Defining Digital Humanities

For
Anthony, Edward and Fergusson,
Clara and Joey,
and
Wonne and Senne

Defining Digital Humanities
A Reader

Edited by

MELISSA TERRAS

JULIANNE NYHAN

EDWARD VANHOUTTE

ASHGATE

Published by
Ashgate Publishing Limited
Wey Court East
Union Road
Farnham
Surrey, GU9 7PT
England

Ashgate Publishing Company
110 Cherry Street
Suite 3-1
Burlington, VT 05401-3818
USA

www.ashgate.com

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

The Library of Congress has cataloged the printed edition as follows:
Defining digital humanities : a reader / [edited] by Melissa Terras, Julianne Nyhan, and Edward Vanhoutte.
pages cm
Includes bibliographical references and index.
ISBN 978-1-4094-6962-9 (hardback) – ISBN 978-1-4094-6963-6 (pbk) – ISBN 978-1-4094-6964-3 (epub) 1. Humanities–Data processing. 2. Humanities–Research–Data processing. 3. Information storage and retrieval systems–Humanities. 4. Humanities–Electronic information resources. I. Terras, Melissa M. II. Nyhan, Julianne. III. Vanhoutte, Edward.

AZ105.D44 2013
001.30285–dc23

2013020285

ISBN 9781409469629 (hbk)
ISBN 9781409469636 (pbk)
ISBN 9781409469643 (ebk – PDF)
ISBN 9781409469650 (ebk – ePUB)

MIX
Paper from responsible sources
FSC www.fsc.org FSC® C013985

Printed in the United Kingdom by Henry Ling Limited, at the Dorset Press, Dorchester, DT1 1HD

CHAPTER 6

The Gates of Hell: History and Definition of Digital | Humanities | Computing[1]

Edward Vanhoutte
Royal Academy of Dutch Language & Literature

1. A Metaphor

In 1879, Edmund Terquet (1836–1914), the French Secretary of State for Fine Arts, commissioned a monumental door from the sculptor Auguste Rodin (1840–1917). Rodin's door would be used as the entrance to the planned Decorative Arts Museum in Paris. The artist was given three years to complete it, but the museum project started to go wrong, and the state cancelled it in 1889. In the meantime, the door had lost its original setting and function and Rodin, freed from the restrictions of designing a functional piece of art, explored the creative possibilities of the surface and created a sculpture which he would constantly revisit until his death. The sculpture, which is on exhibition at the Rodin Museum in Paris,[2] is unmistakingly a door, with its two leaves, sideparts, and tympanum. And yet, the door doesn't open. There is no opening mechanism and, even if there were one, the more than 200 figures and groups on the door are too entangled and prevent any movement of the leaves. Rodin called his sculpture *La Porte de l'Enfer* or *The Gates of Hell*, since his original inspiration was the then very popular theme of Dante's *La Divina Commedia*.

When I was watching the documentary 'A Season in Hell. Rodin's Gate',[3] it struck me that the story of Rodin's sculpture could be used as a metaphor for the field of Humanities Computing.[4] By 'Humanities

[1] This essay is for Ron Van den Branden: *Sine te*...

[2] http://www.musee-rodin.fr/fr/collections/sculptures/la-porte-de-lenfer [accessed 12 January 2013].

[3] http://www.canal-educatif.fr/en/videos/art/2/rodin/gates-of-hell.html [accessed 12 January 2013].

[4] Throughout this essay, Humanities Computing with capitalization refers to the field and humanities computing without capitalization refers to the activity of computing in and

Computing' I mean the practice of using computing *for* and *in* the humanities[5] from the early 1950s to 2004 when 'Digital Humanities' became the prominent name for the field.

Just as Rodin's 'door', Humanities Computing consisted of two clearly separated leaves with their own history and understanding behind them but, when put together, they became so heavily interlinked that they could not be separated without any loss of meaning. Humanities Computing was neither a traditional humanities nor a computing subject. That's why, in the course of time, the self-reflective question what constitutes and defines Humanities Computing has in itself become a research theme.

However, the main reason why Rodin's *Gates of Hell* is such a good metaphor for Humanities Computing is that it is the creative result of failure. The failure on the part of the French government to build the Decorative Arts Museum in Paris freed Rodin's design from the functional restrictions, and paved the way for an almost exuberant creation. Likewise, Humanities Computing is the creative result of failure on the part of the manufacturers of early computers to produce operational machines in time to be used during the Second World War (or, one can argue, of failure on the part of the allied forces to make the war last longer).

2. Failure

Probably the first mention of the application of computing to the Arts is found in the notes to the translation of Luigi Federico Menabrea's (1809–1896) *Notions sur la machine analytique de Charles Babbage* (1842)[6] – translated in 1943 as *Sketch of the analytical engine invented by Charles Babbage*[7] (Menabrea, 1961 [1843]; Lovelace, 1961 [1843]) by Augusta Ada, Countess of Lovelace (1815–1852).[8] With the poet Lord Byron (1788–1824) as her father and the mathematician Anabella Milbanke (1792–1860) as her mother, Ada Lovelace, as she is more frequently called, may well be considered the personification of the humanities computing educational idea. Meditating upon the possible uses

for the humanities.

5 By computing *for* the humanities, I mean the instrumental use of computing for the sake of the humanities. By computing *in* the humanities, I mean the meaning-generating activity of Humanities Computing.

6 Originally published in the *Bibliothèque Universelle de Genève*, 82 (October 1842).

7 These *Notes* were published separately in *Scientific Memoirs, Selections from The Transactions of Foreign Academies and Learned Societies and from Foreign Journals*, edited by Richard Taylor in 1843.

8 See Toole (1996 and 1998) for biographical notes and comments on her work.

of Babbage's Analytical Engine[9] for non-numerical purposes, she wrote that the operating mechanism:

> might act upon other things besides *number*, were objects found whose mutual fundamental relations could be expressed by those of the abstract science of operations, and which should be also susceptible of adaptations to the action of the operating notation and mechanism of the engine. Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent. (Lovelace, 1961 [1843], pp. 248–9)

However, the first computer music wasn't produced before CSIRAC,[10] Australia's first digital computer, was used to perform the Colonel Bogey March in 1950 or 1951, and electronic computer music boomed from 1957 onwards with the release of the first program for sound generation, appropriately called MUSIC.[11] Moreover, the submission of musicologist papers to the journal *Computers and the Humanities* in the 1960s and 1970s is substantial. But Lovelace was right in her observation that computing techniques and devices could have their use in non-numerical applications as well. This was especially realized after the end of the Second World War.

