Scholarship in interaction: case studies at the intersection of codework and textual scholarship

Zundert, J.J. van

Chapter 2

Screwmeneutics and Hermenumericals: The Computationality of Hermeneutics[1]

> "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."
> (Drucker 2012)

Can the computer be a hermeneutical instrument? This question is trivial, for obviously the computer can be. As long as there is a human interpreter any object can be interpreted and can therefore be an instrument of hermeneutical activity. So the question is not if, but how. How can the computer be applied as a hermeneutical instrument of humanities? That question is less trivial, but passes over a number of important precursory questions. First of all: Must the computer be a hermeneutical instrument to humanities? Which again leads to the question: What is the role of hermeneutics in humanities? If we can – at least tentatively – answer that last question, we may progress to evaluate whether digital humanities can and must have a hermeneutics. And if so, the question becomes, how?

[1] An earlier version of this chapter appeared as Van Zundert, Joris J. 2016. "Screwmeneutics and Hermenumericals: The Computationality of Hermeneutics." In *A New Companion to Digital Humanities*, edited by Susan Scheibman, Ray Siemens, and John Unsworth, 331–347. Malden, Oxford, etc.: John Wiley & Sons, Ltd. http://onlinelibrary.wiley.com/doi/10.1002/9781118680605.ch23/summary.

2.1 On Hermeneutics

Hermeneutics is the theory of text interpretation. The very root (Greek ἑρμηνεύω, hermeneuō) means to interpret or to translate. According to folk etymology its origin derives from Hermes, the Greek god-messenger. It is in the nature of Hermes not just to use language as a means of communication, but also to be a corrupter of words, relishing in the confusing power of his messages. He is a god of transitions and boundaries. An apt eponym for hermeneutics, thus – interpretation is the transition of knowledge that happens on the boundary between text and reader. Hermeneutics is already referred to in various ways by classic philosophers, but it is Philo of Alexandria who pulls together a first systematic theory which is aimed at uncovering the deeper allegorical meaning of sacred scripture (Ramberg and Gjesdal 2013). Methodologically connected to the pivotal issue of interpreting the texts of the Bible, hermeneutics plays a central role throughout the history of philosophy, humanistic theory, philology, and literary criticism. There are numerous key works in the development of hermeneutic thinking, but one that should in any case be mentioned is *De Doctrina Christiana* of St. Augustine of Hippo (c.400 CE). In his work Augustine unfolds a methodology to interpret the scriptures. But more importantly, in his methodology he connects semiotics – the theory of signs and symbols – to language, and he connects the interpretation of language to a deeper existential meaning (cf. Green 2008). In his theory words are signs that impart cognitive concepts to an interpreter. Just as a natural sign such as smoke signals "fire" to the interpreter, so do words convey meaning as "given" signs of language. The problem is however that this meaning may be literal or metaphorical. The sun may stand for light of day or for light of vision. The existential aspect is raised when Augustine argues that it is the will and intention of the reader that allows her to address the deeper allegorical interpretation.

From Augustine we take a huge leap through humanistic history and we pass over Thomas Aquinas, Dante, Petrarch, Luther, Spinoza, and many other philosophers and scholars whose names and works stand witness to the profound influence of Augustine's thinking, and of the central role of hermeneutics in the humanistic disciplines (Barolini 2007; Marchesi 2011; Ramberg and Gjesdal 2013). We turn to the early nineteenth century

and Friedrich Schleiermacher's contribution to hermeneutic methodology. Schleiermacher points to an important aspect of interpretation, which is that it is in part emphatic in nature. A reader is able to understand a text not just because of a linguistic code shared with the author, but essentially also by sharing a human nature. Thus, a part of the interpretation and part of the meaning of a text is not based on what is in the text, but on what is external to the text. Following, broadening, and formalizing Schleiermacher's work, Wilhelm Dilthey theorized that works are constructed from the vantage point of a particular worldview held by an author. The interpretation and understanding of a text therefore involves relating the text to the biographic and historical circumstances of its author. For both Dilthey and Schleiermacher a basic assumption is that the meaning of texts is grounded in the intentions and histories of their authors (Mallery, Hurwitz, and Duffy 1986). But more importantly, they believed that these intentions were knowable to later interpreters through reconstruction. Dilthey however recognized that this reconstruction would be tainted by the interpreter's present worldview. Interpretation therefore could in his opinion not be objective in a scientific sense of establishing facts empirically. But he argued that aggregation of multiple interpretations could lead to valid and more generalized interpretations.

Both Schleiermacher and Dilthey point us to the fact that any interpretation necessarily involves information that is not in the data itself. This may be information that is available elsewhere in the form of other explicit data, texts, and so on. But interpretation also involves the unique cognition of the interpreter, which is tacit. Acknowledging the partly tacit nature of interpretation sets hermeneutics apart from other frames of interpretation such as the probabilistic model of information theory inspired by Claude Shannon.

