Humanistic Theory and Digital Scholarship

JOHANNA DRUCKER

Digital humanists have seen themselves within the longer tradition of the humanities, suggesting that the main value of their work resides in the creation, migration, or preservation of cultural materials (McGann). Using new platforms and networked environments, humanists entering the digital arena learned a great deal from the encounter. Expressed succinctly, the tasks of creating metadata, doing markup, and making classification schemes or information architectures forced humanists to make explicit many assumptions often left implicit in our work. Humanities content met digital methods and created projects in which the terms of production were, necessarily, set by technological restraints. (The forms of print media and their rhetorics, by contrast, were established by humanist scholars for whom debate, commentary, and interpretative exposition were so essential they drove the development of the book format and the paratextual apparatus.)

After several decades of digital work, the question remains whether humanists are actually doing anything different or just extending the activities that have always been their core concerns, enabled by advantages of networked digital technology (easier access to primary materials, speed of comparison, searching, etc.). Whatever the answer, the role of humanities scholars is crucial in the production and interpretation of cultural materials. It may turn out that data mining or large corpus processing and distant reading are substantially different from close reading and textual analysis and may bring new insights and techniques into the humanities (Moretti). But my second question frames a very different agenda: Have the humanities had any impact on the digital environment? Can we create graphical interfaces and digital platforms from humanistic methods?

The cultural authority of digital technology is still claimed by the fields that design the platforms and protocols on which we work. These are largely fields in which quantitative, engineering, and computational sensibilities prevail. Tools for humanities work have evolved considerably in the last decade, but during that same period a host of protocols for information visualization, data mining, geospatial representation, and other research instruments have been absorbed from disciplines

whose epistemological foundations and fundamental values are at odds with, or even hostile to, the humanities. Positivistic, strictly quantitative, mechanistic, reductive and literal, these visualization and processing techniques preclude humanistic methods from their operations because of the very assumptions on which they are designed: that objects of knowledge can be understood as self-identical, self-evident, ahistorical, and autonomous.

Within a humanistic theoretical frame, all of these are precepts that have been subject to serious critical rethinking. So can we engage in the design of digital environments that embody specific theoretical principles drawn from the humanities, not merely work within platforms and protocols created by disciplines whose methodological premises are often at odds with—even hostile to—humanistic values and thought? This question is particularly pressing in light of the absorption of these visualization techniques, since they come entirely from realms outside the humanities—management, social sciences, natural sciences, business, economics, military surveillance, entertainment, gaming, and other fields in which the relativistic and comparative methods of the humanities play, at best, a small and accessory role. While it may seem like an extreme statement, I think the ideology of almost all current information visualization is anathema to humanistic thought, antipathetic to its aims and values. The persuasive and seductive rhetorical force of visualization performs such a powerful reification of information that graphics such as Google Maps are taken to be simply a presentation of "what is," as if all critical thought had been precipitously and completely jettisoned. Therefore, this is a critical moment to identify core theoretical issues in the humanities and develop digital platforms that arise from these principles.

At their base—which is to say, in the encoded protocols of operating systems, machine languages, compilers, and programming—computational environments are fundamentally resistant to qualitative approaches. We can cast an interpretative gaze on these instruments from a humanistic perspective, and we can build humanities content on their base; but we have rarely imagined creating computational protocols grounded in humanistic theory and methods. Is this even possible? Desirable? I suggest that it is essential if we are to assert the cultural authority of the humanities in a world whose fundamental medium is digital that we demonstrate that the methods and theory of the humanities have a critical purchase on the design of platforms that embody humanistic values. Humanistic methods are necessarily probabilistic rather than deterministic, performative rather than declarative. To incorporate these methods, more advanced models of simulation than the literal techniques of current visualization will need to be designed.

