The Limits of Modelling: Data Culture and the Humanities

Harvey Quamen
University of Alberta

Abstract

Willard McCarty's Humanities Computing (2005) opens a dialogue about modelling in the humanities. I extend that conversation by complicating his version of the model and by looking at the limits of modelling to find out what it is not. I suggest that, unlike scientific models, humanist models cannot be separated from the mode of Humanism that produced them. I argue that databases are not always models and that the boundary between model and database helps us to understand the advantages and limitations of both. A renewed appreciation of database culture can help to move humanist scholarship in new directions.

The INKE Research Group comprises over 35 researchers (and their research assistants and postdoctoral fellows) at more than 20 universities in Canada, England, the United States, and Ireland, and across 20 partners in the public and private sectors. INKE is a large-scale, long-term, interdisciplinary project to study the future of books and reading, supported by the Social Sciences and Humanities Research Council of Canada as well as contributions from participating universities and partners, and bringing together activities associated with book history and textual scholarship; user experience studies; interface design; and prototyping of digital reading environments.

Harvey Quamen is Associate Professor of English and Humanities Computing at the University of Alberta. Previously a computer programmer, he is currently active in several large digital humanities research projects, including those sponsored by Editing Modernism in Canada, the Implementing New Knowledge Environments projects, and the Canadian Writers Research Collaboratory. Email: hquamen@ ualberta.ca .

*Now, may any regular body, be Transformed into an other, &c. Now,
any regular body: any Sphere, yea any Mixt Solid: and (that more is)
Irregular Solides, may be made (in any proportiõ assigned) like vnto the
body, first giuen. Thus, of a* Manneken, *(as the Dutch Painters terme
it) in the same* Symmetrie, *may a Giant be made: and that, with any
gesture, by the Manneken vsed: and contrarywise. Now, may you, of
any Mould, or Modell of a Ship, make one, of the same Mould (in any
assigned proportion) bigger or lesser. Now, may you, of any *Gunne, or
little peece of ordinaũce, make an other, with the same* Symmetrie *(in all
pointes) as great, and as little, as you will. Marke that: and thinke on it.*

(Dee, 1570, p. C3r)

The text from which this epigraph was taken was published in London in 1570.
It was written by John Dee, perhaps best known as Queen Elizabeth I's resident
mathematician and alchemist. Dee's text served as the preface to Sir Henry Billingsley's
new translation – the very first translation into English, in fact – of Euclid's *Elements
of Geometrie*. Dee's preface situated Euclid's geometry amid the existing branches of
sixteenth century mathematics and articulated why readers might find an English
translation of the *Elements* useful. Dee, a sometime astronomer and navigator, was
well enough acquainted with new sixteenth century shipbuilding techniques to be
able to use the word "modell" as the working shipwrights did. Dee knew that the new
ships – with larger keel, beam, and "depth in hold" (Friel, 2009, para 12) dimensions
in order to increase carrying capacity, or "tons burden" (Friel, 2009, para 14) – were
being rescaled from older, smaller ships. That job was performed not by the carpenters
themselves, however, but rather by designers who constructed paper or wood models
so that they could more easily manipulate the geometrical proportions of the ships.
Ships were no longer built "by eye," as they had been in the old tradition, "but after
a careful process of design" (Barker, 1988, p. 540). In his preface, Dee argued that
Euclid's geometric principles were applicable everywhere (and were therefore worth
reading) by grounding those claims in specific discourses that might be familiar to his
contemporary English readers: proportion, mathematical ratio, painting, weaponry,
and shipbuilding. This new translation of Euclid's *Elements* would explain to a
literate audience the mathematical foundations supporting the newly emergent art of
modelling. Indeed, the *Oxford English Dictionary* (2011) claims Dee's text as the first
published use of the word "modell" in English.

My assertion here is that the humanities – especially the digital humanities – are
increasingly negotiating a terrain not unlike that of sixteenth century shipwrights.
We ponder the relationship between ships and models, between texts and contexts,
between schemas and data, between interpretation and algorithm, between foreground
and background. Coming to terms with that dynamic is an important part of
contemporary digital humanities research. Uncertain where the line falls between
the "D" and the "H" of the digital humanities, we incessantly crisscross the terrain,
seeking it out. Worried that our methodologies are not rigourous enough, we borrow
from both the hard and soft sciences, wondering what might be gained (or lost) from
such interloping. I see this tension most acutely these days in our ongoing dialogues
between models and data – a tension exemplified, at least for my purposes here,
in Willard McCarty's *Humanities Computing* (2005). The argument here is that an

agonistic tension between data and modelling can be a very productive venue for the digital humanities – and there are very good reasons for *sustaining*, rather than *resolving*, that tension. As Dee recommended, let us mark that, and think on it.