Around the time of Dilthey's life and work hermeneutics was still tightly connected to philology, which at the start of the twentieth century was very much geared towards establishing texts according to what was perceived as the intent and ideal of the original author. With Heidegger, that was about to change. For the philosopher Martin Heidegger, the hermeneutic process is not a philological tool. Instead hermeneutics scales to an ontological level and becomes philosophical in nature. Interpretation and understanding pertain to all of us as the interplay between our self-understanding and our

understanding of the world (Ramberg and Gjesdal 2013). Heidegger holds that interpretation and understanding are to a great extent intuitive operations. Our understanding of the world is largely an immediate and unreflective grasp of what we sense, based on a priori knowledge accumulated from experience. Heidegger believed that this understanding is uniquely subjective. We can only "read ourselves" into a text. A few decades later, philosopher Hans-Georg Gadamer would be less pessimistic and would suggest that a human can transcend his own horizon by being exposed to the discourse and linguistics of others. Later again, Jürgen Habermas and Karl-Otto Apel added pragmatics into the equation – that is, a theory of interpretation and understanding must also take into account the intentionality of linguistics.

Hermeneutics, then, turned from a theory of the interpretation of text into an ontological theory of understanding. It can now be understood broadly as the theory of the processes that turn information into knowledge. As such, the role of hermeneutics in humanities cannot be overestimated: humanities practice is primarily hermeneutic, its main theoretical frame is hermeneutics. Consequently, the way that hermeneutics developed over time has significant ramifications for the epistemology of humanities. According to Chambers (2000) humanities "is hermeneutic, intertextual, participatory, value-laden, context dependent, and relatively indeterminate; there are no hierarchical structures of information, no obvious causal explanations and no undisputable truths of any significance to be found."

The highly relativistic nature of post-structuralist hermeneutics problematizes factuality as veritably factitious. This poses problems for those realms of humanities that are concerned with establishing the concrete humanistic record – for instance in the case of philology, ironically a humanistic pursuit most intimately connected to hermeneutics. Jerome McGann rejects the post-structuralist project of, inter alia, Lyotard and Derrida, informed by Heidegger's philosophy, to replace "traditional science with a science of the unknown" (McGann 2013). McGann reasons that philosophy is rather a subroutine of philology concerned with testing, reconstructing, or falsifying its subjects of attention. The primary concern of philology then is with establishing the archive of what is known or has been known: "Philology is the fundamental science of human memory". McGann (2013:345–346)

reduces the impact of post-structuralist hermeneutics to an "after the fact" reinterpretation of established sources: "For the philologian, materials are preserved because their simple existence testifies that they once had value, though what that was we can never know completely or even, perhaps, at all. If our current interests supply them with certain kinds of value, these are but Derridean supplements added for ourselves".

Philology cannot however escape problematic hermeneutics by simply stating that its aim is a factual archive. More often than not, for instance when difficult script is encountered, interpreting medieval manuscript is nontrivial. Thus, even if it poses as merely recording the words glyph by glyph, textual editing involves interpretation. Moreover, a philologist editing a historical text cannot escape actualization without betraying the pragmatics of philology that presupposes making the archive intelligible for a current audience too. A gloss is instrumental in this translation, but therefore also not ahistorical. Any "ahistoric" presupposition of philological hermeneutics is negated by historicality: "Not only is the decision for one possible correction rather than another already interpretation, but the question of which possibilities of correction occur to the philologist, and which don't, also depends upon his own historical horizon. [...] The intention toward the historical meaning changes with changes in the conception of history. [...] Once it has become doubtful that one can experience how it really was, then it is no less doubtful that one is in the position to establish how something was meant once" (Szondi and Bahti 1978).

This severely upsets traditional philology, which "believes itself to be independent of its own historical point of view." Fiormonte and Pusceddu (2006) problematized in a similar vain the temporal dimension of text, arguing that genetic editions also cannot escape fundamental subjectivity: "one might say that up to now we have analyzed the literary text according to the laws of the pre-Heisenbergian universe, i.e. inside a stable system in which the observer does not modify the object observed." Mutatis mutandis this "hermeneutic condition" can be generalized to many subfields of humanities. The study of history, for example, being dependent on a humanistic record as well, is affected similarly.