The humanistic tradition is not a unified monolith, and any values I suggest here, or methods, will simply reflect my own disposition and training. But I think we can fairly say that the intellectual traditions of aesthetics, hermeneutics, and interpretative practices (critical editing, textual studies, historical research) are core to the humanities. These all have their own histories, of course, and in the twentieth

century humanistic assumptions were subjected to systematic critique. The insights gleaned from poststructuralism, postcolonialism, and deconstruction altered our understanding of notions of meaning, truth, authorship, identity, subjectivity, power relations, bodies, minds, nations, intelligence, nature, and almost any other ontological category of cultural thought. Computation and digital techniques have been subject to plenty of rich discussion along deconstructive and postcolonial lines. But the distinction on which I am trying to call for a next phase of digital humanities would synthesize method and theory into ways of doing as thinking. This is different from the (albeit important and insightful) directions set by digital forensics, code studies, and other analyses of digital materiality and its implications. (Kirschenbaum; Wardrip-Fruin)

The challenge is to shift humanistic study from attention to the *effects* of technology (from readings of social media, games, narrative, personae, digital texts, images, environments), to a humanistically informed theory of the *making* of technology (a humanistic computing at the level of design, modeling of information architecture, data types, interface, and protocols). To theorize humanities approaches to digital scholarship we need to consider the role of affect, notions of non–self-identicality of all expressions, the force of a constructivist approach to knowledge as knowing, observer dependent, emergent, and process-driven rather than entity-defined.

Let's backtrack a moment and review some of the ways humanistic thinking has been incorporated into digital scholarship, since some useful insights can be carried forward from this summary.

As the first phase of digital humanities reveals, the exigencies of computational method reflect its origins in the automation of calculation. Counting, sorting, searching, and finding nonambiguous instances of discrete and identifiable strings of information coded in digital form are the capacities on which digital technology performed the tasks of corpus linguistics to create Father Busa's concordance. The power of automation depended on principles that are basically *ahumanistic* principles. The creation of the concordance does not depend on interpretation even if the resulting work supports it, even if it offers a powerful instrument for extending the capacities of humanistic scholars.

The phase of digital humanities that began in the 1990s was characterized by an abundance of critical editing and repository building. Both depended on elaboration of substantial metadata and also engagement with markup languages. The scholars involved in this work certainly understood clearly the rhetorical force of argument that was embodied in the information structures they created to model content types and organize data structures. My own direct involvement with digital humanities came at the end of the 1990s, as the Perseus Project; Blake, Whitman, and Rossetti archives; the Women Writers Project; and other large-scale complex editions and repositories were already well developed. Much critical thinking was involved in the construction of these undertakings, and mature discussions of content modeling, metadata, and markup all had philosophical dimensions to

them. But as one of my colleagues was fond of saying, humanists came into those conversations as relativists and left as positivists out of pragmatic recognition that certain tenets of critical theory could not be sustained in that environment. Getting the work done—putting texts into digital formats with markup that identified content—might be an interpretative exercise, but introducing ambiguity at the level of markup was untenable, not merely impractical.

Discussions about the limits of markup languages were a sign of the maturing field. Debates at the end of the 1990s focused keen attention on the ways the hierarchical structure of XML was antithetical to the formal structure of aesthetic works. The realization did not lead to new methods grounded in humanistic theory. Arguments emerged; discussion was heated, insights profound, then scholars shrugged and went back to coding. The basic conclusion was that to play in a digital sandbox one had to follow the rules of computation: disambiguation and making explicit what was so often implicit in humanities work was the price of entry. The benefits outweighed the liabilities. The capacity to aggregate all versions of Blake's prints, to assemble the many witnesses and expressions of Rossetti's double works in a single web environment in which their relations could be studied, annotated, made evident for the first time—these were remarkable leaps beyond the printed approach that had defined the horizon of possibility just a few years earlier. The combination of digital tools and networked environments pushed scholarship to a new level.

But if this is what the encounter of humanities work and digital tools was like, then what could the encounter of humanities "tools" bring to digital contexts?