In 1943 the US military[12] commissioned the building of the ENIAC (Electronic Numerical Integrator And Computer) to calculate trajectories of World War II artillery guns, a task that involved repetitive sequences of operations on complex mathematical data. The two leading architects of this giant electronic digital calculator were J. Presper Eckert (1919–1995)[13] and John Mauchly (1907–1980)[14] of the University of Pennsylvania's Moore School of Electrical Engineering. Before the ENIAC, these operations had been executed with the use of differential analysers, desk calculators, and punched-card installations, consisting of several serialized punched-card machines (Polachek, 1997), a market dominated by IBM at that time.

[9] Babbage's Analytical Engine was a proposed programmable mechanical calculator with a planned memory of 1,000 numbers of 50 digits. It used punched cards for the input of instructions, the input and output of data, and the storage of data and instructions.

[10] CSIRAC: Council for Scientific and Industrial Research Automatic Computer.

[11] The program was written by Vernon Matthews (b. 1926) at Bell Labs (Doornbusch, 2004 and 2005).

[12] More particularly, the Ordnance Department.

[13] See Eckstein (1996) for an account of the early life of Eckert, and Wilkes (1995) for a tribute to his work.

[14] See Stern (1980) for a biographical note on Mauchly, and Mauchly (1984) for an account of his crucial early years of experimenting and research.

When ENIAC was assembled and delivered to the US army in 1946, its development and production time had exceeded the war and its envisioned purpose for warfare had therefore become redundant. The same happened with the EDVAC (Electronic Discrete Variable Automatic Computer), the first binary stored program computer, which was commissioned from the same team in 1944 and which only became fully operational in 1951.

With the end of the Second World War, the urgent need for computing power for warfare purposes disappeared, although the Cold War kept the importance of (classified) computer research programs at the top of the intelligence agenda till the early 1990s. Among the early thinkers on the social function of computing technology was Warren Weaver (1894–1978)[15] who had been involved with ballistics during the war and who had become director of the Natural Sciences Division of the Rockefeller Foundation afterwards. Inspired by pioneering pre-war computing projects and the developments he witnessed during the war, he started to wonder what sort of applications 'this incredibly powerful tool, the electronic computer' (Weaver, 1970, p. 105) could be used for. The warfare computing practices of ballistics and cryptanalysis convinced him that the computer could be used for two peaceful academic applications in particular: one in the Sciences and one in the Humanities, namely mathematics and machine translation respectively (Weaver, 1970, pp. 104–08).

Just as the failure of the Paris museum project freed Rodin from the restrictions of a functional door, the end of the Second World War freed Weaver from seeing the computer only as a warfare tool for ballistics and cryptanalysis.

3. Machine Translation

Machine Translation (MT) is 'the application of computers to the translation of texts from one natural language into another' (Hutchins, 1986, p. 15). The arguments for research into MT are pragmatic and social (people have to read documents and communicate in languages they do not know), academic and political (international cooperation and globalization through the removal of language barriers in order to promote peace and further knowledge in developing countries), military (to find out what the enemy knows), scholarly (to study the basic mechanisms of language and mind and exploit the possibilities and limits of the computer), and economical (to sell a successful product).

[15] See Weaver (1970) for his autobiography, and Hutchins (2000b) for a brief biographical note.

The early 1950s saw some experiments with word-for-word translations of scientific abstracts by Richard H. Richens (1918–1984)[16] and Andrew D. Booth (1918–2009)[17] using punched cards (Richens and Booth, 1952).[18] Until then, the problem of automating translation was thought of in mechanical terms solely: the development of a dictionary lookup system in aid of the human translator. Andrew Booth, a crystallographer at Birkbeck College (University of London) was probably the first person to refer to the possible use of electronic computers for Machine Translation. In a memorandum to the Rockefeller Foundation dated 12 February 1948, he wrote:

> A concluding example, of possible application of the electronic computer, is that of translating from one language into another. We have considered this problem in some detail, and it transpires that a machine of the type envisaged could perform this function without any modification in its design. (quoted from Weaver, 1965 [1949], p. 19)

A Rockefeller Research Fellow at the Institute for Advanced Study at Princeton, Booth was reporting to Warren Weaver who, as early as 1946, had had several conversations with Booth on the use of automatic digital computers for mechanical translation from one language into another (Booth and Locke, 1965 [1955], p. 2; Booth, 1980, p. 553; Hutchins, 1986, p. 24). In the course of his fellowship, Booth, together with his assistant Kathleen Britten (b. 1922), who later became his wife, developed a detailed code for storing a dictionary in an automatic digital computer's memory to be retrieved from standard teletype input. This idea dated back from 1946 and realized dictionary translation on an automatic computer (Booth, 1958, pp. 92–9).

It was Booth's work and his own experience as a cryptanalyst during the war that formed the basis for Weaver's memorandum 'Translation' which was issued on 15 July 1949 (Weaver, 1965 [1949]). The *Weaver Memorandum* was circulated amongst twenty or thirty 'students of linguistics, logicians, and mathematicians' (Weaver, 1970, p. 107) and up to 200 scholars (Locke and Booth, 1965 [1955], p. 15) in different fields. It was this memorandum which initiated research projects at different

16 See Sparck Jones (2000) for an overview of Richens' work in MT.

17 See Booth (1997) and Booth and Booth (2000) for an overview of Booth's work in MT. See Booth (1980) for details on his work in crystallography and the development of early computers and magnetic storage devices. Booth was also the holder of the British patent on the floppy disk.

18 This paper 'Some methods of mechanized translation' was written in 1948, but not published before 1955 (Sparck Jones, 2000, p. 263). The paper was presented on the first Conference on Mechanical Translation at MIT in June 1952.

universities and generated some early writings on the problems involved with Machine Translation.[19] These problems included ambiguity of words, the semantic function of syntax, and the resolution of word order problems in different languages.