2.2 The Hermeneutics of Digital Humanities

Does digital humanities have a hermeneutics like humanities does? Given that digital humanities is humanities too, the answer must be yes. However, there seems to be no focused program to uncover the hermeneutics of digital humanities. I want to investigate whether a call for attention to this hermeneutics, if not a specific program, is a necessity for digital humanities. Rafael Capurro (2010) seems to have come closest to calling for a programmatic approach to digital hermeneutics. Capurro states that the Internet challenges hermeneutics because of its social relevance for the creation, communication, and interpretation of knowledge. That is, the Internet makes the creation and sharing of knowledge a more open and social activity. A problem in addressing this challenge is that the last part of the twentieth century saw a pseudo-critical rejection of hermeneutics with regard to technology in general and to digital technology in particular. But it is exactly digital technology, and more particularly the Internet, that has ontological implications or implications for how we are and behave as humans: the Internet shapes important parts of human expression and experience, and conversely humans shape the Internet as a technology by expressing themselves through it. According to Capurro, a resulting problem is that humans only very partially control the network that they shape but that is importantly shaping them. A counterargument could be that individual humans also only very partially control their physical environment, and that moreover the power of control is unevenly distributed in the virtual as well as in the physical environment. However, Capurro's more important point is that the network is shaping us in more fundamental ways than we may realize. Our lives are increasingly expressed through digital technologies that function as extensions of our minds and bodies: we are different on Facebook, and Facebook makes us different in real life too. This raises questions of a particularly humanistic nature, and Capurro concludes that current hermeneutics fails to address these questions that "go far beyond the horizon of classic hermeneutics as a theory of text interpretation as well as beyond classic philosophic hermeneutics."

If current hermeneutics is unable to address such questions, this may explain the relative lack of theory on hermeneutics we find in digital humanities.

The dialog surrounding hermeneutics seems not to have developed fully yet in digital humanities – references to hermeneutics are scant and often at a concrete level of the practice of text interpretation, such as when Katherine Hayles (2012) uses the phrase "hermeneutic close reading." Yet from several paragraphs and sections in the literature the emergence of a debate seems traceable. Like Capurro, Fred Gibbs and Trevor Owens (2012) have made programmatic claims for a hermeneutics of history writing. Their argument concentrates on data. Data has always been used by historians, but the vast quantities of it that become available should mean "that 'using' signifies a much broader range of activities." Gibbs and Owens argue that using data is not the same as fully conforming to the epistemic burden of the statistician. A playful iterative approach to quantitative tools, explorative and deliberately without the complete formal mathematical rigor, can serve to use large amounts of data to discover and frame research questions. Data does not always have to be used as evidence; in a variety of forms it can provoke new questions and explorations. Data analytics need not be by definition mathematical. "Historians must treat data as text, which needs to be approached from multiple points of view and as openly as possible" (Gibbs and Owens 2012).

Like many contributions in the theory on digital humanities, the article by Gibbs and Owens refers to the opposition between quantification and narrative as methodological means. They write about the "epistemological jitter" and "hostility to data" on the side of historians. Their solution to this conundrum is that data can be read as text. This is true, but it is also an unsatisfactory and incomplete solution to the problem. The presupposition of data-as-text reduces the hermeneutical act to a post-processing of what remains of data after the processes of curation, analysis, and visualization. However, those processes of curation, analysis, and visualization have a hermeneutics of their own. The dialog on the hermeneutics of digital humanities cannot therefore simply posit a dichotomy between the quantitative and qualitative, and relegate hermeneutics to a qualitative aspect of interpretation of given data as if this data would not be value-laden and interpreted already. It is along these lines too that Federica Frabetti – like Capurro – has argued that new technologies affect cultural understanding. She proposes a re-conceptualization of digital humanities that indeed transcends an

assumed dichotomy between the technical and the cultural aspects. Such "must be pursued through a close, even intimate, engagement with digitality and with software itself" (Frabetti 2012). Thus part of the hermeneutics of digital humanities relates to the hermeneutics of code, computation, and quantification.

A close or even intimate engagement between digitality and hermeneutics has however not been a main concern of digital humanities. Rather, the opposite has been stressed. In a 1995 issue of Literary and Linguistic Computing Lisa Lena Opas-Hänninen writes: "Only where indexing and sampling are concerned does the computer offer useful help in computer-assisted literary studies. So the impact of computer-assisted techniques sets in before the interpretation and evaluation of the text begins." Looking back, Opas-Hänninen's introduction reads like a very careful attempt to avoid stating that computational analysis in the realm of literary studies can go beyond anything but a pre-hermeneutical support tool. Jan Christoph Meister in the same volume – carefully? – formulates that "an intelligent and well-balanced application of literary computing tools allows us to reconcile the two paradigms by measuring and mapping difference in literary structures, and then forwarding them to the ultimate hermeneutic machine, the human mind" (Meister 1995). Both Opas-Hänninen and Meister at the time argued that the hermeneutical potential of digital technology is limited by the fact that "only questions that can be formalized are open to electronic analysis in literary studies and this is why computer-assisted techniques can cover only part of the work of the literary critic in certain, clearly defined areas of application" (Opas-Hänninen 1995). Meister draws a very strong opposition between the numerical and semantic paradigms: the first is connected to computing, the second to hermeneutics. His argument is basically that semantics do not apply in the computational paradigm. Algorithms can manipulate or process objects, but only insofar as they can be formalized and quantified. Computational operations are strictly and unambiguously transformative: "results are effectively nothing but a more or less sophisticated re-formulation [...] of the original data input." But these transparent repetitions and permutations of data are redundant in the semantic paradigm. When it comes to hermeneutics, "only those results that are different, that happen to question the validity or confinements of the procedures which