The question has more in it than a simple shift of emphasis. The tools in scare quotes stands for precepts, theoretical approaches, basic ways of thinking that are fundamentally different in the two realms. Humanities approaches would proceed from a number of very specific principles. The first of these is that interpretation is performative, not mechanistic—in other words, no text is self-identical; each instance or reading constructs a text; discourses create their objects; texts (in the broad sense of linguistic, visual, acoustic, filmic works) are not static objects but encoded provocations for reading. These are familiar arguments within the digital and broader humanities community, and finding ways of showing these principles informed our work at SpecLab, particularly in the realization of the Ivanhoe platform. The project of showing interpretation, modeling it, making a composition space in which ambiguity and contradiction can coexist, where the non–self-identicality of objects can be made as evident as their codependent relation to the social fields of production from which they spring (a relation premised on the constructedness of their identity, rather than the relation of an "object" to a "context") remains undone. Perhaps it is undoable, since the very instantiation of an interpretative act would reify it in ways that create at least the illusion (or delusion) of fixity. That built-in paradox does not obviate the need to experiment with humanistic precepts in the design of digital environments, however; and that takes me to the current situation.

As we turn our attention from that first phase of textual studies and critical editing, we see the phenomena I noted appear on the horizon: data mining with all its attendant visualization techniques, engagement with geospatial tools, and the use of timelines and other graphic conventions from fields rooted in empirical investigation, management, and social sciences where predictable outcomes and repeatable results are at the desired outcomes even if they are not absolute tenets of belief. The graphical tools that are used for statistical display depend, in the first instance, on quantitative data, information that can be parameterized so that it lends itself to display. Virtually no humanistic data lends itself to such parameterization (e.g., what year should a publication be dated to in the long history of its production and reception?), and it is in fact precisely in the impossibility of creating metrics appropriate to humanistic artifacts that the qualitative character of *capta*, that which is taken as interpretation rather than *data*, comes sharply into relief.

But if the premises on which quantitative information might be abstracted from texts or corpora raise one set of issues, the use of graphical techniques from social and natural sciences raise others. Graphs and charts reify statistical information. They give it a look of certainty. Only a naive viewer, unskilled and untrained in matters of statistics or critical thought, would accept an information visualization at face value. But most humanists share with their social and natural science colleagues a willingness to accept the use of standard metrics and conventions without question in the production of these graphs (e.g., using the same sized unit for all hours of the day seems bizarre for a humanist charting the narrative of *Mrs. Dalloway* or film critic analyzing *Memento*). Even the supposedly simple act of counting entities in any study, for instance, raises a host of questions, while creating a chart on which to graph the information immediately throws the untrained humanist into a world where the difference between a bar chart and a continuous graph should be understood in professional terms. Amateur chart making provides a rash of examples of ways not to do things with numbers, but we are still only in the baby pool of information visualization, and already the waters are perilous to navigate. Statisticians are concerned with probabilities, not certainties. They do not count things; they model conditions and possible outcomes. Data mining in the humanities has largely depended on counting, sorting, ordering techniques—in essence, some automated calculations. Statistical modeling has factored less, at least to date, in the analytic tool kit of critical digital work with texts. Stylometrics, attribution studies, natural language processing, and other higher level analyses have long made use of these sophisticated modeling techniques, but graphing ambiguous and partial knowledge is still in its early stages. Archaeologists who work in reconstruction, trying to extrapolate various models of possible form from fragmentary remains, have created spectral palimpsestic methods of portraying uncertainty in their digital imagery. Some conventions from art history and architectural study include an inventory of techniques for indicating distinctions between known and projected or imagined evidence.

Probability is not the same as ambiguity or multivalent possibility within the field of humanistic inquiry. The task of calculating norms, medians, means, and averages will never be the same as the task of engaging with anomalies and taking their details as the basis of an argument. Statistics and pataphysics will never meet on the playing fields of shared understanding. They play different games, not just the same game with different rules. However, the dialogue among probabilistic methods, modeling, and visualization holds much promise for humanistic work ahead as the shift from counting and sorting to predicting and presenting uncertainty advances in the digital humanities.