"Circumstances where no analogical congruence obtains"

Willard McCarty asserts in *Humanities Computing,* that "analytical computing in the humanities *is* modelling" (McCarty, 2005, p. 71) and his text explores the implications of that statement for theory and praxis in the humanities. Indeed, the entirety of his first chapter is devoted to the process of modelling, and seeing it as "an iterative, perfective process" (McCarty, 2005, p. 23), McCarty attempts to work out its dynamics. His text is not only my point of departure but it also forms a kind of locus around which I will orbit in the pages that follow, wrestling with his argument and its implications.

McCarty posits two definitions of model: "a representation of something for purposes of study" and "a design for realizing something new" (McCarty, 2005, p. 24). The distinction, he maintains, is between a model *of* and a model *for,* but the difference makes more sense to me in terms of *representation* and *design*, respectively. Citing a wealth of resources across a wide range of disciplines, McCarty finds his touchstone: modelling is "*the continual process of coming to know by manipulating things*" (McCarty, 2005, p. 28, emphasis in original). He stresses the contingency of models, their temporariness, and he underlines the idea that models are instructive even when they fail. Although the model might go by a variety of names – analogy, representation, diagram, map, simulation, experiment, plan, etc. – they all share the basic premise of working towards "an approximating convergence" (McCarty, 2005, p. 28).

So far so good. His conversation is a very thorough discussion about what a model *is*. But because of the long and vexed relationship between the humanities and modelling, perhaps even more interesting is McCarty's catalogue of what a model is *not*:

- Example 1: A model is not a *concept*, he argues, because a concept tries to "freeze [a phenomenon] into an ahistorical abstraction" (McCarty, 2005, p. 23) and, unlike models, concepts are utterly useless when they fail.
- Example 2: Insofar as computational models remain "temporary states in a process of coming to know," they remain models, but once they become "fixed structures of knowledge," "these would not then be *models* and would no longer have reason to exist in software" (McCarty, 2005, p. 27). I understand McCarty to be arguing here against simple website delivery of humanities scholarship, especially when he helpfully adds that "the greater potential is for *computers as modelling machines, not knowledge jukeboxes.* To think of them as *only* the latter is profoundly to misunderstand human knowledge" (McCarty, 2005, p. 27, emphasis in original). A knowledge jukebox is not, then, a model.
- Example 3: Mapping. "Because its action is bounded and its result defines a world," McCarty writes, "mapping better suits the theoretician's than the experimenter's view." And because mapping evokes "the disciplinary politics of

Quamen, Harvey. (2012). The Limits of Modelling: Data Culture and the Humanities. *Scholarly and Research Communication, 3*(4): 040148, 14 pp.

3

conceptual territory" – a punning reference, which I understand as meaning that the inaccurate map is as useless as a false concept – "it is not the term we need" (McCarty, 2005, p. 33).
- Example 4: Simulation – see below.

A fourth example, simulation, presents perhaps the most interesting challenge to modelling and I'd like to linger upon this example. Citing biologist Glenn Rowe ("a great many systems seem to have an inherent complexity that cannot be simplified" [McCarty, 2005, p. 34]), McCarty demarcates the boundary between modelling and simulation as an issue of complexity and one that has interesting theoretical implications. Simulations lead not only to a "displacement of reality" but also to "a more general problem with the notion of modelling under circumstances where no analogical congruence obtains" (McCarty, 2005, pp. 34–35). Michael S. Mahoney underscores that point in suggesting that, as scientific inquiry has historically moved toward simulation, "the job of the scientist became 'to build models that matched the phenomena, without concern for whether the model was 'true' in any other sense'" (McCarty, 2005, p. 228, note 30). I will cite McCarty at length:

> Like game-playing, simulation tends to forgetfulness of the mechanism by which it is created so long as its terms of engagement (expressed in parameters and algorithms) are fixed. Unfix them . . . and the simulation becomes a modelling exercise directed to exploring the question of that attitude. Thus simulation crosses over into modelling when the constants of the system become variables. Modelling, one might say, is a *self-conscious* simulation, and simulation is *assimilated* modelling. The difference between the two is . . . far more significant than these terms indicate. (McCarty, 2005, p. 35)

For McCarty, the boundary of the model is marked by a degree of self-consciousness and the "unfixing" of parameters and algorithmic constants into variables, at least insofar as these are manageable by our methodologies. This dynamic movement between simulation and modelling has both methodological and theoretical implications that are worth investigating in more detail. Brian Cantwell Smith (1996) suggests that the problem here is both broader and more radical than McCarty's discussion suggests, however. As Smith has argued, we have no reliable means by which to measure what McCarty calls "analogical congruence." The radical suggestion here is that not only do simulations challenge the analogical congruence between model and world, but perhaps all models do too. Reproduced here is a figure from Smith's article, "Limits of Correctness in Computers," and his description of it (Figure 1):
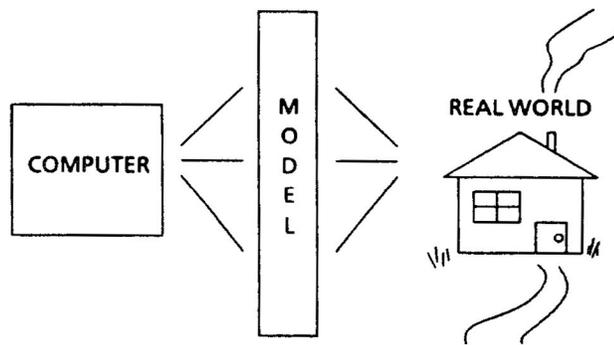
Figure 1. Computers, models, and the embedding world.