In 1952, eighteen scholars, including Booth as the only non-American delegate, gathered on the first 'international' conference on Machine Translation at MIT, followed by a meeting later that year in London where some forty linguists met during the International Linguistic Congress. A year later, Machine Translation appeared for the first time in a scholarly textbook written by Andrew and Kathleen Booth. In their book *Automatic Digital Calculators* (Booth and Booth, 1953), aimed at a readership of computer scientists, the authors published a chapter on 'Some applications of computing machines' in which Machine Translation was discussed at length.[20] In 1954, a widely publicized demonstration took place at IBM headquarters and involved a carefully selected sample of 49 Russian sentences, a limited vocabulary of 250 Russian words from different fields and their English equivalents, and six rules of syntax. The IBM press release quoted: 'A girl who didn't understand a word of the language of the Soviets punched out the Russian messages on IBM cards. The "brain" dashed off its English translations on an automatic printer at the breakneck speed of two and a half lines per second.'[21] In the same year, the first doctoral dissertation on Machine Translation was presented by Anthony Oettinger (b. 1929) at Harvard University (Oettinger, 1954) and the journal *Mechanical Translation* appeared for the first time.

From 1955 to 1966, the field organized itself in groups working mainly on dictionary, lexicographic, and semantic problems and groups working on syntactic problems; in groups that took an empirical approach (mainly in the UK) and others that took a theoretical approach (mainly in the US); and in groups working towards operational systems in the short term, and groups working toward high quality systems in the long term. These years saw a dozen important conferences, gatherings, and sessions on Machine Translation, and the founding of the Association for Machine Translation and Computational Linguistics (AMTCL) on 13 June 1962. Apart from the US and the UK, research was undertaken in e.g. Bulgaria, Canada, former Czechoslovakia, France, Israel, Japan, the former USSR, and the later independent states.[22]

---

19 See the annotated bibliography in Locke and Booth (1965 [1955], pp. 227–36).

20 The book was reprinted twice, in 1956 and 1965, and translated into Russian in 1957.

21 IBM press release, 8 January 1954, http://www-03.ibm.com/ibm/history/exhibits/701/701_translator.html [accessed 12 January 2013].

22 For recollections and overviews of research in these countries, see the different papers in Hutchins (2000a).

Whereas the funding agencies in the US had applauded the importance of Machine Translation to 'the overall intelligence and scientific effort of our nation' in a report compiled by the Committee on Science and Astronautics of the US House of Representatives in June 1960 (Hutchins, 1986, p. 159), six years later the final conclusions of an independent advisory committee installed by request of the funding bodies put an end to the funding of research in Machine Translation in the US. The notorious ALPAC[23] report 'Languages and Machines: Computers in Translation and Linguistics' (ALPAC, 1966) criticized the need, cost, and performance of automatic translations and even suggested that, since English is the dominant language in science, it was more cost efficient to teach heavy users of translated Russian articles Russian than to provide them with a translation service. The final recommendations outlined that funding should be provided for the improvement of translation by developing machine aids for human translators and for Computational Linguistics, which had grown out of Machine Translation. The ALPAC report put the research towards perfect translation to an end and referred its ideal to the realm of utopia.[24] For linguistics in general, and for Computational Linguistics and Humanities Computing in particular, the report put the future research programme on language in a different perspective, or as Victor Yngve put it:

> The future of linguistics is not in philosophy, from which it is emerging, but in standard science, into which it can now move with confidence. This requires that linguistics finally recognize that the true object of study of a scientific linguist is the people that speak and understand and communicate in other ways, and other relevant aspects of the real world. (Yngve, 2000, p. 69)

Roberto Busa (1913–2011) seemed to agree with Yngve when he identified the major problem with research in Machine Translation not as the inadequacy of *computers* to deal with human language, but as *man's* insufficient comprehension of human languages (Busa, 1980, p. 86).

[23] Automatic Language Processing Advisory Committee.

[24] The ALPAC report caused ten years of neglect of Machine Translation from the scientific world, and more from the funding bodies, and it fostered a general belief with the public that Machine Translation was more fiction than science. A renewed interest in Machine Translation can be observed from the 1980s onwards with a new journal, *Computers and Translation* (1986), which changed its name to *Machine Translation* (1989), and a series of international conferences and summits. In 1989, *Literary and Linguistic Computing* (4 (3)) devoted a special section to Machine Translation containing six papers edited and introduced by Antonio Zampolli.

4. Lexical Text Analysis

Machine Translation was highly involved with the electronic processing of humanities data. Early writings on Machine Translation mention the essential use of concordances, frequency lists, and lemmatization, which were, according to Antonio Zampolli (1937–2003) typical products of Lexical Text Analysis (LTA) (Zampolli, 1989). In this respect, it is not surprising to find an article on 'The Computer in Literary Studies' in a volume devoted to Machine Translation (Booth, 1967).

Collaboration between scholars of Machine Translation and of Lexical Text Analysis in the 1950s and early 1960s has been reported by Michael Levison, who joined Booth's laboratory as a PhD student in 1958. Although Booth's humanities-based work was mainly situated in Machine Translation, there was a strong interest in the application of the computer to other linguistic processes from the mid-1950s onwards (Booth et al., 1958). Programs for the statistical analysis of text, stylometry, and the production of concordances were developed in the early 1960s (Levison, 1962) and in his 1967 article on 'The Computer in Literary Studies' Levison describes the following classes of literary problems in which computers can be used successfully: concordances, glossaries, authorship attribution, stylistic studies, relative chronology, fragment problems with papyri, and even a preliminary form of the digital library described as a tape library (Levison, 1967). A 'steady stream of visitors' who came 'seeking help with literary and linguistic problems' (Lessard and Levison, 1998, p. 262) frequented Booth's laboratory to work on all of these literary problems, and even a couple of geographers turned up with a proposal to investigate the possibility of accounting for 'Polynesian settlement by drift voyaging', using simulation (Lessard and Levison, 1998, p. 262).

Although Booth had left the laboratory before all of these projects came to fruition, it is certainly his inspiration and reputation that brought about the cooperative ventures. Two of Booth's students, Leonard Brandwood and John Cleave, may even have been the first PhD students who applied computers to non-translation language problems in the Humanities. Brandwood worked on the chronology and concordance of Plato's works (Booth et al., 1958, pp. 50–65), and Cleave on the mechanical transcription of Braille (Booth et al., 1958, pp. 97–109).

