itself, which is based on an increasingly obsolete notion of solo work in which one is either a builder or interpreter, is the way Heuser and Le-Khac are both builders and interpreters: through rich collaboration. Here we reach the outer limits of the frame of the meaning problem, where it converts into a coextensive frame. Just as meaning is both a metavalue and a metaproblem, so is collaboration as it bears on such urgent issues in the digital humanities as coauthorship, collective project building, multigraph books, open peer review, social media, crowdsourcing, and the hiring and promotion implications of all these. Rather than explore the collaboration problem in its own frame here, I note only that it is fundamentally convertible to the meaning problem. For example, the question of what kind of knowledge is produced by "the wisdom of the crowd," "collective intelligence," "the long tail," "the hive mind," "folksonomy," and so on (dominant memes of Web 2.0) is essentially a question about the meaning of the social version of big data, the big crowd. The mind, or mindlessness, of that crowd has been a core problem of modernity since at least the French Revolution.

With these observations as a guide for relating Heuser and Le-Khac's work to other emphases in the digital humanities, we can now look closely at their pamphlet. The meaning problem comes to the fore there as a consequence of Heuser and Le-Khac's key research innovation. They report on their use of a tool they created called *Correlator*, which, when fed "seed words" suggested in part by "existing literary scholarship" (11), processed their long-nineteenth-century corpus of novels (with the aid of a database designed by Jockers that tabulated the number of occurrences of each word [Heuser and Le-Khac 6]) to find other words that were statistically correlated with the seed words and whose frequency trends, measured longitudinally across the century in decade intervals, closely followed the frequency trends of the seed words. In other words, *Correlator* finds what Heuser and Le-Khac call "word cohorts" in the corpus, consisting of words that kept company with one another and behaved similarly over time, waxing or waning in frequency together and standing out from other companies of words. The final result—which I jump to while skipping the details of Heuser and Le-Khac's algorithmic method as well as, temporarily, an important adjustment step in their method—is the identification in the word cohorts of "rich, consistent semantic fields" that are "both semantically and culturally legible" in historical trends. An example is the 136 words such as *gentle*, *sensible*, *vanity*, *elegant*, *delicacy*, *reserve*, *mild*, and *restraint* that they label the "social restraint field" in the novels (8).

Using this method, Heuser and Le-Khac made two principal discoveries. First, they used the seed words *integrity*, *modesty*, *sensibility*, and *reason* (suited to scholarship on novels of the period) to find a strikingly large cohort of "abstract, socially normative, evaluative, and highly polarized words" whose frequencies declined dramatically over the century. This cohort they subcategorized into semantic fields labeled "social restraint," "moral valuation," "sentiment," and "partiality" (11–19). Then, serendipitously (instead of under the sway of preexisting scholarship), they fed the unlikely seed word *hard* into *Correlator* to discover a strikingly distinct, very large cohort of "concrete description words of a direct, everyday kind" whose frequencies rose dramatically over time. This cohort they subcategorized into semantic fields for action verbs, body parts, colors, numbers, locational and directional adjectives and prepositions, and physical adjectives (19–27). The strongly inverse correlation between these two large and differentiated word cohorts, which they further statistically corroborated (28–29), started them on the path of interpretation. The "abstract" cohort, they concluded, consists of words whose usage, while unanchored in specifics, was "monitored and tightly constrained" by the traditional smaller, rural communities represented more or less earlier in the century in evangelical, gothic, and village novels and the novels of Jane Austen, Walter Scott, and George Eliot (Eliot is a chronological outlier), while the "hard" cohort is populated by words whose stand-alone referentiality and alienation from larger contexts correlated later in the century with the "wider, less constrained social spaces" of the urban centers represented in city, industrial, adventure, fantasy, science fiction, and children's novels (30–34).

This insight finally led Heuser and Le-Khac at the highest level of interpretation to follow Raymond Williams's *Culture and Society* and *The Country and the City* in suggesting that the inverse "abstract" and "hard" trends reveal something significant about the "social space of the novel." They argue that the "values of conduct and social norms" in "knowable communities" (a phrase from Williams) declined in the face of "urbanization, industrialization, and new stages of capitalism" (33, 35–36). The computational verification of this previously known thesis, coupled with the discovery of precise word cohorts

giving genuinely fresh insight into the thesis, enables them in later sections of their pamphlet to offer more recognizably normative literary and cultural criticism, touching on action, setting, and character (37–45). Here they closely read texts (including, for the first time, block quotations), match aggregate trends to "units understandable and familiar to us as readers and literary scholars[:] the actual novels, genres, and authors" (31), and generalize about sweeping changes in cultural history, but with the important distinction that their reading is based not on the usual anecdotal, faux-empirical, or unique-case observations of literary criticism (e.g., noticing that a word appears "often" or in an "important" location) but on lines of interpretation generated by machine observation.