produced them, will ultimately be found to be relevant and noteworthy." In retrospect it is intriguing that Meister did not consider at the time a distinctly hermeneutic consequence of this argument. Algorithmic transformations can in fact lead to identification of results that are different, and thus not "hermeneutically" neutral. Firstly, even a rudimentary indexing algorithm can transform the full text of a book into a list of terms used more than average per chapter, and can subsequently single out the chapter that shows the least terminological overlap. Is this not a hermeneutics expressed through the algorithm? Secondly, we can consider the breakdown of software. As long as the algorithm only transforms data, it may not be a hermeneutical thing. But it may become so when it falters over some input and breaks down or spews inconsistent and unexpected results. This is akin to what we find in Heidegger's work, which holds that only a breakdown in practice leads to theoretical knowledge (Froesse 2006). As long as a hammer is a hammer, it is a hammer; only when it is broken do we consider its function and how it works.

Twenty years on, the consideration of hermeneutics in the digital humanities does not seem to have moved beyond a basic opposition between patterns and narrative, quantification versus interpretation, that can already be discerned in Meister's strong binary opposition between the numerical and the semantic. This opposition often surfaces as an apparent ideological or political opposition between humanities and digital humanities. Stanley Fish for example has qualified digital humanities as just another fad answering to a crisis of legitimization of the humanities (Fish 2010; 2011). Others point to the ideologies and institutional motivations of innovation, which certainly are not neutral (e.g. Piersma and Ribbens 2013). These crises or ideologies, even if they exist, do not relieve us from critically evaluating the ramifications of emerging digital technologies for hermeneutics. In the first place, these technologies are increasingly used to create the humanistic artifacts that are the objects of study in the humanities. In the second place, we are applying these technologies for the capture and analysis of research data. Both of these processes, motivated by digital technologies, affect our modes of interpretation in nontrivial ways. Piersma and Ribbens argue that evaluation of these digital technologies is "even more urgent in view of the frequently implicit claims [...] that technological progress also implies a new historical-scientific

paradigm" – a paradigm based on quantified approaches, on computational analysis of big data, and subsequent serendipitous finds in such big data.

From the perspective of hermeneutics, however, the literature in digital humanities does not seem to justify presupposing an implicit turn to a scientific paradigm. Geoffrey Rockwell (2003), writing on the hermeneutics of text analysis, refers to the French eighteenth-century philosopher Étienne Bonnot de Condillac: analysis merely consists of composing and decomposing our ideas to create new combinations and to discover, by this means, their mutual relations and the new ideas they can produce. Rockwell argues that there is no a priori privilege of any procedure for deconstruction and reconfiguration. But a potential a priori for coherence and homogeneity in computational data analysis may have been inadvertently introduced at the very onset of automated text analysis, which is tied to the computationally constructed concordance by Roberto Busa that was commenced in the late 1940s. Concordancing aims to discover patterns of coherence in a text or corpus – in a hermeneutically naive way because it assumes that a word will have the same meaning and weight wherever it occurs. Yet even the algorithmic creation of concordances shows how deconstruction of a text and subsequent reconfiguration leads to a new text, namely the very concordance. But that is just one method of reconfiguration. To escape naive biases we should shed habitual practices and any axiomatic primacy of unity and coherence. To this end Rockwell – following Gadamer and Huizinga – suggests a hermeneutics of disciplined play that privileges experimentation and modeling, rather than a narrow quantified empirics.

Stephen Ramsay (2011c), even more than Rockwell, emphatically denies a scientific paradigm for hermeneutics: "For decades the dominant assumption within humanities computing [...] has been that if the computer is to be useful to the humanist, its efficacy must necessarily lie in the aptness of the scientific metaphor for humanistic study." Ramsay takes the contrary view, and proposes that the scientific method and metaphor are, for the most part, incompatible with the terms of humanistic endeavor and only lead to a distorted epistemology called "scientism." Ramsay follows Gadamer by stating that the hermeneutic phenomenon is basically not a problem of method at all. Hermeneutics is simply not concerned with amassing verified knowledge of the sort that would satisfy the methodological idea of sci-

ence. Rather, literary criticism operates within a hermeneutical framework in which the specifically scientific meaning of fact, metric, verification, and evidence do not apply. Yet humanities too is concerned with knowledge and with truth, just of a different kind than that of science. Ramsay has also argued that the availability of vastly more digital data essentially does not change the hermeneutic assumptions of humanities. The fact is that there has always been too much information available to synthesize individually in full; the digital age just makes this condition more apparent. But now as ever hermeneutics involves finding a purposely selective and subjective path through too much information. This is the basic assumption underlying what Ramsay (2010) calls the hermeneutics of screwing around. For Ramsay the "screwmeneutical imperative" is nothing more or less than the realization of Roland Barthes' concept of "writerly text", which is the text a reader constructs by reducing all possible meanings of a text to one that is his own interpretation of it.