One of the most important early computing projects which made use of Lexical Text Analysis, however, was Roberto Busa's *Index Thomisticus*, a lemmatized concordance of all the words in the complete works of Thomas Aquinas. Although the first mention of the project was a short project description published in *Speculum* in January 1950 (Busa, 1950),[25] Busa himself dates his original idea of using modern mechanical techniques for the linguistic analysis of written texts between 1941 or 1942 (Busa, 2004b, p. xvi; Busa 2002, p. 49) when he started his PhD research, and 1946 when he completed his dissertation and was looking for a follow-up research project (Busa, 1980, p. 83). The fact is that Busa's dissertation (Busa, 1949) was written without the use of or reference to any computer technology.[26] But in 1951 Busa teamed up with people from IBM in New York to automatically compile a concordance of the poetry of Thomas Aquinas, which was the first example of a word index printed by punched-card machines (Busa, 1951). However, this proof of concept exercise used no computing and no programming. The main innovation was Busa's insight that commercial accounting machines could be used for humanities purposes with good results. The result of the 1951 project offered six scholarly tools: an alphabetical frequency list of the words; a retrograde frequency list of the words; an alphabetical frequency list of words set out under their lemmata; the lemmata; an index of the words; and a KWIC Concordance (Winter, 1999).

For his complete *Index Thomisticus*, Busa calculated that the stack of punchcards would have weighed 500 tonnes, occupying 108 m³ with a length of 90 m, a depth of 1 m, and a height of 1.20 m. By 1975, when the *Index Thomisticus* was completed and started to appear on 65,000 pages in 56 volumes (Busa, 1974–1980) some 10,631,973 tokens were processed.[27] This processing consisted of inputting, verifying, and interpreting with references and codes which specify the values within the levels of the morphology – the 'internal hypertext' in Busa's terminology (Busa, 2002 and 2004a). The work was done by a team of keypunch operators who were trained in Busa's own training school which ran from 1954 to 1967 (Busa, 1980, p. 85).

Whereas Busa was using keypunch technology in close cooperation with IBM, John W. Ellison completed his *Computerized Concordance to the Revised Standard Version of the Bible* with the computing facilities offered by Remington Rand, namely magnetic tape technology and the UNIVAC I mainframe computer (UNIVersal Automatic Computer)[28] in 1957. The story goes that Busa met Ellison around 1954, congratulated

[25] Although this publication is mentioned in Busa's Bibliography published in Busa (2002), it is not clear whether Busa is really the author of the piece, which is written in the third person.

[26] 1949 is used as the symbolic start of computational work in the humanities by several authors. Cf. recently Burdick et al. (2012, p. 123).

[27] Figure according to the project report *Opera quae in indicem thomisticum sunt redacta* (1975, revised 1980), privately made available to me.

[28] The UNIVAC computers were built by the same team which built the ENIAC and EDVAC computers.

him on his computing work, and went back to IBM to transfer the punch cards onto magnetic tape and use computer technology and programming[29] for the publication of his Dead Sea Scrolls project in 1957.[30] For the *Index Thomisticus*, Busa was working on 1,800 tapes, each one 2,400 feet long, and their combined length was 1,500 km (Busa, 2004b, p. xvii).

Ellison dates his original idea of using 'modern mechanical devices' back to 1945 when he realized that distinguished scholars 'having two or three earned doctorates, were essentially counting on their fingers as they studied manuscripts' (Ellison, 1965, p. 64). In 1950, he asked for computing time at the Harvard Computation Laboratory, which was granted in 1951. His proof of concept exercise was the internal collation of 309 manuscripts of the St. Luke gospel, printed against the standard text with a classification of eight kinds of variant readings with the MARK IV computer in 1952 or 1953. This was the first example of a manuscript collation carried out and printed by a computer.

Up to the publication of the infamous ALPAC report in 1966, Computational Linguistics and Lexical Text Analysis were not separated fields, and used statistical analysis for the creation of indexes, concordances, corpora, and dictionaries. But from then onwards, Computational Linguistics embraced the symbolic approach and abandoned statistical analysis which has been at the heart of Humanities Computing.

5. Literary and Linguistic Computing and Computing in/for the Humanities

The history of both Machine Translation and Lexical Text Analysis are closely related to the technological development of computing machinery, program languages, and software and the economic opportunities identified by their manufacturers. In the years following the end of the Second World War, traditional manufacturers and suppliers of analog tabulating equipment changed their core business to digital computing equipment and services, and were prospecting new markets. This is why key players like Remington Rand and IBM teamed up with humanities scholars and funded conferences and projects that explored new applications of computing. One such early conference was held at Yale University in January 1965 under the hesitating title *Computers for the Humanities?*

[29] Since FORTRAN was only released in 1975, the programming was still in card management.

[30] Also in 1957, and independently of the work of Busa or Ellison which hadn't appeared yet, Cornell University launched a program for a computer-produced series of concordances, with Stephen M. Parrish as general editor (Parrish, 1962, p. 3).

The cover of the proceedings, published under the same title (Pierson, 1965), shows a silhouette drawing of Rodin's *Le Penseur* (The Thinker) punched like a punched card to indicate the link between computing and the history of ideas. The proceedings contain papers on the history of computing and the use of computers in the Sciences; on computers and words; language and literature; computers and history; computers and the Arts; and a discussion of some possibilities and speculations on future computer projects. This book is probably the earliest volume surveying the early use of computing in the humanities beyond Machine Translation. Two years later, the selected papers from six such conferences sponsored by IBM,[31] which were attended by some 1,200 academics from all over the US in 1964 and 1965, were published under the not so hesitating title *Computers in Humanistic Research. Readings and Perspectives* (Bowles, 1967). The papers in this book deal with computational applications in anthropology, archaeology, history, political sciences, language, literature, and musicology.

In the UK, the Literary and Linguistic Computing Centre (LLCC) at the University of Cambridge was set up with Roy Wisbey (b. 1929) as its first director in 1964. It was also Wisbey who organized the first international conference on the use of the computer in literary and linguistic research which brought together British scholars with participants from Australia, Canada, continental Europe and the US in 1970.[32] In 1972, a second such conference was organized in Edinburgh.[33] The emphasis on literary and linguistic computing was also reflected in the name of the *Association for Literary and Linguistic Computing* (ALLC)[34] which he co-founded in 1973 and chaired from 1973 to 1978. The ALLC published a periodical called *ALLC Bulletin* from 1973 to 1985 and the *ALLC Journal* from 1980 to 1985. In 1986 both publications were replaced by the journal *Literary and Linguistic Computing* (*LLC*) which in 2005 changed its name to *LLC: The Journal of Digital Scholarship in the Humanities*. The ALLC started to organize a series of biannual conferences on literary and linguistic computing under its own name and the two previous conferences were added to the list.[35] From 1973 onwards, these conferences alternated with an American series of biannual conferences called International

31 The conferences were held in 1964 and 1965 at Rutgers, Yale, UCLA, the Consortium of Universities in Washington DC, Purdue, and Boston University.