Here we reach the crux of the meaning problem in the digital humanities. While Heuser and Le-Khac mix "supervised and unsupervised procedures" (28), the newest, boldest, and most interesting part of their methodology is unsupervised. They help advance an important, general digital humanities goal that might be called tabula rasa interpretation—the initiation of interpretation through the hypothesis-free discovery of phenomena. Also shared by such digital humanities methods as topic modeling (a mathematics-based way to discern differentiated clusters of words that Heuser and Le-Khac use to corroborate their findings), the ideal in its purest form is what Heuser and Le-Khac call "an unsupervised method that generates topics without subjective input from users" (28).[11] That is, a computer should be able to read texts algorithmically and discover word cohorts or clusters leading to themes without acting on an initial concept from an interpreter looking to confirm a particular theme. Of course, Heuser and Le-Khac assume that there *are* preexisting themes to be found in the word cohorts of primary materials and also that the main mission is to discover them. But tabula rasa interpretation is equally a goal of the more postmodern side of the digital humanities, which argues that critics should use algorithmic methods to play with texts experimentally, generatively, or "deformatively" to discover alternative ways of meaning that are not so much true to preexisting signals as riffs on those signals. The common goal is to banish, or at least crucially delay, human ideation at the formative onset of interpretation.

However, tabula rasa interpretation puts in question Heuser and Le-Khac's ultimate goal, which is to get from numbers to humanistic meaning ("qualitative arguments and insights about humanistic subjects—culture, literature, art, etc."). It is not clear epistemologically, cognitively, or socially how human beings can take a signal discovered by machine and develop an interpretation leading to a humanly understandable concept unless that signal (in order to be recognized as a signal at all) contains a coeval conceptual origin that is knowable in principle because, at a minimum, the human interpreter *has* known its form or position (the slot or approximate locus in the semantic system where its meaning, or at least its membership in the system, is expected to come clear).[12] If the machine can discover word cohorts triggered by seed words, in other words, then what seed concepts—which is to say seed semantics (using "semantics," for the moment, as overlapping with "concepts" in a manner consistent with Heuser and Le-Khac's usage)—lurk in the background as a latent, classificatory form of relational semantic positions able to make word cohorts into "proto-semantic fields" (7)?

Thus the immense importance of the adjustment step in Heuser and Le-Khac's method that I earlier elided. In fact, Heuser and Le-Khac used *Correlator* by itself to produce only initial word cohorts and not finished semantic fields because they realized that they needed to ensure that their cohorts had a semantic consistency that quantitative correlation alone could not offer. Some word cohorts discovered

by machine, for example, suggested only fuzzy semantic fields that seemed blurred by extraneous words or, inversely, to lack words that should have been there. In other words, word cohorts had to be filtered and filled out in ways that made sense. Heuser and Le-Khac thus realized that they needed not just seed words but seed semantic concepts suspended precisely in what I above called a classificatory form of relational semantic positions, endowing cohorts with a sense of what is and is not proximate in meaning.

In short, Heuser and Le-Khac needed a thesaurus, and not just any thesaurus but one offering a historical semantics matched to the longitudinal dimension of their word cohorts. After they had already begun using *Correlator*, they pulled a rabbit out of the hat. They turned to the remarkable *Historical Thesaurus of the Oxford English Dictionary* (2009; hereafter *HTOED*), which had just been published, and borrowed its historical semantic classifications through what they call "a dialogic method that drew on both quantitative historical data and qualitative semantic rubrics to construct semantic fields with precision and nuance":

> Having moved through an empirically and historically focused stage of semantic field development, we needed to return to the semantic focus in order to make such purely empirical word cohorts interpretable and meaningful. Our initial approach was to filter through these words for groups that seemed semantically coherent, but this proved too loose and subjective. . . . Finally we turned to the [HT]OED. . . . It's nearly exhaustive, its categories are nuanced and specific, and it's truly organized around meaning. We used this powerful taxonomy to do two things: first, be more specific in identifying the semantic categories that constituted our word cohorts; second, to expand these word cohorts with many more words. (7)[13]

Created at the University of Glasgow beginning in 1965, the *HTOED* taxonomizes the English language into three master semantic concepts, "the external world," "the mind," and "society," and then, descending its classificatory tree by stages, into myriad ramifications. On the lower branches of the taxonomy, synonyms appear in the chronological order in which they entered the language.[14] At first glance, therefore, the *HTOED* is the perfect concept-hunting guide for Heuser and Le-Khac's word-cohort-hunting machine. This seems even more apparent when we realize just how human-powered the semantic interpretation involved in making the *HTOED* was. As documented in reports by Christian Kay and Irené Wotherspoon, among its chief editors, and in a detailed e-mail to me from the current *HTOED* associate director, Marc Alexander, the editors' sorting of words by meaning and chronology recapitulated the famous use of paper slips in the compilation of the *OED* itself.[15] Started before humanities computing was practical, the *HTOED* required human beings over decades to write down individual words from the *OED* on paper slips with meanings, usage dates, and sparse metadata, then to sort, bundle, and file the slips in conceptual groupings and hierarchies. When computation entered the picture, it did so originally in a secondary capacity (to drive the print run of the work). In its formative state, the *HTOED* was a human labor of semantic ordering.

By installing the *HTOED* as what amounts to a plug-in for *Correlator*, Heuser and Le-Khac sowed their hermeneutical process with a coseed of human semantic interpretation. They thus "solved" the meaning problem only by deftly turning the aporia between tabula rasa quantitative interpretation and humanly meaningful qualitative interpretation into its own apparent solution: a "dialogic approach that oscillates between the historical and the semantic, between empirical word frequencies that reveal the historical trends of words and semantic taxonomies that help us identify the meaning and content of those trends" (9). They add, "Strictly

speaking, the methods developed here are not looking at word cohorts, which have historical consistency but may lack semantic coherence, or semantic fields, which have semantic coherence but may have an ahistorical relationship. The real object of study is a hybrid one that satisfies both requirements" (9–10). Such hybridity is a prevalent feature of digital humanities method.[16]

How might such hybrid method be better grounded theoretically? I return to my earlier differentiation of the digital humanities from neighboring fields. Two of the deficiencies in the digital humanities revealed by that ethnographic map of fields are relevant.

