

Table of Contents

<i>Preface: Digital Humanities at a political turn?</i>	ix
<i>Introduction</i>	15
1. Digital Humanities, and beyond	15
2. Do we still need humanists, and why?	18
3. How this book is organized	19
<i>PART I: THE SOCIO-HISTORICAL ROOTS</i>	23
<i>Chapter 1 – Technology and the humanities: A history of interaction</i>	25
1.1 From Alan Turing to the modern computer	26
1.2 What computers cannot do: from analog to digital	29
1.3 Bush’s visionary dream	32
1.4 A mathematician with a Ph.D. in philosophy	35
1.5 Wiener’s ethics and politics of the computer	37
1.6 Licklider and the man-machine symbiosis	40
1.7 Libraries and information processing	43
1.8 Conclusion	46
<i>Chapter 2 – Internet, or the humanistic machine</i>	49
2.1 The design of the intergalactic network	49
2.2 The computer as a communication device	51
2.2.1 The birth of the ARPANET	53
2.2.2 The WWW: an authoring system in the heart of Europe	55
2.3 Web 2.0 and beyond	59
2.4 Leibniz’s <i>Lingua Characteristica</i> and the Semantic Web	62
2.5 Social and cultural inequalities on the Web	67
2.5.1 The digital divide	67
2.5.2 Geopolitics of the network	71
2.5.3 The value of cultural and linguistic diversity	73
2.6 The challenge of open knowledge	79
2.6.1 Big Data	81
2.6.2 Open data and the humanities	83
2.6.3 Open access	84
<i>Summary of Part I.</i>	90

PART II – THEORETICAL AND PRACTICAL DIMENSIONS 95

Chapter 3 – Writing and content production. 97

- 3.1 Writing, technology and culture 97
- 3.2 Writing from the margins. 99
- 3.3 Modes of production: layers, forms and genres 101
- 3.4 Rhetoric and the Internet 107
 - 3.4.1 Reconfigurations 107
- 3.5 Time in writing 108
 - 3.5.1 Technology and textuality 109
 - 3.5.2 Paratexts, microtexts, metatexts 112
- 3.6 Content usability and accessibility 114
 - 3.6.1. Elements of “interaction design” for the Web 116
- 3.7 Digital ethnographies 118
 - 3.7.1 Cultural interfaces and the ethnoscience of writing 118
 - 3.7.2. The Machine is Us. 120
 - 3.7.3 Goodbye Windows?. 120
 - 3.7.4 Behind the screens: the languages of the Web 121
 - 3.7.5 The seduction of discretion 122
- 3.8 Identity on the Web 123
 - 3.8.1 My Website, outsourced 123
 - 3.8.2 Digital literacy 124
- 3.9 Transitions. The edited human 126

Chapter 4 – Representing and archiving 129

- 4.1 The longevity of digital information 131
 - 4.1.1 Degradation and obsolescence. 132
- 4.2 Balancing tradition and innovation 133
 - 4.2.1 Proposals for preservation 134
 - 4.2.2 The role of languages and metadata 135
- 4.3 Markup standards and languages 136
 - 4.3.1 Marking-up a document 137
 - 4.3.2 XML and the OHCO theory. 138
 - 4.3.3 XML Schemas and the “document type” approach. 140
 - 4.3.4 TEI: A standard for the humanistic domain 141
 - 4.3.5 Schemas and namespaces: why we need formal vocabularies 142
 - 4.3.6 Beyond text: using annotations 143
- 4.4 Metadata and the description of the document 144
 - 4.4.1 The unambiguous identification of resources 144
 - 4.4.2 Metadata and modeling 144
 - 4.4.3 A Model for understanding metadata: FRBR. 147
 - 4.4.4 Tools for metadata: the role of Dublin Core 149
 - 4.4.5 Expressing metadata formally: RDF 151
 - 4.4.6 Taxonomies, thesauri, ontologies: towards semantics 154

4.4.7 Metadata and folksonomy: the user experience	155
4.5 Open archives	156
4.5.1 The open archives initiative	156
4.6 Digital libraries	157
4.7 Semantic repositories and networking	159
4.8 Text analysis and text mining	161
4.8.1 Performance or character string?	162
4.8.2 From text retrieval to text analysis	163
4.8.3 Towards text mining	164
4.9 New applied technologies in the digital humanities	166
<i>Chapter 5 – Searching and organizing</i>	<i>169</i>
5.1 The paradox of search according to Plato	169
5.2 Web topology and the (in)equality of nodes	171
5.3 The role of search engines on the Web	172
5.4 How search engines work	175
5.5 The trouble with search engines	180
5.6 Ethical and social implications	184
5.6.1 Copyright	185
5.6.2 Privacy	187
5.6.3 Politics and censorship	189
5.7 Cloud computing and the search for truth	191
5.8 Google, AI and Turing’s social definition of intelligence	195
5.9 Communication and freedom	198
5.9.1 Corporate knowledge or the end of science?	199
5.9.2 The power of the archive	201
<i>Summary of Part II</i>	<i>204</i>
<i>Conclusions – DH in a global perspective</i>	<i>207</i>
1. The periphery-center effect	207
2. Research and teaching experiences	208
3. Associations, journals and centers	215
<i>Notes</i>	<i>219</i>
<i>References.</i>	<i>229</i>

Preface
Digital Humanities at a political turn?