But once we depart from the realms of probability into the world of empirical or positivist representation, the friendly atmosphere evaporates. The theoretical underpinnings of humanistic interpretation are fundamentally at odds with the empirical approaches on which certain conventions of temporal and spatial modeling are based. All maps are constructions, and the history of cartography, like the histories of other graphical forms of knowledge, is filled with productive debates about the ideological and rhetorical force of mapping techniques. Humanists are certainly well aware of these complexities, but the tendency to wave them aside as mere nuance in the rush to adopt Google Maps or a standard projection, even knowing full well that any presentation of an oblate spheroid (the earth) onto a flat surface (map or screen), is fraught. But more is at stake for a humanist than the technical problems of projection. Both space and time are constructs, not givens. As constructs they come into being in a codependent relation with their discursive or experiential production. If I am anxious, spatial and temporal dimensions are distinctly different than when I am not. When the world was bounded by the Mediterranean it was a different world than the one seen from space. These are not different versions of the same thing. The entire theoretical weight of constructivist approaches to knowledge stands against such a reductive pronouncement.

But the social, cultural, experiential, psychological, and phenomenological aspects of spatial and temporal modeling have had very little play in the history of graphical expressions. No space is a given, and no standard metric adequately serves the analysis and presentation of temporal experience. When we expand this discussion to include the spatial and temporal richness of cultural artifacts, the vision of a complex worldview begins to beg for digital means of exposure.

Start with a simple example to make the case. What kind of map should be used to show the locations in which cuneiform tablets were produced in the ancient near east? The configuration of linguistic boundaries and social groupings within the Mesopotamian region would offer one version of that geography. A simple outline of landmasses and landforms would offer another. A map with modern nation-states outlined for easy reference to current events and contemporary place names would offer another. None are commensurate with each other. None present the experience of geography within which the tablets were produced. How do we understand the geographically specific references to lands or cities within those texts if

they are only points on a contemporary map? Most pernicious of all, because it is so ubiquitous, is the already cited Google Maps. The reification performed by a digital photography imaging system trumps in the "this is what is" game. Put the average person in front of one of those screens, and they believe they are seeing *the* world, not a constructed version of it. Savvy as we are, the concessions made in the name of efficacy may turn out to have too heavy a price if we cannot imagine alternative spatial models for experienced and situated approaches. Once those aerial maps come into play, they set a baseline against which historical maps are referenced. Then the assumption built into the analysis is that the other images are deviations from that photographic record of Google satellite. My point is not so much that we are naive users, as we may or may not be, but that the force that photographic renderings carry when they are used as the ground on which other images are referenced makes it very difficult to dislodge the subtler conviction that these satellite images are the "real" earth.

If spatial location is problematic and needs historical and cultural relativization from the outset, then navigation and motion through and across landscapes as a construction of space introduce another level of challenge. In a striking visualization created for the digital project *Mapping the Republic of Letters*, letters between eighteenth-century correspondents move from one geographic location to another as if by airmail. The perfect lines of light, looking for all the world like a contemporary graphic of air traffic, make the connection between the point of origin and place of delivery a smooth, seamless, unitary motion. All letters move at the same speed in this simulation, and all leave and arrive as if by airmail. No matter that the shape of coastlines and borders have changed or that methods of travel involved many kinds of detours, checks, delays, and roundabout routes at varying speeds and through multiple levels of diplomatic negotiations with official or unofficial interlopers. How would the variable spatial relations of communications be modeled to show these complexities? If that were possible, then the humanistic principles of cultural and historical constructions of space would be the foundation from which the visualization would arise. As it is, humanities scholars attach their analysis to the armature of preexisting graphical conventions. The primary strategy for undoing the force of reification is to introduce parallax and difference, thus taking apart any possible claim to *the* self-evident or self-identical presentation of knowledge and replacing this with a recognition of the made-ness and constructedness that inhere in *any* representation of knowledge.