One is design theory and practice, which Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp in their recent *Digital_Humanities* declare with manifesto-like boldness to be central to the field (esp. 12–16, 117–19). Currently, it seems to me, the union between the digital humanities and new-media-studies communities that would be needed for full realization of this vision is more a goal than a reality, existing in specific projects and not programmatically. But the book's coauthors, expressing a West Coast view of the digital humanities rooted in media and design arts, are right to aim for design as a principle of knowledge discovery and generation rather than (more typical in the digital humanities) as an after-the-fact rendering of data in scatter plots, social-network graphs, and other stale visualizations or, equally tired, book-like or blog-like publication interfaces. As the coauthors put it, when "used to pose and frame questions about knowledge," design is "an intellectual method," an "embodiment of a project's argument and methodology," "an act of thinking," and a "new foundation for the conceptualization and production of knowledge" (13, 14, 15, 117). Interactive, multimodal, dynamic, and participatory design in the digital age is a method not just of pattern recognition but of pattern *understanding*. Seeing design in data is a method for knowing meaning in the digital humanities.

The other deficiency I refer to is science and technology studies (STS), which digital humanists often occlude even as they speak of "digital technology," "media technology," and so on, as if technology were an indivisible part of the digital and media without its own history, philosophy, sociology, politics, economy, and aesthetics all tangled up with, yet also distinct from, science. I invoke especially the postmodern branch of STS (e.g., Feyerabend; Latour; Pickering), whose "against method" view of science (especially in its weird relations with technology) is that any quest for stable method in understanding how knowledge is generated by human beings using machines founders on the initial fallacy that there are immaculately separate human and machinic orders, each with an ontological, epistemological, and pragmatic purity that allows it to be brought into a knowable methodological relation with the other—whether a relation of master and slave, cause and effect, agent and instrument, or another. What could we learn from STS if we took the Stanford Literary Lab and other digital humanities centers and programs at their word and studied them as labs, much as Andrew Pickering studied the "hunting of the quark" in a physics lab (68–112)? The answer is likely that digital humanities method—converging with, but also sometimes diverging from, scientific method—consists in repeatedly coadjusting human concepts and machine technologies until (as in Pickering's thesis about "the mangle of practice") the two stabilize each other in temporary postures of truth that neither by itself could sustain. Knowledge is an ice-skater's dance on a slippery epistemic surface, on which neither the human nor the machine—the dancer nor the skates—alone can stand. STS, in other words, is another method for knowing meaning in the digital humanities. In fact, it can be thought of as complementing the method

of design. The difference between a mangle of practice and a dance, after all, is design.

Indeed, an STS approach opens a fascinating chapter in our reading of Heuser and Le-Khac's work that I can only briefly relate.[17] While I have not conducted STS-style ethnographic and documentary research on the Stanford Literary Lab (except through informal observations during a daylong visit at the invitation of the lab),[18] I have looked from a distance into the "lab" of the making of the *HTOED*—that is, the human-semantics lab that Heuser and Le-Khac position as the partner to their computational-analysis lab. When we view the *HTOED* in this way, we realize that Heuser and Le-Khac pulled not just any rabbit out of a hat but a special rabbit much like the one that led Alice down the rabbit hole. The *HTOED* is less a solution to the meaning problem than a recursive, looking-glass version of the very same problem. At least five aspects of its making are relevant in this regard:

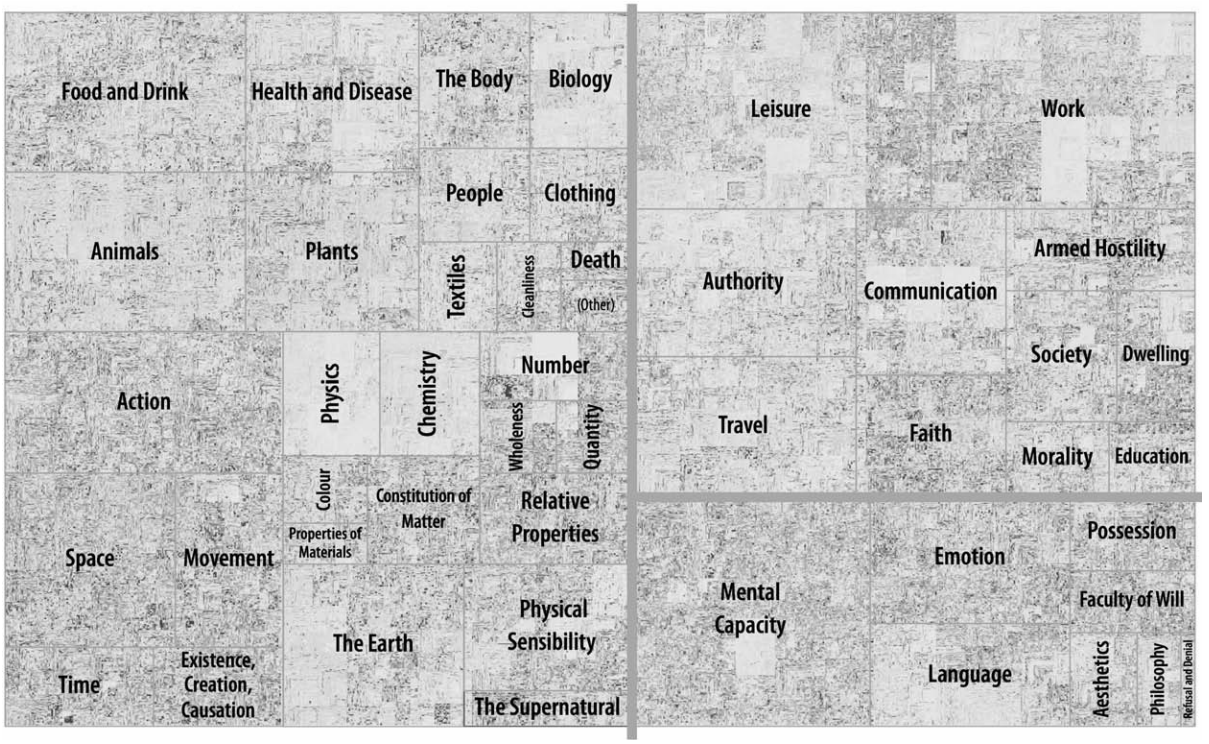
1. While the creation of the *HTOED* was essentially precomputational, it was not pretechnological. Requiring human beings to write words, meanings, and metadata on paper slips and then to sort the slips in drawers (and later in databases), the *HTOED* originated as a thoroughly entangled human-technological, rather than simply human, semantic act.
2. Following Heuser and Le-Khac, I have so far treated "semantics" and "concepts" as coincident. But in reflecting on the *HTOED*'s method, Kay separates out lexical semantics (meaning relations among words) and conceptual semantics (meanings linked to external-world referents). A focus on the former, she says, was "essentially the procedure adopted in *HTOED*, and is what we mean by saying that the classification should 'emerge' from the data," meaning the purely lexical data in their source text, the *OED* ("Classification" 265–66).
3. The phrase "'emerge' from the data" is the cue for an important tenet of the *HTOED*. As Kay argues, "Our theoretical position on *HTOED* has always been that the classification at whatever level should develop from the data rather than be imposed upon it using some predetermined schema" ("Classification" 258). The underlying authority of the *HTOED*, it turns out, is "data" (the record of the English language observed through the instrument of the *OED*). That is, the *HTOED* itself adhered to the principle of the tabula rasa discovery of phenomena and initiation of interpretation (in this case, taxonomic interpretation).
4. It was thus after surveying their lexical data that the *HTOED* editors contravened the ordering scheme of their canonical predecessor, *Roget's Thesaurus*, which had put "abstract relations" first. The *HTOED* puts the "external world" first in its trinary taxonomic structure. That's because concrete and near-to-hand word senses relating to the external world entered the language earlier (Kay, "Classification" 257–62). The implications for our reading of Heuser and Le-Khac are startling. The long historical trend identified by the *HTOED*, in which concrete words precede abstract ones, is the reverse of the long-nineteenth-century novelistic trend that Heuser and Le-Khac identify, in which "abstract" words dominate earlier and "hard" words later. Of course, this by no means contradicts their thesis, since in more fully stated form their argument may well be that, because of urbanization, the nineteenth century (and the novel form) was exceptional, or at least matched the trend of only a few other centuries (and forms) in history. But a commitment to reading nineteenth-century novels in conjunction with the *HTOED*'s larger corpus would require further testing to see how exceptional the nineteenth-century trend really was and thus whether additional sociocultural or other phenomena must be factored in to explain its specificity.[19]
5. Finally, as a grace note, I add that while computation was an afterthought in making the *HTOED*, today it is crucial for advanced uses of the work. The *HTOED*

editors eventually migrated their content into a succession of relational databases because they realized that computation might support advanced, real-time querying (Wotherspoon, "Historical Thesaurus" 218). Even more dramatically, a recent essay by Alexander shows that querying the *HTOED* for insight into the history of the language can itself be a form of the digital humanities ("'Various Forms'"). Alexander conducts quantitative computational analysis of the *HTOED* to generate visualizations leading to hypotheses about language changes (see the figure below). Connecting the STS and design approaches I outlined earlier, we might even say that at this point using the *HTOED* becomes an experiment in digital design. In the end, the *HTOED* is not the "other" that Heuser and Le-Khac need to help make the work of their *Correlator* meaningful; it is the precursor of *Correlator*.

But going down a rabbit hole, while necessary in pursuing any case of the digital humanities to its methodological foundation, is not how we must conclude. I thus climb out of the specific purview of my example, and even of the digital humanities field, to open my argument to its most general extent. Apropos is the following insight from N. Katherine Hayles:

> The further one goes along the spectrum that ends with "machine reading," the more one implicitly accepts the belief that large-scale multicausal events are caused by confluences that include a multitude of forces interacting simultaneously, many of which are nonhuman. . . . If events occur at a magnitude far exceeding individual actors and far surpassing the ability of humans to absorb the relevant information, however, "machine reading" might be a first pass toward making visible patterns that human reading could then interpret. (29)

It is not accidental, I can now reveal, that at the beginning of this essay I alluded to Lévi-Strauss and structural anthropology. Structuralism is a midpoint on the long modern path toward understanding the world as system



Marc Alexander's tree-map visualization of present-day English in the *HTOED*. Each dot represents a word, and the shade of the dot corresponds to when the word entered the language (darker dots show earlier words). Words are arranged by semantic proximity as indicated in the labels.