Geoffrey Rockwell

So what exactly is that new insurgency? What rough beast has slouched into the neighborhood threatening to upset everyone's applecart? The [MLA] program's statistics deliver a clear answer. Upward of 40 sessions are devoted to what is called the "digital humanities," an umbrella term for new and fast-moving developments across a range of topics: the organization and administration of libraries, the rethinking of peer review, the study of social networks, the expansion of digital archives, the refining of search engines, the production of scholarly editions, the restructuring of undergraduate instruction, the transformation of scholarly publishing, the re-conception of the doctoral dissertation, the teaching of foreign languages, the proliferation of online journals, the redefinition of what it means to be a text, the changing face of tenure — in short, everything.¹

THERE HAS BEEN a surge of interest in the digital humanities and its place in the liberal arts in the English-speaking world as represented by the Modern Language Association Annual Convention.² Much of the interest is coming from engaged new scholars in North America who are comfortable with new media as they grew up with it. Interest is also coming from outside the Anglo-American world as humanists in Europe and Asia reflect on this field and its opportunities in their academic traditions. I think of the Manifesto for the Digital Humanities that came out of THATCamp Paris in 2011,³ Patrick Svensson's articles in *DHQ*,⁴ or Wang and Inaba's article analyzing the language of the digital humanities.⁵ Of particular

interest are books not written in English or for an English audience because they introduce the field in subtly different ways. One such work is *The Digital Humanist: A Critical Inquiry* by Teresa Numerico, Domenico Fiormonte and Francesca Tomasi, translated from the Italian by Desmond Schmidt and Christopher Ferguson. This is by no means the first work in Italian about computing in the humanities. The “informatica umanistica” (humanities informatics) school is rooted in the pioneering work of Father Busa, and all three authors have been active in the field since the mid-nineties at well-established research centres in Rome and Bologna.⁶ The new version of this book⁷ is current, accessible, and argues that humanists need to engage in not only the development of online content but also with ethical issues around computing, especially issues around language, search engines, open access and censorship. The authors call on humanists to acquire the skills to become digital humanists:

[H]umanists must complete a *paso doble*, a double step: to rediscover the roots of their own discipline and to consider the changes necessary for its renewal. The start of this process is the realization that humanists have indeed played a role in the history of informatics. (Introduction: Do we still need humanists, and why?)

The Digital Humanist is a work of five chapters, introduction, and a conclusion that is designed to introduce humanists to the digital, its human history and the cultural challenges that concern us. The first chapter, “Technology and the humanities: a history of interaction” is a deft tour through the history of computing that emphasizes the importance of human issues while still covering many of the important moments from Turing to social media. The authors start with the computer as a symbol manipulator as opposed to a mere calculator. They introduce Vannevar Bush and the importance of human association in the organization of human knowledge. They write about cybernetics and Wiener’s ethical concerns that computers might control us. They focus on Licklider and man-machine symbiosis as an alternative model to our relationship to computers – an alternative to the AI model where computers replace human work. This is linked to Licklider’s work on information processing and libraries. They then turn to the development of the ARPANET in the second chapter and its evolution. This leads to a discussion of the Web and Web 2.0 ideas. The second chapter ends with two paragraphs discussing how humanists can contribute to cultural diversity on the Web and make the Web more socially inclusive.

Chapter 5, “Searching and organizing” looks closely at the role of search engines, especially Google, in the organization of our knowledge. The first two chapters and chapter 5 frame the two internal chapters that are about digital philology and textual representation, which is why I will deal with them separately. The chapter titled, “Searching and organizing” starts with the old philosophical question of how you can ask about that which you don’t know and connects it to a discussion of how search engines work. The authors argue in the end that a) search engines are important to how knowledge is being organized, even more so now that Google is digitiz-

ing scholarship on a large scale, and b) that they are not neutral – that their algorithms are biased against information that isn't popular or in the dominant language of the Web, English. They return to the access issues of the first couple of chapters and ask if we are comfortable with commercial organizations organizing the human knowledge we in the humanities care about. Without preaching a solution they try to show the humanities reader how high the digital stakes are. This ethical-political turn is perhaps one of the features of *The Digital Humanist* that differentiates it from the more enthusiastic discourse around the digital humanities in the English-speaking world which tends to concentrate on modeling knowledge outside the political.⁸ *The Digital Humanist* addresses an audience concerned with cultural issues that still believes in political action and still believes the humanities are caretakers of a body of knowledge with political value.