32 Proceedings published in Wisbey (1971).

33 Proceedings published in Aitken et al. (1973).

34 http://www.allc.org [accessed 12 January 2013]. The ALLC was recently rebranded as EADH: The European Association for Digital Humanities.

35 This explains why the first two ALLC conferences listed on the EADH website were organized before the founding of the Association.

Conference on Computers in the Humanities (ICCH) in the odd years.[36] The professional association which was founded in the US in 1978 was hence called the Association for Computers and the Humanities (ACH).[37] Twelve years before, its founding president Joseph Raben (b. 1924) had started to edit the journal, *Computers and the Humanities* (*CHum*) which ran from 1966 to 2004.[38] Whereas the scope in Europe was mainly on literary and linguistic studies of language in literary form, the American conferences, journal and association showed a broader interest in computer-based studies of language in literary and non-literary form. This is reflected in the titles of the proceedings from the conferences held in the 1970s: *The Computer in Literary and Linguistic Research* (Wisbey, 1971); *The Computer and Literary Studies* (Aitken et al., 1973); *The Computer in Literary and Linguistic Studies* (Jones and Churchhouse, 1976); and *Advances in Computer-aided Literary and Linguistic Research* (Ager et al., 1979) in Europe, and *Computers in the Humanities* (Mitchell, 1974) and *Computing in the Humanities* (Lusignan and North, 1977) in North America. From the start, the ICCH conferences also included papers on history, musicology, computer assisted instruction, and creative arts (dance, music, poetry).[39]

The first monographs about computers in the humanities, however, came from the computer industry. In 1971, IBM published a series of application manuals on computing in the Humanities: *Introduction to Computers in the Humanities* (IBM, 1971a); *Literary Data Processing* (IBM, 1971b); and *Computers in Anthropology and Archaeology* (IBM, 1971c). Almost a decade later, and after thirty years of computing in the humanities, supporters on both sides of the Atlantic were treated to two textbooks on the topic which appeared in the same week in January 1980. Susan Hockey's *A Guide to Computer Applications in the Humanities* (Hockey, 1980a) and Robert Oakman's *Computer Methods for Literary Research* (Oakman, 1980)[40] provided the first consistent overviews from

[36] With the exception of ALLC conferences in 1985, 1986, 1987, and 1988. From 1989 a joint conference was organized yearly in Europe in the even and in the US or Canada in the odd years. From 1990 to 2005, the conference was called ALLC/ACH or ACH/ALLC depending on the location. With the foundation of ADHO in 2005, the conference was renamed 'Digital Humanities'.

[37] http://www.ach.org [accessed 12 January 2013].

[38] In 1968 another journal was launched: *Computer Studies in the Humanities and Verbal Behaviour*.

[39] In the UK, however, a separate series of conferences on Computer Assisted Teaching in the Humanities (CATH) were organized.

[40] Notice that it's now Hockey who uses 'computers in the Humanities' in the title of her book, and Oakman narrowed it down to 'literary research'.

an academic point of view.[41] Although both books filled an urgent need for a surveying textbook in the field of literary and linguistic computing, they were not explicitly conceived with a didactic point of view. The authors brought together the issues raised in the journal papers, the several collected volumes of conference proceedings, the available project reports, and the scarce manuals for specific programming languages and applications 'from the unifying perspective of one observer' (Oakman, 1980, p. x) and were very much alike.[42] In synthetizing thirty years of research, the books became reference points for further writing on the history in the field. In this respect it's relevant to notice that Hockey identified Busa as the pioneer of humanities computing, whereas Oakman named Ellison.

One of the first mentions of 'humanities computing' to name the activity of computing in and for the humanities was in an article in the second issue of *CHum* about the use of PL/I as a programming language for humanities research in 1966 (Heller and Logemann, 1966).[43] In 1968 Aldo Duro published a survey of 'Humanities Computing Activities in Italy' (Duro, 1968) which suggests that that the term was already well known, though not dominant in the community.[44] Whereas the late 1960s saw the introduction of the term to name the computing activity, the term began to mark the field in the early 1970s, as we can see in Stacey Tanner's report on the ALLC conference of 1974 published in *Dataweek* and reprinted in *ALLC Bulletin* (Tanner, 1975). Tanner reported on Busa's address to the conference by paraphrasing – the term is not used by Busa himself – that he talked about 'the future of humanities computing' and about 'projecting the programs of humanities computing' (Tanner, 1975, p. 54). By the 1980s, the use of the term for the field was widespread, as demonstrated in Busa's retrospective paper 'The Annals of Humanities Computing: The Index Thomisticus' (Busa, 1980) – although neither Hockey (1980a) nor Oakman (1980) use the term to name the field. From the mid-1980s onwards, 'Humanities Computing' started to appear in the

41 Although Howard-Hill's *Literary Concordances* (Howard-Hill, 1979) claimed to be 'A Complete Handbook for the Preparation of Manual and Computer Concordances', the book deals very little with computing. It was, however, published a year before Hockey's and Oakman's books. Hockey had finished writing her book in 1978, but the publisher sat on the manuscript for quite a while. She heard about Oakman writing his book when her manuscript was already at the publishers (Susan Hockey, personal communication, 5 June 2005). Oakman completed his manuscript in early 1978 (Oakman, 1984, p. xv) and explicitly mentioned Hockey's book in the revised reprint from 1984.

42 'I think the similarities are due to the fact that there wasn't a lot of material to draw on, only the proceedings of some conferences and *CHum* and the ALLC publications' (Susan Hockey, personal communication, 5 June 2005).

43 With thanks to Willard McCarty for providing me with a copy of this article.

44 The University of Colorado already had an operational Humanities Computing Facility around the same time.

names of North American teaching programmes (Ide, 1987), computing centres (University of Washingon and McMaster University) and facilities (Arizona State University, Duke University and UCLA). Although the use of the name to delimit a distinctive and coherent discipline was a frequent matter of debate (Miall, 1990, p. 3), the publication of two volumes of the *Humanities Computing Yearbook* (Lancashire and McCarty, 1988 and Lancashire, 1991) and five volumes of *Research in Humanities Computing* (1991–1996) established the name after almost two decades of hesitating use.