Figure 1: Brian Cantwell Smith diagrams how models mediate between computers and the real world.

Illustrating the relationship between the computer, the model, and the real world, Smith writes that

> In philosophy and logic, as it happens, there is a precise mathematical theory called "model theory." You might think that it would be a theory about what models are, what they are good for, how they correspond to the worlds they are models of, and so forth. . . . Unfortunately, however, model theory doesn't address the model-world relationship at all. Rather, what model theory does is to tell you how your descriptions, representations, and programs *correspond to your model*. (Smith, 1996, p. 817, emphasis in original)

"Simulation," he continues, "as our diagram should make clear, *tests only the left-hand side relationship.* Simulations are defined in terms of models; they don't test the relationship between the model and the world. That is exactly why simulations and tests can never replace embedding a program in the real world" (Smith, 1996, pp. 824–825, emphasis in original).

Smith's position, then, is much more radical than McCarty's insofar as the necessary "act[s] of violence" that render all models "inherently partial" (Smith, 1996, p. 816) also imply that "we don't have a theory of the right-hand side" (Smith, 1996, p. 818). At all. The entire process of modelling becomes problematic, therefore, and is always so – not just under exemplary, exceptional circumstances but in all cases. Smith's caveat reminds us of the kind of self-consciousness model-building enterprise that McCarty calls for. It's tempting to assume that our iteratively generated models present us with more and more accurate representations of the real world, but Smith cautions us to remember that we're actually bringing our model by iterations into alignment with our presumptions, assumptions, and theories, not with the real world itself.

That is no cause for despair. It does not occasion the abandonment of models. It does, however, introduce a new set of variables into our model-building enterprises – namely, the subject-position of the modeller, the partialness and contingency of our theories, and, as Smith says, "as well as modelling the artifact itself, you have to

Quamen, Harvey. (2012). The Limits of Modelling: Data Culture and the Humanities. *Scholarly and Research Communication, 3*(4): 040148, 14 pp.

5

model the relevant part of the world in which it will be embedded" (Smith, 1996, p. 822). As a way of introducing those ideas into a more general process of modelling, I would like to examine and to extend an example that McCarty presents: a model of personification tropes in Ovid's *Metamorphoses*.

Romans, romantics, and neu-romancer

To provide an example of "laboratory practice," McCarty details a model he hopes to be "persuasive yet with enough imperfections to raise a number of unanswered questions for research" (McCarty, 2005, p. 53). It is those unanswered questions I find intriguing and useful for a discussion about the limits of modelling. Personification – "any rhetorical act that transgresses normal ontology by investing an entity with a human imaginative component, whether or not the result is anthropomorphic" (McCarty, 2005, p. 57) – is a rich thematic in Ovid's work because "the trope is one of his principal instruments for depicting the world as he imagined it" (McCarty, 2005, p. 55). McCarty's result is a hybrid database/spreadsheet model that produces a visual chart of degrees of personification (Figure 2):
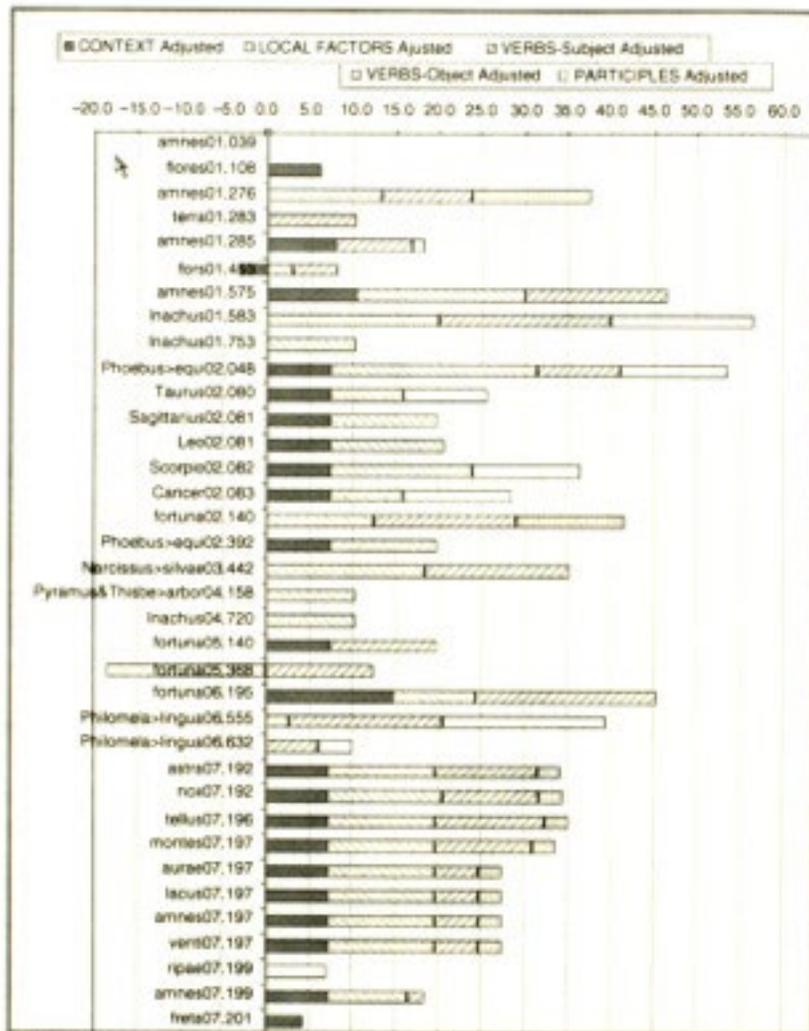