(e.g., as modes of production; Weberian bureaucracy; Saussurean language; mass, media, and corporate society; neoliberalism; and so on) that has forced the progressive side of the humanities to split off from earlier humanities of the human spirit (*Geist*) and human self to adopt a worldview in which, as Hayles says, "large-scale multicausal events are caused by confluences that include a multitude of forces . . . many of which are nonhuman." This is the backdrop against which we can see how the meaning problem in the digital humanities registers today's general crisis of the meaningfulness of the humanities. The general crisis is that humanistic meaning, with its residual yearnings for spirit, humanity, and self—or, as we now say, identity and subjectivity—must compete in the world system with social, economic, science-engineering, workplace, and popular-culture knowledges that do not necessarily value meaning or, even more threatening, value meaning but frame it systemically in ways that alienate or co-opt humanistic meaning. Humanistic knowledge today is thus increasingly assimilated to what humanists themselves call research, evidence, analysis, method, productivity, and "impact" (as this term is institutionalized in "research assessment exercises" in British universities), with no unfilled time and space left for any old ghosts in the machine—unless, as I have argued in *Laws of Cool*, there remains a yearning, nowhere keener than among our students, caught in the educational mangle, to be cool. Cool people say nix to today's knowledge-work system even as they walk into the cubicles.

Of course, if this were only a problem of research methodology, then I would be extravagant to call it a "crisis" in the meaningfulness of the humanities. But "crisis" is appropriate when we realize that the meaning problem also affects pedagogy and jobs in the wake of economic recession, which brings the problem cruelly to bear on individual humanists in training or seeking jobs (not to mention on the humanities programs that nurture and employ them). Here opens a set of topics that I cannot deal with in present limits but that more fully demonstrate how the digital humanities register the larger issues of the humanities.

One topic is the way digital pedagogy—as witnessed in current controversies over massive open online courses (MOOCs), all-digital "campuses" in public universities, and so on—registers the possibility of gigantic changes in the aims, practices, audiences, infrastructure, and staff of humanities teaching. The bluff is now called on decades of defensively legitimating the humanities as thinking and language "skills" added on top of traditionally meaningful humanistic knowledge. If the humanities fit that mold in part, then maybe—some administrators and legislators think—they should fit it entirely so that their content can be "delivered" modularly through the Internet in the manner of the MOOCs or Khan Academy courses in science, technology, engineering, and mathematics that have made the biggest public impression.

Another topic is the job market for digital humanists, especially in regard to tenurable faculty lines versus "alt-ac" (alternative academic career) adjunct, staff, support, and cultural-institution positions. The turmoil and uncertainty in the nature of digital humanities jobs register the larger uncertainty of employment in the humanities as "meaningful" jobs transition away from tenure and toward a corporatized ideal of reconfigurable and replaceable professional-managerial knowledge workers perpetually threatened with restructuring layoffs in favor of even more exploited "permatemp" and outsourced labor.

In all these ways, the digital humanities register the crisis of the humanities. For that reason, I and others started the *4Humanities* advocacy initiative, "powered by the digital humanities community," so that the digital humanities can try to advocate for the humanities and not just register their crisis. I do not know how much difference that initiative and others like it will make in the

meaningfulness of the humanities to the world. But I do know that such an effort—dedicating the digital humanities to the soul of the humanities—is what is meaningful for a humanist, digital or otherwise, now.[20]

Notes

1. Among recent surveys and overviews of the digital humanities are Svensson's articles; Hayles; Kirschenbaum, "Digital Humanities" and "What Is Digital Humanities"; Liu, "State" and "Where Is Cultural Criticism"; Burdick, Drucker, Lunenfeld, Presner, and Schnapp; and "Digital Humanities Manifesto 2.0." Recent or forthcoming essay collections about the digital humanities include Schreibman, Siemens, and Unsworth; Gold; Price and Siemens; and Goldberg and Svensson. On whether the digital humanities are a field and on the linguistic usage of the phrase *digital humanities*, see Liu, "Is Digital Humanities a Field?"

2. These expressions in recent conferences, blogs, tweets, discussion threads, papers, etc., have become important "memes" shaping the digital humanities community. On the issues of theory, building, and hack versus yack, see also Cecire, "Introduction."

3. Since the question of who is included in the digital humanities can be a sensitive one, I should clarify my understanding of the scope of the field. While I focus in this essay on digital literary studies and one concrete research example, the digital humanities are much broader. I thus take care to relate my example to shared methods and problems across the field (including, e.g., digital history). However, while my goal is to address a fairly broad notion of the digital humanities as they are commonly practiced and discussed, I do not try to make the tent so big—by, for instance, covering what should or could be part of it but so far is not commonly recognized to be so (e.g., the design field)—that the term *digital humanities* becomes formless or aspirational.

Also, I do not specifically discuss digital work in such near humanities as the branches of the social sciences, archaeology, and anthropology that have joined the humanities in the so-called linguistic and cultural turns and related trends. Whether such work is considered digital humanities depends on the prior issue of whether, digital methods aside, it is humanistic (fitting, e.g., the loose definition in the 1965 National Foundation on the Arts and Humanities Act: "those aspects of social sciences which have humanistic content and employ humanistic methods" [National Endowment 1]).

Caveats also apply to my generalizations about new media studies and to the gross simplification here (given nuance below) by which I separate them from the digital humanities.