Humanistic conventions for the graphical production of spatial knowledge and interpretation would spring from the premises of situatedness and enunciation. By situatedness, I am gesturing toward the principle that humanistic expression is always observer dependent. A text is produced by a reading; it exists within the hermeneutic circle of production. Scientific knowledge is no different, of course, except that its aims are toward a consensus of repeatable results that allow us to posit with a degree of certainty what the qualities of a frog are by contrast

with those of a cat. We take such distinctions to be observer independent even if the system according to which we represent, name, analyze, arrive at, and value these distinctions is subject to the same analytic principles as any other expression of human knowledge. By enunciation I am suggesting we bring the idea of spatial relations in digital humanities back into conversation with the principles elaborated in film studies, architecture, cultural geography, and visual and textual studies that took seriously the analysis of ways representational systems produce a spoken subject as well as engage a speaking subject within their "regimes," as we used to say. Situatedness and enunciation are intimately bound to each other, though the first resides more squarely in individual experience and its representation, the second in the cultural systems that speak through us and produce us as social subjects. This is not the place to go back through deconstruction and poststructuralist thought and its legacy but, rather, to sketch ways these principles could inform digital humanities work going forward.

If we embark on a project to study inscriptions in the Roman forum, for instance, should we simply take a virtual simulation, a fly-through model made on a standard 3-D software platform and place the writing on the variously appropriate surfaces? Or should the ways signage works as an articulation of space, the "species of spaces" described by Georges Perec, be merged in a mash-up of Christopher Alexander's "pattern language" with Gaston Bachelard's poetics of space and the Foucauldian analysis of spatialized regimes of disciplinary control structured into and structuring architectural and social relations? A ridiculous proposition emerges from that impossible sentence. Or does it? The principles of humanistic spatial modeling arise from a system of enunciation, as a system, in which eyelines, scale, proportion, surface texture, lighting, color, depth, openings, routes of egress, degrees of security, and other tractable elements might be given a value so that the space is constructed as an effect, rather than a basis, of experience. Without minimalizing the complexity involved in realizations and aware of the risk of cartoonish special effects, my point is that by not addressing these possibilities we cede the virtual ground to a design sensibility based in engineering and games. The risk is that the persuasive force of such representations tips the balance of belief; the Roman forum of a virtual world becomes far more satisfying and real than the fragmentary evidence and analytic complexity on which humanistic scholarship arises. But the partial nature of knowledge is another crucial tenet of humanistic belief, and rendering a seamless image of "what is" prevents the imaginative critical faculties from engaging with the all important question of "what if?" The parallax views that arise in the interstices of fragmentary evidence are what give humanistic thought its purchase on the real, even with full acknowledgment that knowing is always a process of codependencies.

To finish the examples and summarize these premises, I return to the unfinished project of modeling temporality. Temporal coordinates, like spatial ones, contain

many variations across history, culture, and individual experience. The experience of temporality, like that of space, is already inflected by cultural circumstance. We feel time differently, because it is a different kind of time, than other generations. Life expectancy, control over lighting, and mechanical clocks and their presence in daily routines are all structuring features of the contemporary experience of time that are fundamentally different from those of, say, citizens of Augustan Rome, pre-conquest Mayan communities, or seventeenth-century French aristocrats. Understanding the basic experience of time is a subject of humanistic research that already differentiates it from the "time" imagined by researchers checking the frequency of urination in lab rats or number of commuters on a stretch of highway during certain hours of a commute.

Within humanistic documents, time is as frequently noted in relativistic terms as in absolute ones. As the community of "tensers" makes clear, the notions of before and after, expressions of temporal sequence in the forms of verbs and auxiliaries, are evidence of the richness of expression in relation to the tensions between ambiguity and certainty in linguistic communications. The time described in aesthetic texts has other features, with its marked discontinuities, the discrepancy between times of telling and times of the told, between flashbacks and foreshadowings, recollection and anticipation. These are all features of temporality within the horizon of humanistic inquiry and study. We know these characteristics and work with them without any difficulty, except when it comes to using graphical means of expression or analysis. Then the distinction between time and temporality evaporates in the fallback position of taking timelines as an adequate measure on which to map the complexities of historical and aesthetic experience.