Figure 2: Willard McCarty charts various degrees of personification in Ovid's *Metamorphoses***.**

McCarty's description includes a detailed analysis of the database schema and the encoding techniques used to populate it with data. But in the end, perhaps predictably, the model proves inadequate, yet it does so in interesting ways. The data of Ovid – the specific details themselves – challenge the model, illuminating its flaws so that McCarty can refine it.

But the model's context – Smith reminded us that systems are always "embedded in complex social settings" (Smith, 1996, p. 813) – remains problematic. Clearly, the version of "personification" modeled here is time- and culture-specific, but those factors are missing from the kind of model presented here. In fact, the model appeals to McCarty precisely because of its lack of context: "With computing it becomes much more practical not only to describe and find patterns in nature and culture but also *almost* – a quantum qualification of the greatest importance – to detach them from their contexts and give them independent life" (p. 172). That self-conscious qualifier, "almost," is enough to bring the model back down to earth. It underscores the limitations of the model, asserting a kind of asymptotic boundary beyond which the model cannot pass. Approaching it, the model is doomed to fail.

Here, the categories "human" and "non-human" – and, more importantly, the qualities we attribute to each of those two categories – have shifted across the centuries and across cultures. The personification model may indeed illuminate something about the ways in which those two categories worked for first century Romans but the model does not, and cannot, stand alone autonomously as an ahistorical, transcendent description of how personification has worked in the centuries since Ovid set pen to parchment. The gendered nature of Latin (I wonder: are gendered nouns examples of personification?), the influences of Roman pantheistic and Stoic religious tropes (which divided agency and will across the world in ways more complex than simply human and non-human), and indeed the legal and civic status of the "person" in Roman culture should all influence the construction of the model at hand. They are the immediate context into which that model is embedded.

A model of personification, therefore, needs to be contextualized within a longer history of "personhood" – a fascinating project that I find more intelligible when articulated as part of the history of Humanism – in order to see how the ever-shifting characteristics of the "person" have been (mis)applied to things non-human. Personification is, as McCarty makes clear, a fundamental category mistake. It "violates ontology" (McCarty, 2005, p. 56). Arguments we have accrued over the years about the pathetic fallacy (a concept invented by John Ruskin in his 1856 book, *Modern Painters*) are attempts to police and to contain any contamination between the two categories of human and non-human. When the Luddites of the early nineteenth century protested the mechanization of the British textile industry, they contributed not only to an escalating cultural phobia about mechanical agency but also to a concomitant anxiety about the purview of the Human. The purification of such abstract categories has been, as Bruno Latour (1993) has argued, the hallmark of modernity itself.

Taking Mary Shelley's *Frankenstein* (1818) as a source-text of personification ought to yield results quite different from Ovid's. Fast-forward yet again to William Gibson's cyberspace thriller, *Neuromancer* (1984), and we find another shifting chapter within

Quamen, Harvey. (2012). The Limits of Modelling: Data Culture and the Humanities. *Scholarly and Research Communication, 3*(4): 040148, 14 pp.

7

the long saga of human vs. machine. That the two major Artificial Intelligences in Gibson's novel display the dominant traits of *improvisation* and *personality* signal two mechanical qualities that would have been wholly alien to the Luddites nearly two centuries earlier. However – and this is important to note – only those contemporary readers who still subscribe to Romantic notions of the *machine* can be persuaded that Gibson has committed the unforgivable humanist *faux pas*: the pathetic fallacy. Only from neo-Romantic perspectives can Gibson be seen as having forgotten that computers are nothing but cold, mechanistic calculators that could never display any qualities that we might, even on a bad day, accidentally misinterpret as a personality. But – to ask explicitly here the implicit question – have McCarty's categories of personification been sufficiently inoculated against that particular Romantic ideology? Did it make sense to the Romans to see these phenomena in the way that McCarty does: as ontological *violations*? A perfectly fine, instructive model about how personification worked *for the Romans* is perhaps less illuminating when we try to use it as a model about how personification worked *for the Romantics*. But that in no way lessens its impact as an explanation about how first-century Roman culture worked when Ovid tried to navigate its intricacies.