For helping me think about these inclusion issues, I am grateful to members of the audiences at talks where I delivered versions of this essay who pointed out that they do not entirely recognize their field or work in my description of the digital humanities. Ultimately, of course, the "who is in the digital humanities?" issue will be adjusted on the ground through normal professional processes of adjudication—where one gets a job or places one's students, where one publishes in print or online, which forums or blogs one posts on, which *Twitter* hashtags one is associated with, which conferences one goes to, which grants one gets, etc.

4. On algorithmic criticism, see Ramsay's essay by that title.

5. On #transformDH, see Phillips. The collective now appears through its hashtag on *Twitter*, has a *Tumblr* page, and is a HASTAC group.

6. Examples of the textual-editing, bibliography-as-sociology-of-texts, and materiality-of-the-digital approaches in the digital humanities include, respectively, McGann, *Radiant Textuality*; the Text Encoding Initiative's *TEI: P5 Guidelines*, with its attention to prosopography and social relations ("Names, Dates, People, and Places"); and Kirschenbaum, *Mechanisms*.

7. Examples of the historical sociology I refer to include Franzosi; Bearman and Stovel; and Mohr and Duquenne. An example of social-network analysis in digital literary studies is Moretti, *Network Theory*. (On such analysis in general, see my "From Reading.") An example of social-network analysis in digital history is Lemercier and Rosental. (For other historians using the methods and tools of social-network analysis, see "Bibliography.")

8. See also Heuser and Le-Khac's "Learning to Read Data," which summarizes the research reported in their pamphlet.

9. For a discussion of "saving the phenomena" that bears centrally on the issue of data, see Bogen and Woodward.

10. The full list of the novels in Heuser and Le-Khac's corpus can be found in their "Online Companion." Information on the source of the digitized texts is from an e-mail from Heuser to the author.

11. Goldstone and Underwood offer an explanation and example of topic modeling of special interest to readers of *PMLA*. Separately, Underwood more fully explains the methodology. For a succinct discussion of the difference between supervised and unsupervised data mining—a distinction originating in the field of machine learning—see "Analytics."

12. Sculley and Pasanek consider the problem of "circularity" (and other issues) in humanities data mining.

13. For a fuller description of the way Heuser and Le-Khac used the *HTOED*, see appendix C in their *Quantitative Literary History*.

14. For this essay, I have consulted the online *HTOED*, which now appears as part of the online *OED*. On the

HTOED's classification system, see "Structure of the *Historical Thesaurus*." For additional general information about the work, see Kay, "What Is the *Historical Thesaurus*."

15. For information about the making of the *HTOED* referred to below, I have consulted Kay, "Classification" and "What Is the *Historical Thesaurus*"; Kay and Chase; and Wotherspoon, "Historical Thesaurus" and "Making"; as well as Alexander's e-mail. For photos of the paper slips, filing drawers, and computers used, see Historical Thesaurus *Photo Gallery*. My special thanks to Alexander for helping me gather resources and for his e-mail filled with details and reflections on the manual and technological processes used in making the *HTOED*.

16. See, e.g., Gibbs and Cohen's discussion of their hybrid method (70, 76).

17. I abbreviate here a longer discussion of the *HTOED* to be included in a version of this essay for my book in progress on the digital humanities. My thanks to Clare Birchall, who, in a conversation with me after I presented an early version of this essay at King's College, asked a question that made me start looking into the technological dimensions of the *HTOED*.

18. My thanks to the Stanford Literary Lab for inviting me to visit on 21 May 2012. On the genesis and principles of the lab, see Jockers.

19. In their statistical analysis of diction in English literary works over a longer period (the eighteenth and nineteenth centuries), Underwood and Sellers make a related point when comparing their findings with Heuser and Le-Khac's. They track a rise in the nineteenth century in the proportional incidence of earlier-vintage, often Anglo-Saxon words in literary genres (including prose fiction) that is "largely consubstantial" with the rising incidence of "hard" words in novels of that century found by Heuser and Le-Khac. However, they add that their longer historical baseline shows that the trajectory of such words "had recently reversed direction" as part of a new way of being "literary." "The relative scarcity of simple action verbs in early-nineteenth-century writing, for instance," they note, "was a recent development" in literary language that reveals social transformations only through the mediation of "competing ideals of literary refinement." My thanks to Le-Khac for his e-mail calling my attention to Underwood and Sellers's article, which I had not seen.

20. Advocating for the humanities does not necessarily only mean defending older or disciplinary notions of the humanities. My "Humanities and Tomorrow's Discoveries" attempts to reframe the humanities in common cause with other disciplines and with public needs.

Works Cited

Adams, Michael, ed. *"Cunning Passages, Contrived Corridors": Unexpected Essays in the History of Lexicography*. Electronic ed. Monza: Polimetrica, 2010. PDF file.

Alexander, Marc. Message to the author. 6 Oct. 2012. E-mail.

———. "'The Various Forms of Civilization Arranged in Chronological Strata': Manipulating the *HTOED*." Adams 309–23.

"Analytics: Supervised vs. Unsupervised Learning." *MONK: Metadata Offer New Knowledge*. N.p., n.d. Web. 1 Dec. 2012.

Bauer, Jean. "Who You Calling Untheoretical?" *Journal of Digital Humanities* 1.1 (2011): 68–74. Web. 23 Nov. 2012.

Bearman, Peter S., and Katherine Stovel. "Becoming a Nazi: A Model for Narrative Networks." *Poetics* 27 (2000): 69–90. Print.

"Bibliography." Ed. Marten Düring. *Historical Network Research*. N.p., 18 Sept. 2012. Web. 26 Nov. 2012.