6. Text Encoding

One of the main problems since the earliest uses of computers and computational techniques in the humanities was the representation of data for input, processing, and output. Computers, as Michael Sperberg-McQueen has reminded us, are binary machines that 'can contain and operate on patterns of electronic charges, but they cannot contain numbers, which are abstract mathematical objects not electronic charges, nor texts, which are complex, abstract cultural and linguistic objects' (Sperberg-McQueen, 1991, p. 34). This is clearly seen in the mechanics of early input devices such as punched cards where a hole at a certain coordinate actually meant a I or 0 (true or false) for the character or numerical represented by this coordinate according to the specific character set of the computer used. Because different computer systems used different character sets with a different number of characters, texts first had to be transcribed into that proprietary character set. All characters, punctuation marks, diacritics, and significant changes of type style had to be encoded with an inadequate budget of characters. This resulted in a complex set of 'flags' for distinguishing upper-case and lower-case letters, for coding accented characters, the start of a new chapter, paragraph, sentence, or word. These 'flags' were also used for adding analytical information to the text such as word classes, morphological, syntactic, and lexical information. Ideally, each project used its own set of conventions consistently throughout. Since this set of conventions was usually designed on the basis of an analysis of the textual material to be transcribed to machine-readable text, another corpus of textual material would possibly need another set of conventions. The design of these sets of conventions was also heavily dependent on the nature and infrastructure of the project, such as the hardware and software.

Although several projects were able to produce meaningful scholarly results with this internally consistent approach, the particular nature of each set of conventions or encoding scheme had lots of disadvantages. Texts prepared in such a proprietary scheme by one project could not readily be used by other projects; software developed for the analysis of such texts could hence not be used outside the project due to an incompatibility of encoding schemes and non-standardization of hardware. However, with the increase in texts being prepared in machine-readable format, the call for an economic use of resources increased as well. Already in 1967, Michael Kay argued in favour of a 'standard code in which any text received from an outside source can be assumed to be' (Kay, 1967, p. 171). Ideally, this code would behave as an exchange format which allowed the users to use their own conventions at output and at input (Kay, 1967, p. 172).

Some sort of standardization of markup for the encoding and analysis of literary texts was reached by the COCOA encoding scheme originally developed for the COCOA program in the 1960s and 1970s (Russell, 1967) but used as an input standard by the Oxford Concordance Program (OCP) in the 1980s (Hockey, 1980b) and by the Textual Analysis Computing Tools (TACT) in the 1990s (Lancashire et al., 1996). For the transcription and encoding of classical Greek texts, the Beta-transcription/encoding system reached some level of standardized use (Berkowitz and Squiter, 1987).

In 1987, a group of thirty-two humanities scholars[45] gathered at Vassar College in Poughkeepsie, New York in a two-day meeting (11 and 12 November 1987) called for by the ACH and convened by Nancy Ide and Michael Sperberg-McQueen. The main topic of the meeting was the question how and whether an encoding standard for machine-readable texts intended for scholarly research should be developed. The conclusions of the meeting were formulated as a set of methodological principles – the so-called 'Poughkeepsie Principles'[46] – for the preparation of text encoding guidelines for literary, linguistic, and historical research (Burnard, 1988, pp. 132–3; Ide and Sperberg-McQueen, 1988, pp. E.6–4, and 1995, p. 6).

For the implementation of these principles the ACH was joined by the ALLC and the Association for Computational Linguistics (ACL).[47] Together they established the Text Encoding Initiative (TEI) whose mission it was to develop workable text encoding guidelines. The TEI very soon came to adopt the Standard Generalized Markup Language (SGML), an ISO standard published in 1986 (Goldfarb, 1990), as the recommended

[45] Amongst the delegates were representatives from the main European text archives and from important North American academic and commercial research centres.

[46] I am quoting The Poughkeepsie Principles in 'Module 0: Introduction' of *TEI by Example*, http://www.teibyexample.org on which this section on text encoding is based [accessed 12 January 2013].

[47] http://www.aclweb.org [accessed 12 January 2013].

encoding format for electronic texts.[48] Michael Sperberg-McQueen was appointed editor-in-chief and Lou Burnard as European editor of the Guidelines.

The first public proposal for the TEI Guidelines was published in July 1990 under the title *Guidelines for the Encoding and Interchange of Machine Readable Texts* with the TEI document number TEI P1 – for Proposal 1 (Sperberg-McQueen and Burnard, 1990). Further development of the TEI Guidelines was done by four Working Committees (Text Documentation, Text Representation, Text Analysis and Interpretation, Metalanguage and Syntax) and a number of specialist Working Groups.[49] The results of that work included substantial amounts of new material and were published chapter by chapter as TEI P2 between March 1992 and the end of 1993 (Sperberg-McQueen and Burnard, 1992–1993).

In 1999, the initial development work was concluded with the publication of a 1,292-page documentation of the definitive guidelines as the *TEI P3 Guidelines for Electronic Text Encoding and Interchange* (Sperberg-McQueen and Burnard, 1999), defining some 439 elements. With this work, the Poughkeepsie Principles were met by providing a framework for the encoding of texts in any natural language, of any date, in any literary genre or text type, without restriction on form or content and treating both continuous materials ('running text') and discontinuous materials such as dictionaries and linguistic corpora.

The advent and the success of the eXtensible Markup Language (XML)[50] as an industry standard replacing SGML from 1999 onwards called for an XML-compatible edition[51] of the Guidelines (Sperberg-McQueen and Burnard, 2002), published in 2002 by the newly formed TEI Consortium.[52]

---

[48] Initial funding was provided by the US National Endowment for the Humanities, Directorate General XIII of the Commission of the European Communities, the Canadian Social Science and Humanities Research Council, and the Andrew W. Mellon Foundation.

[49] Amongst which were groups on character sets, textual criticism, hypertext and hypermedia, formulæ, tables, figures, and graphics, language corpora, manuscripts and codicology, verse, drama and performance texts, literary prose, linguistic description, spoken text, literary studies, historical studies, print dictionaries, machine lexica, and terminological data.

[50] http://www.w3.org/TR/REC-xml [accessed 12 January 2013].

[51] The XML support was realized by the expression of the TEI Guidelines in XML and the conformation to a TEI conformant XML DTD. The TEI Consortium generated a set of DTD fragments that can be combined together to form either SGML or XML DTDs and thus achieved backwards compatibility with TEI P3 encoded texts. In other words, any document conforming to the TEI P3 SGML DTD was guaranteed to conform to the TEI P4 XML version of it. This 'double awareness' of the TEI P4 is the reason why this version was called an 'XML-compatible edition' rather than an 'XML edition'.

[52] The TEI Consortium was established in 2000 as a not-for-profit membership organization to sustain and develop the Text Encoding Initiative (TEI), http://www.tei-c.org [accessed 12 January 2013].