Those nineteenth century Romantic dynamics are still commonplace today, of course, but these mythologies are coming under increasing scrutiny as computers, robots, and artificial intelligence perform more and more activities that we would once have labelled exclusively human. The classical humanist stance is represented poignantly (and even ruefully) in Harold Bloom's foreword to a recent edition of Northrop Frye's *Anatomy of Criticism*: "Overpopulation in literature has gone beyond Malthusian dimensions, and soon the world's computers will enhance a Noah's flood of productivity. If I live long enough, I fully expect individual computers themselves to declare their possession of personality and genius, and to bombard me with the epics and romances of artificial intelligence" (Bloom, 2000, p. xi). How preposterous, Bloom seems to intimate, that a machine could reveal anything interesting – let alone reveal anything at all – about the deep structures of the universal human condition. "Where shall I turn?" Bloom asks plaintively (Bloom, 2000, p. xi). Data accrues and challenges the model, even when that model is Humanism itself.

In contrast, scholars like Donna Haraway (1991) and Katherine Hayles (1999) have argued that contemporary technology is unravelling the strict machine/human dichotomy that's become familiar to us over the past two centuries and that we are entering the age of Posthumanism. I understand their point to be less that Humanism is going extinct, and rather that a revised Humanism (version 2.0?) is in the making. Following Thomas Kuhn's nomenclature in *The Structure of Scientific Revolutions* (1996), we might say that Humanism is undergoing a paradigm shift. But what I find extraordinarily appealing about McCarty's argument is that a nice set of personification models tracking chronologically across our humanist timescape might help us to illuminate that paradigm shift in all its baroque, self-contradictory ways.

What might that look like? In *Humanism*, author Tony Davies remarks that Humanism has effected "striking semantic reversals" that, in the sixteenth century, "rendered words like *individual* (originally 'inseparable') and *identity* ('sameness') over to meanings almost exactly the opposite of their traditional ones" (Davies, 2008, p.

Quamen, Harvey. (2012). The Limits of Modelling: Data Culture and the Humanities. *Scholarly and Research Communication, 3*(4): 040148, 14 pp.

17). The fact that today the watchwords of Humanism point in the other direction – towards an independent, autonomous, unique personhood whose major task is the expression of an inner self – leads Davies to conclude that "'Renaissance humanism,' expressive of an essential humanity unconditioned by time, place or circumstance, is a nineteenth-century anachronism" (Davies, 2008, p. 24). Davies sees contemporary Humanism, too, as a distinctively nineteenth century invention that "institut[ed] a new and distinctively modern notion of human individuality, a notion projected back onto the writings of fifteenth- and sixteenth-century *umanisti*, but demonstrably shaped by and inseparable from nineteenth-century conditions and concerns" (Davies, 2008, p. 19). Davies' thesis is illustrative: Humanism is an ideological model that makes the most sense when seen against its materialistic background of emergent modernity, politics, industrialism, and empire. What if we tried to model *that*? McCarty's model of personification is one step in that direction – but it remains a foreground without a background.

"The study of individual cases"

One serious challenge to modelling in the humanities is that the humanities historically have adopted methodologies that are less amenable to abstract generalization than those used by the sciences. In his article "Clues: Roots of an Evidential Paradigm" (1989), Carlo Ginzburg delineates a fascinating history of what he alternately calls "evidential" or "conjectural" or "semiotic" methods – methods that have, he argues, informed humanities scholarship since at least the late nineteenth century. He places the Italian astronomer Galileo at a decisive turning point in history after which certain disciplines favoured the *telling detail* whereas the sciences, following Galileo, sought *abstract generalizations*. Beginning with the image of the prehistoric hunter following the tracks of wild prey, Ginzburg constructs a sweeping "evidential paradigm" that includes medicine, textual studies, paleography, art history, psychoanalysis, criminology and, ultimately, the whole of the humanities themselves. He summarizes:

> It should be clear by now that the group of disciplines which we have called evidential and conjectural (medicine included) are totally unrelated to the scientific criteria that can be claimed for the Galileian paradigm. In fact, they are highly qualitative disciplines, in which the object is the study of individual cases, situations, and documents, precisely because they are individual. . . . Galileian science, which could have taken as its own the Scholastic motto *Individuum est ineffabile* ("We cannot speak about what is individual"), is endowed with totally different characteristics. (Ginzburg, 1989, p. 106)

The kinds of insights the humanities could produce, then, would have immediate and profound impact upon the concreteness of people's daily lived experience (that's still a popular trope among humanists), but at a severe cost: "the inability to make use of the powerful and terrible weapon of abstraction" (Ginzburg, 1989, p. 115). Humanists sacrificed the abstract model by privileging the individual semiotic sign – the clue.