In the realm of markup, in the digital humanities predominantly represented by the Text Encoding Initiative (TEI), possible scientism seems not to be a very relevant issue either. This may be partly due to the descriptive rather than analytic nature of markup. The hermeneutic dialog within this domain concerns itself more with the issue of multi-perspectivity. Like Ramsay, Lou Burnard (1998) points to post-structuralist ideas: "Texts, and other artifacts alike, are invested with meaning by our use of them, and it is therefore interpretation alone which confers value on them. Small wonder that Derrida, citing Montaigne, takes it as self-evident that 'We need to interpret interpretations more than to interpret things' " Authorial intention, reconstruction, and original reading are concepts that have become unfashionable, Burnard admits, but he follows Dilthey by saying that there "is ample evidence that not all interpretations are equally useful or have equal explanatory force." He suggests that canonicity is in this sense a hermeneutics of aggregation. Burnard also embraces the post-structuralist idea of intertextuality: the reading and the meaning of a text is in part constructed by the references made to other texts. The rationale for markup then is that it provides a single formalized semiotic system that is able to function as an interlingua for the sharing of the multitude of individual interpretations that through aggregation can lead to a critical consensus.

The claim that a single all-encompassing semiotic system is possible, and that technologies such as SGML/XML and DTDs could be an implementation of it, has since been severely contested. Many theorists and practitioners (e.g. Buzzetti and McGann 2006; Fiormonte and Pusceddu 2006; Schmidt and Colomb 2009) find that the single-hierarchy approach to text structuring that the TEI enforces does not fit well with a multitude of possible structural and semantic interpretations. In itself this dialog testifies to the fact that the approach to text encoding within the textual scholarship and digital humanities communities is primarily hermeneutically oriented.

Thus a computational turn does not automatically imply a turn to empiricism and scientism, or a disregard for hermeneutic tradition. Stylometry and the "school" of distant reading (Moretti 2007; Jockers 2013) may lean in their approaches more towards an empiricist or scientistic attitude. This is mainly to be attributed to the intensive use of quantification and – more importantly – statistics in those avenues of research. The work of researchers such as David Hoover (2013), Ted Underwood (2019), Allen Riddell and Karina van Dalen-Oskam (2018), Matthew Jockers, and Franco Moretti is methodologically strongly based in statistics, corpus linguistics, and natural language processing. Those methodologies are numerically inclined indeed, but this does not preclude hermeneutics – numbers of course allow interpretation too. Quantification does however introduce the problem of reduction. Current statistic approaches to stylometry, for example, are based predominantly on word frequencies and co-occurrence analyses of the surface structures of text. But aggregating words based solely on their form usually blinds these methods to more subtle semantic relations such as homonymy, metaphors, anaphors, and so on, that are also hermeneutically important. This does not however discredit numerical approaches as hermeneutical instruments. In fact they may contribute very strong hermeneutical support.

For instance, Mike Kestemont (2012) has shown, using statistical means such as principal component analysis, that the medieval Dutch Arthurian novel *Moriaen* stylistically stands out from the medieval compilation of Arthurian texts that it is a member of. The text forms a much closer stylistic unit with two other texts, one of which is not even an Arthurian novel but a story in the realm of the so-called matière de France pertaining to the culture, court,

and principal personae during the reign of Charlemagne. Based on all we know about medieval Dutch genre and literary history this claim would be outrageous, were it not for a 1970s posthumously published work by a Dutch philologist that had already alluded to these possibilities. Most interestingly, that philologist and poet, Klaas Heeroma, based his conjecture on a fundamentally hermeneutic principle: he claimed he "heard" the kinship between the novels. Somewhat ironically, what is now often perceived of as one of the least hermeneutical instruments – number-crunching-based principle component analysis – indicates that Heeroma's hermeneutical "sixth sense" was right.

The example above draws our attention to another problem inherent in current quantified approaches in digital humanities. As Gibbs and Owens (2012) also point out, neighbor joining, maximum-parsimony phylogenetic trees, z-scores and such probabilistic methods that are used in stemma reconstruction, authorship attribution, and various other computational approaches seem foremost to be used as instruments of reaffirmation. They verify authorship, and they confirm canonicity and genre. They do not answer new questions, but rather solidify existing answers. This may very well be a simple sign of a field in development, of relatively immature application. However, if this confirmation bias were a genuine trait of a specific angle on quantified approaches by digital humanities, then again this would set it apart from the scientific paradigm of falsification rather than import it wholesale into the humanities. So far, quantified approaches in the digital humanities also show a relative lack of explanatory power. Stylometry, for example, can tell us – or rather indicate to us – that there are two authors of a certain text (Van Dalen-Oskam and Van Zundert 2007). But it tells us unsatisfactorily little about how and why the individual styles differ. Engaging and uncovering the "black box" effect of such methods could in due course turn the practice of stylometry into the pursuit of a literary hermeneutics – like conventional hermeneutics but with different means.