Bogen, James, and James Woodward. "Saving the Phenomena." *Philosophical Review* 97.3 (1988): 303–52. Print.

Burdick, Anne, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp. *Digital_ Humanities*. Cambridge: MIT P, 2012. Print.

Cecire, Natalia. "Introduction: Theory and the Virtues of Digital Humanities." *Journal of Digital Humanities* 1.1 (2011): 44–53. Web. 23 Nov. 2012.

———. "When Digital Humanities Was in Vogue." *Journal of Digital Humanities* 1.1 (2011): 54–58. Web. 23 Nov. 2012.

"The Digital Humanities Manifesto 2.0." *Humanities Blast*. N.p., 2009. Web. 9 Sept. 2012.

Feyerabend, Paul. *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: Verso, 1978. Print.

4Humanities: Advocating for the Humanities. Ed. Christine Henseler, Alan Liu, Geoffrey Rockwell, Stéfan Sinclair, and Melissa Terras, with Lindsay Thomas. 4Humanities, n.d. Web. 16 Jan. 2013.

Franzosi, Roberto. *Quantitative Narrative Analysis*. Los Angeles: Sage, 2010. Print.

Gibbs, Frederick W., and Daniel J. Cohen. "A Conversation with Data: Prospecting Victorian Words and Ideas." *Victorian Studies* 54.1 (2011): 69–77. Print.

Gold, Matthew K., ed. *Debates in the Digital Humanities*. Minneapolis: U of Minnesota P, 2012. Print.

Goldberg, David Theo, and Patrik Svensson, ed. *Humanities and the Digital*. Cambridge: MIT P, forthcoming.

Goldstone, Andrew, and Ted Underwood. "What Can Topic Models of *PMLA* Teach Us about the History of Literary Scholarship?" *The Stone and the Shell: Historical Questions Raised by a Quantitative Approach to Language*. N.p., 14 Dec. 2012. Web. 9 Dec. 2012.

Hayles, N. Katherine. *How We Think: Digital Media and Contemporary Technogenesis*. Chicago: U of Chicago P, 2012. Print.

Heuser, Ryan. Message to the author. 13 Sept. 2012. E-mail.

Heuser, Ryan, and Long Le-Khac. "Learning to Read Data: Bringing Out the Humanistic in the Digital Humanities." *Victorian Novels* 54.1 (2011): 79–86. Print.

———. "Online Companion to 'Learning to Read Data: Bringing Out the Humanistic in the Digital Humanities,' *Victorian Studies* 54.1; and Stanford Literary Lab Pamphlet 4." *Stanford Literary Lab*. Stanford Lit. Lab, May 2012. Web. 13 Sept. 2012.

———. *A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method*. *Stanford Literary Lab*. Stanford Lit. Lab, May 2012. Web. 1 Sept. 2012. Pamphlet 4.

Historical Thesaurus. Oxford English Dictionary. Oxford UP, 2012. Web. 25 Nov. 2012.

Historical Thesaurus of the Oxford English Dictionary, with Additional Material from "A Thesaurus of Old English." Ed. Christian Kay, Jane Roberts, Michael Samuels, and Irené Wotherspoon. Oxford: Oxford UP, 2009. Print.

Historical Thesaurus *Photo Gallery*. U of Glasgow, n. d. Web. 25 Sept. 2012.

Jockers, Matthew L. "Thoughts on a Literary Lab." *Matthew L. Jockers*. N.p., 4 Jan. 2013. Web. 10 Jan. 2013.

Jockers, Matthew, and Glen Worthey. "Introduction: Welcome to the Big Tent." *Digital Humanities 2011*. N.p., 2011. Web. 23 Nov. 2012.

Kay, Christian. "Classification: Principles and Practice." Adams 255–70.

———. "What Is the *Historical Thesaurus* of the *OED*." *Oxford English Dictionary*. Oxford UP, 2012. Web. 25 Nov. 2012.

Kay, Christian, and T[homas] J. P. Chase. "Constructing a Thesaurus Database." *Literary and Linguistic Computing* 2.3 (1987): 161–63. Print.

Kirschenbaum, Matthew G. "Digital Humanities As/Is a Tactical Term." Gold 415–28.

———. *Mechanisms: New Media and the Forensic Imagination*. Cambridge: MIT P, 2008. Print.

———. "What Is Digital Humanities and What's It Doing in English Departments?" *ADE Bulletin* 150 (2010): 55–61. Print.

Koh, Adeline. "More Hack, Less Yack? Modularity, Theory and Habitus in the Digital Humanities." *Adeline Koh*. N.p., 21 May 2012. Web. 23 Nov. 2012.

Latour, Bruno. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford UP, 2005. Print.

Le-Khac, Long. Message to the author. 4 Feb. 2013. E-mail.

Lemercier, Claire, and Paul-André Rosental. "The Structure and Dynamics of Migration Patterns in Nineteenth-Century Northern France." Vers. 3. *HAL: Sciences de l'homme et de la société*. Centre pour la Communication Scientifique Directe, 5 May 2010. Web. 24 Nov. 2012.

Liu, Alan. "From Reading to Social Computing." Price and Siemens.

———. "The Humanities and Tomorrow's Discoveries." *4Humanities: Advocating for the Humanities*. 4Humanities, 25 July 2012. Web. 14 Mar. 2013.