With this XML-compatible version of the P4 Guidelines, equal support was provided for XML and SGML applications using the TEI scheme, while ensuring that documents produced to earlier TEI specifications remained usable with the new version.

In 2003 the TEI Consortium asked their membership to convene Special Interest Groups (SIGs) whose aim could be to advise revision of certain chapters of the Guidelines and suggest changes and improvements in view of a new P5 version. With the establishment of the new TEI Council, which superintends the technical work of the TEI Consortium, it became possible to agree on an agenda to enhance and modify the Guidelines more fundamentally, which resulted in a full revision of the Guidelines published as TEI P5 (TEI Consortium, 2007). TEI P5 contains a full XML expression of the TEI Guidelines and introduces new elements, revises content models, and reorganizes elements in a modular class system that facilitates flexible adaptations to users' needs. Contrary to its predecessor, TEI P5 does not offer backwards compatibility with previous versions of the TEI.[53]

7. Back to the Metaphor

Curiously, Rodin's *La Porte de l'Enfer* is called *The Gates of Hell* in English and not the 'Gate' of Hell as the French would suggest. But actually it's a better translation. Not only are there many interpretations possible of Rodin's vision of hell, there is also more than one sculpture with that name. When Rodin failed to enter his sculpture for the 1900 Universal Exposition in Paris, he set up an independent exhibition with the painter Claude Monet (1840–1926). For this exhibition, he created a plaster model[54] of the sculpture from which most figures and groups are deliberately stripped, leaving nothing but undulations in the surface with no clear focal points. By doing this, he moved from narration to expression. Since the bronze door was state property, and was thus unsellable, he regrouped, enlarged, and cast some of the stripped figures and groups as separate, marketable works, such as The Thinker, The Kiss, Fleeting Love, and The Three Shades.[55] Others were reworked in different materials and sizes, such as Crouching Woman. The abstract character of

[53] The TEI Consortium has, however, maintained and corrected errors in the P4 Guidelines for five more years, up to the end of 2012. Since that date, the TEI Consortium has ceased official support for TEI P4, and deprecated it in favour of TEI P5.

[54] On display at The Rodin Museum, in Meudon.

[55] By that time, Rodin needed to make a living by selling the individual sculptures.

the plaster representation of the work also created a renewed interest in his work.

Also in this respect, *The Gates of Hell* is an appropriate metaphor for Humanities Computing and Digital Humanities. Over the course of time, Humanities Computing struggled with problems of self-representation and marketing itself as a discipline or a field. Finding a common practice, theoretical principles, methodology, or philosophy across humanities disciplines which employed computational techniques didn't seem to be that straightforward. The doubts about the validity of a distinctive and coherent discipline which were raised in the mid-1980s still remain 'while most of what is called "Humanities Computing" is carried out within specific Humanities subjects' (Miall, 1990, p. 3). As an applied method, Humanities Computing sold itself as an archipelago (McCarty, 2006; Chapter 5 in this volume) of humanities disciplines, as demonstrated in disciplinary organized teaching programmes, chapters in collected volumes, and strands on conferences. The (hi)story of Humanities Computing has long been the (hi)story of specific subjects, such as authorship studies, electronic textual editing, narratology, and multimedia studies or of the use of computing in broader fields such as history, musicology, lexicography, or performing arts.

The isolation of these subjects paved the way for Humanities Computing to rebrand itself with the more non-jargon-like but more abstract term 'Digital Humanities', which generated a new interest in the field, especially from the broader audience. The hermetic activity of humanities computing was replaced by a convenient hipster qualification of the humanities. The real problems of self-representation and definition, however, remained the same.

8. Self-Representation

If we know what it is that we do in Humanities Computing or Digital Humanities, the argument goes, we should be able to communicate about that research for the purpose of identifying our work, gaining acknowledgement and academic kudos, and furthering research through (interdisciplinary) collaboration and the development of advanced strategies and tools. As Melissa Terras warned us in her opening keynote address to Interface 2011, 'we should be careful what view of ourselves we are projecting into the wider academic world' (Terras, 2011a). In her DH2010 closing plenary 'Present, Not Voting: Digital Humanities in the Panopticon' (Terras, 2011b; Chapter 18 in this volume) Terras threw at us how bad we are at representing ourselves as a field and as a community. Towards the end of her lecture, she presented an agenda for the digital identity, impact, and sustainability of the field. Central to this agenda is the development of a definition of the field, the articulation of the field's relevance, success, and impact, the historical knowledge of the field as a discipline, and the preservation of the discipline's heritage.

Willard McCarty agrees with Terras that historical knowledge about and definition of the field are central issues for the awareness and self-representation of the field. 'A genuine history of the digital humanities in its first half-century', McCarty argued recently, 'would greatly help us turn pitiful laments and dull facts into the stimulating questions we should be asking now' (McCarty, 2012). More challenging than writing a historiography of humanities computing based on existing chronologies is the writing of an historical account for which the historian 'would have to locate practitioners' minority concerns within the broad cultural landscape of the time and then describe the complex pattern of confluence and divergence of numerous interrelated developments' (McCarty, 2012).

However, self-representation has long been restricted to the presentation of chronological overviews and surveys of the field. McCarty himself, for example, in explaining the title of his seminal book *Humanities Computing*, writes that it 'names a field of study and practice found both inside and beyond the academy in several parts of the world' (McCarty, 2005, p. 2). As an illustration he refers to a much longer description that is established in his and Matthew Kirschenbaum's 'Institutional Models for Humanities Computing' (McCarty and Kirschenbaum, 2003), a structured list of 'departments, centres, institutes and other institutional forms that variously instantiate humanities computing' (McCarty and Kirschenbaum, 2003, p. 465).[56]

9. Humanities Computing

This reflex to refer to a list of instantiations of what is covered by the name, and thus to provide enumerative descriptions rather than definitions, is typical for attempts to define fields of (scholarly) activities. Etymologically, definition comes from the Latin 'definitio' which literally means demarcation or fencing. A definition therefore formally freezes the meaning of a term and since Humanities Computing as a field of activity was in constant flux, a formal description was therefore impossible. This

[56] Apart from the fact that the URL in the book is broken as a result of the ALLC website's redesign, the list has been superseded by the information on digital humanities centres published by Centernet, http://digitalhumanities.org/centernet [accessed 12 January 2013]. See also Dan Cohen's list of digital humanities scholars and centres on Twitter, https://twitter.com/#!/dancohen/digitalhumanities/members [accessed 12 January 2013].

impossibility has thus been bypassed by providing enumerations of these activities, as in chronologies or overviews of the field. In the late 1980s and early 1990s such overviews have been provided, for instance for history and computing (Adman, 1987), for computing in musicology from 1966 to 1991 (Hewlett and Selfridge-Field, 1991), and for publications in *CHum* on statistical analysis of literature between 1966 and 1990 (Potter, 1991). More recently *A Companion to Digital Humanities* (Schreibman et al., 2004) published such surveys of archaeology, art history, classics, history, lexicography, linguistics, literary studies, music, multimedia, performing arts, and philosophy and religion, albeit mainly from an Anglo-American point of view.