2.3 The Computationality of Hermeneutics

Quantified approaches and distant reading currently have good press. But we should be careful not to identify digital humanities solely with these approaches. The field is decidedly broader (cf. for instance Alvarado 2012). There is a tendency in debates to reduce the potential of computation to a methodology of quantification. The nature of digital humanities is hybrid, however, and there is not an a priori discontinuity with the hermeneutic traditions. We still maintain that knowledge has an interpretative character – that the state of an object is determined by its context and is dependent on the observer's interpretation. Computer-mediated text turned text into something computationally tractable. Starting with the work of Father Busa, this made the application of quantified approaches to text feasible and practical. Computational tractability, however, does not dictate quantification and a probabilistic approach. These approaches have been inspired by their success in computational linguistics, a field informed substantially by a positivist and structuralist tradition. This tradition holds that knowledge has a causal deterministic character so that the state of any given object is necessarily determined by its prior states. Probabilistics and, for instance, the Markov models that underpin many natural language processing algorithms derive ultimately from such a positivist deterministic philosophy (cf. Vandoulakis 2011). Johanna Drucker unequivocally denied the applicability of deterministic computational methods to problems of humanistic nature: "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" (Drucker 2012).

Drucker's quote summarizes quite succinctly the problems inherent in probabilistic approach that can only lead to "naive empiricism" (Drucker 2010). Grounding the bulk of digital humanities methodology in quantification and deterministic reasoning may have far-reaching disruptive implications. Katherine Hayles pointed out that digital humanities as a field may converge towards traditional humanities or diverge from it as its own field, depending on how digital humanities articulates itself with respect to conventional hu-

manities. "The kinds of articulation that emerge have strong implications for the future: will the digital humanities become a separate field whose interests are increasingly remote from the traditional humanities, or will it on the contrary become so deeply entwined with questions of hermeneutic interpretation that no self-respecting traditional scholar could remain ignorant of its results?" (Hayles 2012) Thus Hayles ties a successful interaction of digital humanities with the traditional humanities to the question of how well digital humanities will be able to cater to hermeneutics. The extent to which the hermeneutic approach is fundamental to the humanities is, however, not always well understood. In his recent history of the humanities, Rens Bod dedicates a mere two pages to the concept and history of hermeneutics, in a section titled "Hermeneutics and the anticipatory 'method' " (Bod 2013a:333–334). He disposes of the "method" as being based on guesswork and premonitions. This dismissal might be cast aside as anecdotal were it not for Bod's position as professor of computational and digital humanities, investigating the humanities from both a computational and a historical perspective. Within the dichotomy between patterns and narrative, Bod has decidedly opted for patterns as a primary principle of investigation. Leaning strongly towards a deterministic paradigm, he concludes that inferences can only be valid based on patterns to be discovered in the researched data. Another example of a dialog between the realms of computation and humanities reveals an interesting "computational" perspective on the fundamental importance of the concept of context to hermeneutics: "We do not exclude the possibility that there may be other relationships that can constitute a valid narrative. [...] However, such examples are context-dependent, and not easily generalizable, we therefore [...] limit our focus to the prototypical narrative structures described" (Akker et al. 2011). This quote derives from a project of which the particular aim was to find a suitable formalization for (historical) events and to build narratives – i.e. historical accounts – from these. The statement reveals the clear tension between hermeneutic context-dependency and the thrust towards the generalization needed for computational tractability. The generalization requires events to be formalized or modeled so they can be computationally traced and quantified. Researchers try to escape the problematic hermeneutics by reducing the number and type of relations that events can maintain. But the problem stubbornly persists, because formalizations and patterns are not hermeneutics-free. Just as philo-

logical practice cannot escape a certain hermeneutics, neither modeling nor quantification can escape the hermeneutics involved in choosing the basic assumptions onto which the formalizations are founded. Pasanek and Sculley, in their article on "Mining millions of metaphors" (2008), point out that in this respect there is no such a thing as a free lunch: "It is important to avoid the illusion that automated analysis is somehow more objective or less biased than traditional methods. There is no new infallible science of literature forthcoming. As the "No Free Lunch" Theorem states, every machine learning method requires the acceptance of base level assumptions, such as the appropriate choice of distance metric or the shape of the probability distribution underlying the data. These assumptions must, at some level, be taken on faith, and influence the results of automated analysis, just as cultural and theoretical biases influence traditional analysis."