Ginzburg's analysis helps to explain why McCarty's argument about modelling in the humanities seems an imperfect fit. As if some kind of impedance mismatch were at work, the model of personification in Ovid's *Metamorphoses* is best critiqued – even in McCarty's own analysis – by the experiential, the idiosyncratic, the cultural. "Adjustments to accommodate special pleading . . . are clues," he says, "to an as yet undiscovered factor, one in need of refinement, a structural difficulty in the database or simply a misreading" (McCarty, 2005, p. 63). Context, he warns, cannot be easily resolved; we must maintain a sense of "justice" towards the text (McCarty, 2005, p. 56); a "degree of objectivity" and "absolute consistency" must be achieved (hence, "subjectivity" and "inconsistency" ruin the model [McCarty, 2005, p. 61]); and telltale incongruous clues pose special challenges ("the situation is quite odd," for example, when Cephalus kills Procris [McCarty, 2005, p. 64]). In the end, the model "ill accords with a reasonable sense of the phenomenon" (McCarty, 2005, p. 65) – it simply does not fit the data. Ginzburg again: "Individualizing knowledge is always anthropocentric, ethnocentric, and so on" (Ginzburg, 1989, p. 112). Charges of anthropocentrism or ethnocentrism can often be read, especially by scientists, as a need for greater objectivity, but in humanistic circles they are more often assertions that the *anthropos* and the *ethnos* remain central to scholarly inquiry. Like readiness, context is all.

An awareness of context – that which the model cannot adequately ingest – may be more than just an ideological or disciplinary construction, however. Cognitive scientist Andy Clark, author of *Natural-Born Cyborgs* (2003), details the patterns with which our eyes scan our environments for details. These *saccades*, as they are called, are neither smooth nor linear. "We now begin to understand," writes Clark,

> why the patterns of saccade are not cumulative – why we visit and repeatedly *revisit* the same locations. It is because our brains just don't bother to create rich inner models. Why should they? The world itself is still there, a complex and perfect store of all that data, nicely poised for swift retrieval as and when needed by the simple expedient of visual saccade to a selected location. The kind of knowledge that counts, it begins to seem, is not detailed *knowledge* of what's out there, so much as a broad *idea* of what's out there: one capable of then informing on-the-spot processes of information retrieval and use. (Clark, 2003, p. 66)

> It is the strategy of preferring meta-knowledge over baseline knowledge," says Clark. "Meta-knowledge is knowledge about how to acquire and exploit information, rather than basic knowledge about the world. It is not knowing so much as knowing how to find out. (Clark, 2003, p. 67)

As Ginzburg and Clark both demonstrate, a tension exists between rich fields of data on the one hand and simplified, abstract models on the other. Neither, it seems, can fully encompass the other and there is an ongoing symbiotic relationship between the two. The humanities may exist, then, not *in spite of* that tension, but precisely *because* of it. As Ginzburg (1989) has argued, the humanities are the place where that tension is held in highest suspension, in a kind of unresolved *agon*, in a philosophical aporia from which even methodology cannot satisfyingly extricate us.

This dynamic has significant implications for modelling in the humanities because it begins to render intelligible the vexed position of models within a centuries-long history of humanistic inquiry. At this critical juncture in the history of the humanities, achieving a balance between the "D" and the "H" in "digital humanities" is crucially important, and models may find themselves perched precariously between the two sides. Susan Brown has argued recently, for example, that "[s]ince digital modes of scholarly production are highly experimental, working at the interface of the digital-humanities divide constitutes, in itself, research, *provided that the two aspects of the research seriously engage with one another*" (Brown, 2011, p. 218, emphasis added). In stressing the importance of "'core' humanities research activities," Brown calls for a research agenda "to move digital humanities work forward, to ensure it serves the humanities at large, and to counter the pressure created by larger trends or technical advances" (Brown, 2011, p. 218). Modelling, then, might be a useful addition to the humanist's toolkit. We might quibble about the emphasis to be placed upon it (McCarty does seem to suggest, after all, that it should perhaps be the only game in town), but theorizing models – in particular, models useful for humanists – is one way to investigate what lies at the intersection of the "digital" and the "humanities."

Toward that end, I would like to offer a few suggestions. First, an entire discourse of science studies literature lives and works at precisely the intersection between science and the humanities towards which McCarty has pointed us. Scholars like Thomas Kuhn, Ian Hacking, Donna Haraway, Katherine Hayles, Bruno Latour, and others who are grouped generally under the heading of "science and technology studies" will help to navigate those theoretical shoals. In *How We Became Posthuman,* Katherine Hayles (1999) has written eloquently on the gendered nature of embodiment and the subjective stance of the modeller. In the "Cyborg Manifesto" (1991) and elsewhere, Donna Haraway has explicated the "translation of the world into a problem of coding" via not only digitization technologies but also genetic DNA (Haraway, 1991, p. 164). Bruno Latour's work, like *Science in Action* (1987), for example, has offered the methodological caveat that our current representations of both Nature and Society are the *effects* of technoscientific controversies, not the *causes* of them, and so neither can be used "to explain how and why a controversy has been settled" (Latour, 1987, p. 258). Texts such as these can improve our theorizations about the model-building enterprise upon which we're embarking.