The humanities heart of the book is the third chapter on “Writing and content production.” This chapter tackles the digital text through a number of theoretical questions starting from reflections on orality/literacy to questions about how we define our identity through online writing in blogs and other social media. Three moves that the authors make in this dense chapter are interesting.

Layers of Digital Textuality. The authors present a typology of digital texts that illustrates just how difficult it is to talk about digital textuality. The typology starts from what we typically call the Text In Itself (email, blog entries, wiki pages). They then shift to the Coded Text (ASCII, HTML, XML) that underlies the text itself, but, of course, is also a text we write. Then they move to the Processed Text like that text generated by Google when you query it or texts mashed up through social media. Finally they move to the Text Which Writes Us (credit cards, debit cards, text games) and draw our attention to the ways in which we are defined by texts from our credit rating to the interactive games we play. The authors are aware of the limits to this layered typology, but it serves well to break open our idea of exactly what a text is on the computer. One could add other layers like the Inscribed Text, which would be the material ways a text is written on a hard drive or CD-ROM.

Time and Space of Writing. Starting with section 3.5 there is a very interesting discussion of the shift from the temporality of modern narrative writing to the ways in which the Web (which is, after all, mostly writing) is seen spatially. For the authors this shift in metaphor is important to understanding online textuality. They follow this up by describing how in web-writing it is the paratext, microtexts, and metatexts that are important — more important, and stand in for the text itself. It is the metadata keywords you provide for a page that Google uses, it is the headings that people read, and it is all the navigating text that people use to understand what your site is about. The point is that if you look at web-writing advice it isn't really any more the old rhetorical advice about how to write your paragraphs – it is about how to contextualize your text and make it easy to navigate.

Usability and Ethnography. The reflections on the importance of the paratext lead to an argument that the interface is the new face of text and therefore usability is the new metric for studying the interface/text. This leads to a discussion of the place of ethnography as a method for digital humanists who are studying digital tex-

tuality as they write it. There is a future in the digital humanities for the way of doing philology which operates as an interface to our cultural identity.

At the end of the chapter the authors return again to identity and the ways the digital writes us as we write it. The importance of the digital humanities is that with the digital text it is not enough to simply study the text as linguistic meaning (layer 1), humanists need to understand the technologies of computing and culture of computing in order to get at the text in all its layers. We need to deconstruct the system that manages the social text from Google to Facebook. We need to ask who owns our text (which writes our identity), who manages it, and who provides access. The good old days when the technology was just a tool are over and *The Digital Humanist* calls for a new hermeneutic for the humanities that can study culture in the digital.

The fourth chapter is the practical sequel to the chapter about writing. “Representing and archiving,” as the chapter is entitled, focuses on the pragmatics of scholarly electronic editions and digital libraries. It provides a tour that starts by describing the problems around preserving electronic texts and moves up to big data. The chapter takes us through the uses of metadata, markup and digital libraries. It is a concise review of the key technologies we use to represent and preserve information. It is the sort of practical overview humanists need to have to understand digital humanities electronic text projects.

It is worth noting that the title is not about the “digital humanities,” but the “digital humanist.” It is about the formation of a new and engaged humanist. This is a work calling for and about the formation of a new persona in the tradition of the humanities. It tries to convince humanities students that they need to engage the digital and then provides a tour through what they need to know from the history of computing and the human to the importance of search engines. It calls them to question the digital infrastructure being built — infrastructure which, to someone outside the English-speaking world, is biased. We need digital humanists who don’t just use what is at hand, but inquire critically into what is in their hands. We need humanists that ask about how it might bias the representation, conservation and interpretation of the cultural record.

Beyond Big Data, mega-platforms and the mass archiving of data, the true innovation of the next decade of DH appears to be its geographic expansion and the consequent enlargement (and deepening) of these questions. (Conclusions: The periphery-center effect)

Above all, we need forms of innovation which are not of the bigger and even bigger kind. The authors call for innovation from the periphery and for the periphery rather than the dominant centralized variety characterized by large centers and mega-projects. *The Digital Humanist* itself comes from outside the loop of English-speaking centres, though I’m not sure I would call Italy a periphery. It imagines a way of doing digital work which doesn’t necessarily involve grants. What could we do

with the resources at hand? How could we imagine philological projects that could be adapted by others, whatever their resources and wherever they are?