As with quantified approaches, there is a hermeneutics to any formalization. Textual scholars from Bernard Cerquiglini (1999) to Peter Shillingsburg (2013) hold that an edition of a text is not that text itself, but an intellectual argument about it. A digital edition is an interpretation, and in exactly the same sense formalizations and models are interpretations. A simple example for this is a database field, which is nothing more or less than a category label. Category labels, databases, and data models: all are models, necessarily narrow representations of aspects of reality. Confronting any database with reality, one will encounter observations that will not fit to any of the defined database fields. Therefore most data models exclude certain properties of data, which poses problems in a field such as humanities that works primarily with highly complex, heterogeneous, and non-concrete data. To fit the observations to the chosen categories or properties of the model is to fit a subjectively observed reality to the interpretation expressed by the model. The effect is that the chosen formalization imposes a particular interpretation on a set of data that does not really fit, reducing to a certain extent the richness and complexity of the body of information. The quantitative model or data model is an impressionistic primer onto which more interpretation is painted. Thus statistics and models inform interpretative narrative on the basis of formalized reductive interpretations.

If formalizations, models, and quantifications have hermeneutics too, we can concur with Katherine Hayles (2012) when she states that the tension

between algorithmic analysis and hermeneutic close reading should not be overstated. Hayles argues that often there is not an opposition but a synergetic interaction between algorithmic analysis and close interpretative reading. She points to the example of what Matthew Kirschenbaum has called "rapid shuttling," which involves a repetitive switching between the modes of close reading and of interpretation of big data analysis results, comparing the interpretations those different modes yield. Ramsay (2011c), when talking about "algorithmic criticism," also points to this recursive interaction between corpus analytics and close reading that can inform humanistic inquiry of texts. What these views share is that the act of interpretation is postponed to a post-algorithmic phase. Only when the computation has been done and the algorithms and number crunching produce visualizations does the interpretative act come to the fore. This type of digital humanities hermeneutics therefore faces outward and away from the computational model, the math, and the code. It interprets only the results of the algorithmic or quantitative phase. But if it is true that algorithms and models have hermeneutics too, then should these not somehow be taken into account in establishing the validity of interpretations done in algorithmic analysis?

David Berry (2012), like Katherine Hayles, does not "want to overplay the distinction between pattern and narrative as differing modes of analysis. Indeed, patterns implicitly require narrative in order to be understood, and it can be argued that code itself consists of a narrative form that allows databases, collections and archives to function at all." Instead of dismissing code and algorithm as hermeneutic domains, Berry is arguing for a more intertwined articulation of humanities and computer science in this respect. He proposes that digital humanities in part should also concentrate on the underlying computationality of the forms held within a computational medium. To "understand the contemporary born-digital culture and the everyday practices that populate it [...] we need a corresponding focus on the computer code that is entangled with all aspects of our lives." According to Berry there is an "undeniable" cultural dimension to computation as well, which points to the importance of engaging with and understanding code: "Understanding digital humanities is in some sense then understanding code." Berry argues that computational techniques are not merely an instrument wielded by traditional methods. Rather, they have profound

effects on all aspects of the disciplines because the computational logic is entangled with the digital representations of physical objects, texts and "born-digital" artifacts. But the way in which the digital archive is deeply computational and the ramifications of that computationality are currently not well understood, and cannot be understood without a deep dialog between humanities and computer science. Federica Frabetti (2012), reasoning along similar lines, concludes that such "an understanding must be pursued through a close, even intimate, engagement with digitality and with software itself" – which is not without problems, as digital humanities and computer science have no readily available mutually informed way of examining software, and because it is: "especially difficult for those not active in the field of the digital humanities to see how the creation of digital surrogates of analog materials, the development of tools to support visualization and analysis, and the contribution of high-end computing skills [...] constitute research" (Schreibman, Mandell, and Olsen 2011).

In the domain of textual scholarship Elena Pierazzo has drawn attention to a similar need to understand coding intimately. Like others, she holds that editing a text is "interpretative and irreversible." She follows Claus Huitfeldt and Michael Sperberg-McQueen in stating that a transcription of a text consists of "a systematic program of selective alteration." Thus it is very unlikely that two scholars, even given the same transcriptional criteria, will produce the same transcription of the same exemplar (Pierazzo 2011). As scholarly editing moves into a digital environment, computational approaches and programming acquire substantial roles and responsibilities in the creation of digital scholarly editions. Pierazzo therefore argues that this role of programming should not be underestimated, and, more importantly, "neither [should] its implicit scholarly content." Coming from a different angle but reaching a similar conclusion, Alan Galey and Stan Ruecker (2010) call attention to the design of artifacts as a critical and hermeneutical act. They argue that digital humanities must not lose sight of design as an act that shapes the meanings of artifacts, and that is no less vital to the interpretative potential of digital artifacts. Galey and Ruecker draw an analogy between software design and the textual and material design involved in book production: "By understanding how fields like book history take the design decisions embedded in physical artifacts as interpretive objects, we can begin to see digital hu-

manists' creation of new digital artifacts as interpretive acts." Digital humanities as yet lacks a deep understanding of digital text production and software design, whereas we have a well-defined understanding of the roles of non-authorial agents in print and manuscript book production, such as scribes, binders, typographers, compositors, correctors, and illustrators (Galey and Ruecker 2010).