Second, re-reading the theoretical transition from structuralism to post-structuralism has, for us, more than just historic interest. Jacques Derrida's career-long work on the "structurality of structure" (Derrida, 1978, p. 278), for example, suggests ways in which we might begin to re-appreciate the limitations of modelling. Deconstruction has articulated the ways in which our discursive concepts shape the very models we make. The "structurality of structure," Derrida wrote,

> has always been neutralized or reduced . . . by a process of giving it
> a center or of referring it to a point of presence, a fixed origin. The
> function of this center was not only to orient, balance, and organize

Quamen, Harvey. (2012). The Limits of Modelling: Data Culture and the Humanities. *Scholarly and Research Communication, 3*(4): 040148, 14 pp.

11

the structure – one cannot in fact conceive of an unorganized
structure – but above all to make sure that the organizing principle of
the structure would limit what we might call the *play* of the structure
(Derrida, 1978, p. 278)

The role of the centre – the very goal of structuralism itself – is to fix instability, to
locate an essence, to move the model towards an approximating convergence. Brian
Cantwell Smith (1996) argued that all models are partial; Derrida would add that they
are so precisely because they can never fully eliminate an instability that he calls "play."

One implication of this debate between models and data (increasingly, especially in the
digital age, "big data") is, I'd like to suggest, the growing prominence of certain kinds of
humanist scholarship that, to date, have often been too difficult or too time-consuming
to pursue diligently. Two examples might suffice: James Shapiro's *1599: A Year in the Life
of William Shakespeare* (2005) and Robert Darnton's investigation into the publication
of Voltaire's *Questions sur l'Encyclopédie,* first published in 1982 as "What is the History
of Books?" and republished in his recent volume, *The Case for Books* (2009).

Both works engage massive quantities of original archival material in order to draw,
in Shapiro's case, a rich description of Elizabethan England in the year 1599 and,
in Darnton's, a model of eighteenth century authorship in Paris under a regime
of censorship, during which time manuscripts had to be smuggled out of Paris
to Switzerland for publication and then re-smuggled back into Paris to be sold to
interested readers. Shapiro's work encompasses not only political history (England's
ongoing war against Ireland), but also theatre receipts, travelogues, Lancelot Andrewes'
religious sermons, letters, and even contemporaneous weather reports that, for
example, portended an ill omen upon the Earl of Essex's military departure from
London for Ireland on 27 March 1599 (Shapiro, 2005). Darnton, similarly, reconstructs
the dangerous passage of literary materials across an illicit underground network that
makes possible one single, but deceptively unremarkable, event in the life of a Parisian
bookseller: "On August 16, 1770, [Isaac-Pierre] Rigaud ordered thirty copies of the
nine-volume octavo edition of the *Questions,* which the Société typographique de
Neuchâtel (STN) had recently begun to print in the Prussian principality of Neuchâtel
on the Swiss side of the French-Swiss border" (Darnton, 2009, p. 183). The stories
behind the writing, printing, and shipping of these volumes make good reading and
they also suggest ways in which new kinds of database-rich humanities research might
proceed.

Darnton points out one of the problems, however: reading his way through the papers
of the Société typographique de Neuchâtel ("50,000 letters, the only complete archive
of a publishing house from the eighteenth century that has survived" [Darnton, 2009,
p. 62]) and "most of the documents in two of the largest collections at the Bibliothèque
Nationale de France" (Darnton, 2009, p. 62), Darnton had spent "eleven summers
and three winters over a period of twenty-five years" pursuing his research (Darnton,
2009, p. 62). The point should be obvious: these are not the kinds of projects in which
either graduate students or junior scholars can engage. "Not that I suffered," Darnton
confesses, "Neuchâtel is a lovely city in good wine country on the edge of a lake behind
a handsome range of mountains, and Paris is paradise" (Darnton, 2009, p. 62). The

work of reading and scholarship in this scope and breadth should not intimidate us. But the implication that closing off certain kinds of research questions – reserving them for the most senior and privileged scholars who can afford the luxuries of time in, and travel to, good wine country – should probably give us pause. Digital techniques – and, more specifically, a digital, data-centric culture – can open up these kinds of research questions to energetic new scholars who have the smarts and the savvy, but not necessarily the means or the money, to pursue them. Digital infrastructure projects – like creating a database of sixteenth century weather reports, for example – might not catch the eye of funding agencies, but that's one kind of "data accessibility" project that can help to revitalize humanistic inquiry and can lead to the generation of new models that can draw attention back to the humanities.