———. "Is Digital Humanities a Field? An Answer from the Point of View of Language." *Alan Liu*. N.p., 6 Mar. 2013. Web. 6 Mar. 2013.

———. *The Laws of Cool: Knowledge Work and the Culture of Information*. Chicago: U of Chicago P, 2004. Print.

———. "The State of the Digital Humanities: A Report and a Critique." *Arts and Humanities in Higher Education* 11.1–2 (2012): 8–41. Print.

———. "Where Is Cultural Criticism in the Digital Humanities?" Gold 490–509.

McCarty, Willard. *Humanities Computing*. Basingstoke: Palgrave, 2005. Print.

McGann, Jerome J. "Position Statement from McGann." *Panel: What Is Text? A Debate on the Philosophical and Epistemological Nature of Text in the Light of Humanities Computing Research*. By Susan Hockey, Allen Renear, and McGann. N.p., 1999. Web. 25 Nov. 2012.

———. *Radiant Textuality: Literature after the World Wide Web*. New York: Palgrave, 2001. Print.

Mohr, John W., and Vincent Duquenne. "The Duality of Culture and Practice: Poverty Relief in New York City, 1888–1917." *Theory and Society* 26.2–3 (1997): 305–56. Print.

Moretti, Franco. *Network Theory, Plot Analysis*. *Stanford Literary Lab*. Stanford Lit. Lab, 1 May 2011. Web. 6 July 2012. Pamphlet 2.

"Names, Dates, People, and Places." *P5: Guidelines for Electronic Text Encoding and Interchange*. Vers. 2.2.0. TEI, 25 Oct. 2012. Web. 14 Jan. 2013.

National Endowment for the Humanities, Office of Digital Humanities. *Digital Humanities Start-Up Grants: Frequently Asked Questions*. NEH, n.d. Web. 17 Nov. 2012.

Owens, J. B. "Jack." "What Historians Want from GIS." *ArcNews*. Esri, 2007. Web. 15 Jan. 2013.

Phillips, Amanda. "#transformDH—a Call to Action Following ASA 2011." *HASTAC*. N.p., 26 Oct. 2011. Web. 24 Nov. 2012.

Pickering, Andrew. *The Mangle of Practice: Time, Agency, and Science*. Chicago: U of Chicago P, 1995. Print.

Price, Kenneth M., and Ray Siemens, eds. *Literary Studies in the Digital Age: An Evolving Anthology*. *MLA Commons*. MLA, 2013. Web. 12 Jan. 2013.

Ramsay, Stephen. "Algorithmic Criticism." *A Companion to Digital Literary Studies*. Ed. Susan Schreibman and Ray Siemens. Oxford: Blackwell, 2008. *Alliance of Digital Humanities Organizations*, n.d. Web. 25 Nov. 2012.

———. "On Building." *Stephen Ramsay*. N.p., 11 Jan. 2011. Web. 20 June 2011.

———. "Who's In and Who's Out." *Stephen Ramsay*. N.p., 8 Jan. 2011. Web. 23 Nov. 2012.

Schreibman, Susan, Ray Siemens, and John Unsworth, eds. *A Companion to Digital Humanities*. Oxford: Blackwell, 2004. *Alliance of Digital Humanities Organizations*, n.d. Web. 27 Nov. 2012.

Sculley, D., and Bradley M. Pasanek. "Meaning and Mining: The Impact of Implicit Assumptions in Data

Mining for the Humanities." *Literary and Linguistic Computing* 23.4 (2008): 409–24. Print.

Seefeldt, Douglas, and William G. Thomas III. "What Is Digital History? A Look at Some Exemplar Projects." *Faculty Publications*. Dept. of History, U of Nebraska, Lincoln, 1 May 2009. Web. 12 Oct. 2012. Paper 98.

Stallybrass, Peter. "Books and Scrolls: Navigating the Bible." *Books and Readers in Early Modern England: Material Studies*. Ed. Jennifer Andersen and Elizabeth Sauer. Philadelphia: U of Pennsylvania P, 2002. 42–79. Print.

"The Structure of the *Historical Thesaurus* of the *OED*." *Oxford University Press*. Oxford UP, 2009. Web. 25 Nov. 2012.

Svensson, Patrik. "Envisioning the Digital Humanities." *Digital Humanities Quarterly* 6.1 (2012): n. pag. Web. 23 Nov. 2012.

———. "Humanities Computing as Digital Humanities." *Digital Humanities Quarterly* 3.3 (2009): n. pag. Web. 30 Nov. 2010.

———. "The Landscape of Digital Humanities." *Digital Humanities Quarterly* 4.1 (2010): n. pag. Web. 23 Nov. 2012.

Underwood, Ted. "Topic Modeling Made Just Simple Enough." *The Stone and the Shell: Historical Questions Raised by a Quantitative Approach to Language*. N.p., 7 Apr. 2012. Web. 14 Jan. 2013.

Underwood, Ted, and Jordan Sellers. "The Emergence of Literary Diction." *Journal of Digital Humanities* 1.2 (2012): n. pag. Web. 4 Feb. 2013.

Vismann, Cornelia. *Files: Law and Media Technology*. Trans. Geoffrey Winthrop-Young. Stanford: Stanford UP, 2008. Print.

Wotherspoon, Irené. "Historical Thesaurus Database Using Ingres." *Literary and Linguistic Computing* 7.4 (1992): 218–25. Print.

———. "The Making of *The Historical Thesaurus of the Oxford English Dictionary*." Adams 271–87.