The choices and methods involved in software design do shape the hermeneutics of digital humanities. Modeling encompasses the worldview of the model designer, her context, and her subjective decisions. Data models are anything but neutral – on the contrary, they are a purposefully specific selection of semantic categories and properties. Programming languages have paradigms that affect hermeneutics. Moreover, the reciprocal shaping of the hermeneutics of digital humanities by the methods of computer science extends beyond software design. The choices made in the analytical conception of any given digital humanities project affect its hermeneutic makeup. The choices of what properties to quantify, what probability distribution functions are chosen, which statistical tests are used, are in essence hermeneutically informed. Arguably these choices are currently in large part left implicitly to the experts and professionals of software design and computer science. Computer science as a field, however, is grounded not in a problematizing paradigm but in a problem-solving one. Computer scientists and software engineers have a strong generalizing proclivity. Their reasoning tends toward the inductive: solve a specific problem in a specific context and then scale the solution to general applicability. This propensity invites positivistic reasoning and reductive determinism that favors patterns and relegates the exception to the status of "corner case." These characteristics fit poorly with a humanities that is accustomed to reasoning from heterogeneous information, that favors multi-perspectivity, and that problematizes as a means to create knowledge, perspectives, and understanding. The eventual articulation of digital humanities with respect to conventional humanities – and the implications for the future that Katherine Hayles described – will depend to a great extent on how well the intimate dialog between humanities and computer science as discussed by Frabetti and Berry is established. As Galey and Ruecker showed, little attention is currently paid to the hermeneutical implications of the software

design aspect. Similarly little attention is given to the hermeneutical implications of data modeling and of analytical models applied in computer science and other fields that inform the digital humanities, such as mathematics and artificial intelligence. Thus at a very fundamental level and in a substantial part of its research chain the hermeneutics of digital humanities is driven by software designers and computer scientists. This means that in practice the hermeneutic choices of digital humanities are made substantially by software designers and computer scientists. Failure to reflect critically on these choices may all too easily lead to a naive scientism permeating the digital humanities, born from the generalizing and problem-solving nature of computer science and software engineering.

Stephen Ramsay (2011c) argues that it is possible to make algorithmic procedures conform to the hermeneutical methodology of humanistic critical inquiry without transforming the nature of computation. Be that as it may, this conformity will not come about without a fundamental dialog between humanities and computer science – a dialog that is not part of Ramsay's hermeneutics for digital humanities, focused as this is on post-algorithmic acts of interpretation. However, a substantial part of the specific nature of digital humanities hermeneutics arises exactly from the nature of computation. This nature need not be reductive, deterministic, absolute, and quantified, as is so often implied. Rather, we have here a rationale for exploring "hermenumericals," a hermeneutics of computation that could complement Ramsay's post-algorithmic "screwmeneutics." Computation need not be a domain of absolute numbers and binary logic. In the field of artificial intelligence, non-binary reasoning and expression of uncertainty has progressed considerably (cf. Russell and Norvig 2009). There are subtler computing logics than the first-order logic that currently makes up the bulk of commonly used computer languages (cf. Forbus 2008; Pratt 1976). Some are concerned, for example, with modeling intuitive notions of truth and validity. Their nature may be a much closer fit for the hermeneutics of humanities. Exploration of the hermeneutic potential of computation is a challenge that digital humanities could pose, to itself and to computer science on behalf of the humanities. This need not imply transforming the nature of computation, but it must involve remediating the nature of hermeneutics through computational logic and design, informed strongly by a dialog with

humanities. It is apparent that scientific methods deriving from the humanities would be more appropriate than scientism for artificial intelligence and computer science when interacting with the humanities (cf. Mallery, Hurwitz, and Duffy 1986).

2.4 Conclusion

Unquestionably there is a role for hermeneutics in digital humanities. Thus the question becomes: What does such a hermeneutics look like? Capurro has shown how profound the ontological implications of digitality are for cultural dynamics and for the creation of humanistic artifacts. From this it follows that humanities must consider the extent and characteristics of a hermeneutics that takes digitality and computationality into consideration. Current practice shows, if it was not already self-evident, that conventional hermeneutics in its form of "post-algorithmic" interpretation takes up a large and undeniable part. At the same time, as we apply algorithms, models, and quantification, there arises an urgent need to understand the effect of these analytic methods on our hermeneutics. We have seen that the design of analytic methods is not free of its own hermeneutics. The effects and ramifications of these implicit hermeneutics on humanistic interpretation and reasoning are nevertheless unclear, poorly understood, and hardly studied. To understand these effects more fully – that is, to understand the hermeneutics of algorithmic and quantified approaches – we need a constructive and intimate dialog with the domains of computer science and software design. We cannot simply face outward after the algorithmic fact and interpret its results without implicitly but unconscientiously being a proxy to its hermeneutics. The profound effects of the digital on human culture and the humanities demands that we fully grasp its potential for hermeneutics.