The experience of temporality; the representation of that experience; the temporal dimensions of narration and mutabilities of duration, extension, telescoping contraction, and explosive expansion all need a graphical language for analysis as well as for the presentation of argument. Relations among various documents and artifacts within a corpus raise another set of tangled theoretical issues, since rates at which different members of a communication network send or receive information produce the circumstances in which they collectively imagine a sequence of events. The very notion of an "event horizon" cannot be standardized within a single frame. Each participant's sense of when an event occurred will differ, and hence the liabilities of diplomacy and delicacies of policy decisions within varying temporal assumptions.

And what of incommensurate temporalities within a single frame? If I chart my life according to segments based on places I have lived, the units of time and the relation among them has one shape, still not consistent since, as with most humans, I assign the eons of childhood a different value from the fleeting units of adult life. But what if I try to divide my life into segments according to relationships, affective connections, and emotional ties? What are the seasons of the

heart, the dimensions of trauma? How long does something last? This segmentation, or impossibility of even making a segmentation, does not match the first exercise. These are incommensurate chronologies and yet within a single life, and some shared milestones connect these different versions of the lived.

The original project of temporal modeling, conceived a decade ago, engaged these issues and developed a basis on which to make a graphical authoring space that would express some of these principles. Our computational muscle was not adequate for the task. I return to the points raised earlier about engaging with probabilistic systems rather than deterministic ones that rely on discrete or fixed frames of reference or standard metrics. Flexible metrics, variable, discontinuous, and multidimensional will be necessary to realize a humanistic system for the graphical analysis of temporal relations.

Just as the attempt to place James Joyce's spatial references onto a literal street map of Dublin defeats the metaphoric and allusive use of spatial reference in *Ulysses*, flattening all the imaginative spatial experience that infuses the text with images of an urban imaginary irreducible to its material counterpart, so the task of putting pins into a Google map or charting the times of lived experience on a single unvarying linear scale is a grotesque distortion—not merely of humanistic approaches to knowledge as interpretation, but the very foundation from which they arise. The modern paradigms of logic and rationality, now thoroughly subjected to post-Cartesian critique, underpin those visualizations, those seemingly self-evident maps and charts. But the humanities are not a mere afterthought, simply studying and critiquing the effects of computational methods.

Humanistic theory provides ways of thinking differently, otherwise, specific to the problems and precepts of interpretative knowing—partial, situated, enunciative, subjective, and performative. Our challenge is to take up these theoretical principles and engage them in the production of methods, ways of doing our work on an appropriate foundation. The question is not, Does digital humanities need theory? but rather, How will digital scholarship be humanistic without it?

BIBLIOGRAPHY

Kirschenbaum, Matthew. *Mechanisms: New Media and the Forensic Imagination.* Cambridge, Mass.: MIT Press, 2008.

McGann, Jerome J. *Radiant Textuality.* New York: Palgrave Macmillan, 2001.

Moretti, Franco. "Conjectures on World Literature." *New Left Review.* January–February 2000. http://www.newleftreview.org/A2094.

The Perseus Archive. http://www.perseus.tufts.edu.

The Republic of Letters. http://shc.stanford.edu/collaborations/supported-projects/mapping-republic-letters.

The Rossetti Archive. http://www.rossettiarchive.org.

The Walt Whitman Archive. http://www.whitmanarchive.org/.
Wardrip-Fruin, Noah. *Expressive Processing*. Cambridge, Mass.: MIT Press, 2009.
The Willam Blake Archive. http://www.blakearchive.org/.
The Women Writers Project. http://www.wwp.brown.edu.