Bigger ships, please

To bring us back full circle to John Dee: "Now, may you, of any Mould, or Modell of a Ship, make one, of the same Mould (in any assigned proportion) bigger or lesser" (Dee, 1570, p. C3r). Our fascination with modelling begins in a pragmatic recognition that our time, our energy, and our resources are all finite. We scale, we downshift, we ratchet back because the challenges of data management sometimes appear intractable. But no data management structure to this point in history has had the potential to make manageable the vast scales of data that we can wield via the computer. Ours may not be a new utopian age, but it certainly seems that the problems of humanistic scholarship are shifting under our feet. Questions that even a short time ago seemed unanswerable – and, unfortunately, deemed by some probably not even worth asking – now look much different. Our digital ships are bigger. They have larger draughts. They carry more cargo, more data, more "tons burden." The humanist methodology privileged data over the model. The scientific methodology privileged the model over data. The next humanist challenge may be how to sustain those two paradigms in a productive agonistic tension.

References

Barker, Richard. (1988). 'Many may peruse us': Ribbands, moulds and models in the Dockyards. *Revista da Universidade de Coimbra, 34,* 539–559.

Bloom, Harold. (2000). Foreword: Northrop Frye in retrospect. In Frye, N., *Anatomy of criticism: Four essays* (pp. vii-xi). Princeton: Princeton University Press.

Brown, Susan. (2011). Don't mind the gap: Evolving digital modes of scholarly production across the digital-humanities divide. In Coleman & Kamboureli (Eds.), *Retooling the humanities* (pp. 203–231). Edmonton, AB: University of Alberta Press.

Clark, Andy. (2003). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence.* Oxford, UK: Oxford University Press.

Darnton, Robert. (2009). *The case for books: Past, present, and future.* New York, NY: Public Affairs.

Davies, Tony. (2008). *Humanism.* 2nd ed. New York, NY: Routledge.

Dee, John. (1570). Preface. In Sir Henry Billingsley (translated), *The elements of geometrie of the most auncient philosopher Euclide of Megara.* London, UK: John Daye. URL: http://www.gutenberg.org/files/22062/22062-h/main.html [October 15, 2012]

Derrida, Jacques. (1978). Structure, sign and play in the discourse of the human sciences. Alan Bass (translated), In *Writing and Difference* (pp. 278–293). Chicago: University of Chicago Press.

Friel, Ian. (2009). *Elizabethan merchant ships and shipbuilding.* URL: http://www.gresham.ac.uk/print/2420 [July 17, 2012]

Gibson, William. (1984). *Neuromancer.* New York, NY: Ace Books.

Ginzburg, Carlo. (1989). Clues: Roots of an evidential paradigm. In John Tedeschi and Anne C. Tedeschi (translators), *Clues, myths, and the historical method* (pp. 96–125). Baltimore, MD: Johns Hopkins University Press.

Haraway, Donna. (1991). A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In *Simians, cyborgs and women: The reinvention of nature* (pp 149–181). New York, NY: Routledge.

Hayles, N. Katherine. (1999). *How we became posthuman: virtual bodies in cybernetics, literature, and informatics.* Chicago, IL: University of Chicago Press.

Kuhn. Thomas S. (1996). *The structure of scientific revolutions.* Chicago, IL: University of Chicago Press.

Latour, Bruno. (1987). *Science in action: How to follow scientists and engineers through society.* Cambridge, MA: Harvard University Press.

Latour, Bruno. (1993). *We have never been modern.* Cambridge, MA: Harvard University Press.

Mahoney, Michael S. (2002). Software as science – Science as software. In Ulf Hashhagen, Reinhard Keil-Slawik & Arthur L. Norberg (Eds.) *History of computing: Software issues. International conference on the history of computing ICHC 2000, April 5–7, 2000*, Heinz Nixdorf MuseumsForum, Paderborn, Germany, pp. 25–48, Berlin, Germany: Springer Verlag.

McCarty, Willard. (2005). *Humanities computing.* London, UK: Palgrave.

Oxford English Dictionary. 2011. Online Edition. Oxford University Press. URL: http://www.oed.com.login.ezproxy.library.ualberta.ca/view/Entry/120577?rskey=zGF1JB&result=1&isAdvanced=false [October 28, 2011].

Ruskin, John. (n.d.). *Modern painters.* 2nd ed. (Vols 1–5), Boston, MA: Dana Estes & Co. (Original work published in 1856).

Shapiro, James. (2005). *1599: A Year in the Life of William Shakespeare.* London, UK: Faber and Faber.

Shelley, Mary. (1999). *Frankenstein; or, the modern prometheus.* (D.L. Macdonald and Kathleen Scherf, Eds.). 2nd ed. Peterborough, ON: Broadview. (Original work published 1818).

Smith, Brian Cantwell. (1996). Limits of correctness in computers. In Rob Kling (Ed.), *Computerization and Controversy.* 2nd ed.(pp. 810–825). New York, NY: Academic Press.

Quamen, Harvey. (2012). The Limits of Modelling: Data Culture and the Humanities. *Scholarly and Research Communication, 3*(4): 040148, 14 pp.

14