More general surveys of developments in Humanities Computing have reflected on the community's activities for a specific purpose and are either addressed to the community itself or intended for a broader audience. In 1987, Susan Hockey briefly discussed the availability of hardware, software, textual data, and courses for the humanist, and the acceptance of computational techniques in the humanities in an assessment of the significant impact that Humanities Computing developments had on teaching (Hockey, 1987).[57] Three years later, Hockey called for a critical appraisal of the activities in the field and she advised a shift of the meta-critical emphasis from methodology to modelling in her conclusion to a chronological survey of the available and emerging tools since the early 1960s (Hockey, 1990). Ian Lancashire used his reflections on the activities in literary and linguistic computing in the period from 1968 to 1988 to develop strategies for the future. As a way of rethinking the purpose of the present by a reflection on the past, he promoted both the transformation of research into teaching, and the study of meaning as an important agenda for the future of Humanities Computing (Lancashire, 1990). As a last example I mention here the report on computers and the humanities published by the European Science Foundation in 1992 (Genet and Zampolli, 1992) which was conceived as an introduction for the research communities and policy makers in the humanities, the social, and natural sciences to the challenges and the potential of the transversal and interdisciplinary characteristics of computer-based humanities research. This book, submitted as a memorandum to the Standing Committee for the Humanities of the ESF, covers the state of the art of computing methods in humanistic research, and presents overviews of journals, institutions,

[57] Especially the rapid development of the independently operable 'microcomputers' as opposed to the 'mainframes' and the rapid increase of storage capacity revolutionized the way in which Humanities Computing interacted with teaching in higher education, e.g. with respect to searching through large collections of texts.

and projects in humanities computing, next to critical reflections on the development and future agenda of the field.[58]

Apart from the fact that such endeavours seem to approach Humanities Computing as a semantic primitive, these more or less chronological surveys all honour the implicit premise that historical knowledge about a field provides that field with its definition or at least enables the detection of theoretical, methodological, and philosophical commons which are both formative and indicative to the field under study. They mainly concentrate on performance and, apart from the concluding research agendas, avoid any involvement with predictions. In other words, they don't provide a definition of 'Humanities Computing', only surveys of activities and tools in the field.[59]

One could argue of course that by studying what is being done in the field, we may better understand it (Warwick et al., 2012, p. xiii) and that chronologies provide Humanities Computing with a definition in use or a contextual definition by chronologically reporting on its activities. The problem here is the chronology of the definition. From its definition it follows that a contextual definition may not contain the expression that is defined but must use an equivalent not containing that expression (*OED online*, definition, 4.c). In recounting the chronology of Humanities Computing, this equivalent is a virtual contraction consisting of all given contextual definitions or descriptions of the activities which are, at the moment of defining, considered as belonging to the field. McCarty's reference to the list of *Institutional Models* at the beginning of his book (2005) thus defines all elements of that list more in terms of Humanities Computing than it defines Humanities Computing in terms of its (alleged) activities.

However, even if a genuine history of the field existed, it would still need to be complemented with methodological awareness, as McCarty argues in *Humanities Computing* (McCarty, 2005). Methodology is at the basis of any transfer of knowledge about computing in the humanities, which is where Terras and McCarty locate the problem for a fruitful debate about the interdisciplinarity of the field. This is, however, not a recent problem, since Raben already pointed out in 1973 that the funding bodies' general ignorance of the methods of computing in the humanities was the greatest hindrance to its development and success: 'In their eyes, the preparation

[58] The cited overviews were an attempt to scope out the activities at a time when the field was still surveyable and when networked infrastructure and resources were inexistent.

[59] In these chronologies, it's the chronologist's intuition, not a definition of the field, which determines the scope and focus of the inventory. They argue falsely that pointing at the field provides it with a definition.

of a text seems like secretarial work, but the publication of a book comes within the definition of scholarship' (Raben, 1973, p. 5).

With his book, McCarty has tried to fill in this knowledge vacuum by attempting to 'anatomize the method of humanities computing into four perspectives: analysis, synthesis, context and profession' (McCarty, 2005, p. 6). Analysis and synthesis are the conventional methods of all humanities disciplines, the first of which is the realm of the private scholar, whereas the second is essential to the sociological role of the scholar in the academy and preferably also in the outside world. Since Humanities Computing is *in* the humanities, as McCarty (2005) sufficiently argues, its general method does not differ from those of the conventional humanities disciplines. The computational aspect offers the humanities scholar the opportunity to develop alternative analytical approaches towards the subject matter. The difference between computing *for* the humanities (instrumental) and computing *in* the humanities (methodological) is exactly the lack (in the former case) or the importance (in the latter case) of modelling as the most essential analytical method of the many forms of computing. Whereas the latter is the realm of Humanities Computing, both exist side by side in Digital Humanities.

By modelling, McCarty means the 'heuristic process of constructing and manipulating models'; a model, McCarty takes to be either 'a representation of something for purposes of study' (denotative model) or 'a design for realising something new' (exemplary model) (McCarty, 2003b; 2004, p. 255; 2005, p. 24). The purpose of modelling is never to establish the truth directly, but it 'is to achieve failure so as to raise and point the question of how we know what we know' (McCarty, 1999b), 'what we do not know,' and 'to give us what we do not yet have' (McCarty 2004, p. 255). Humanities Computing shares this methodological characteristic with, for instance, computer science, but reverses the model. Humanities Computing starts from the modelling of 'imperfectly articulated knowledge' (McCarty, 2005, p. 194), and works its way up through further steps of computational modelling till it reaches the stage of a deeper understanding of the world. Computer science, and programming in particular, starts from a real world problem and travels down to its implementation in hardware.[60]

The method shared by the humanities and computer science that Manfred Thaller and Tito Orlandi argued for in their respective defences of a 'Humanistic Computer Science' (Thaller, 2001 and 2006)[61] or

[60] See McCarty's 'Stages of modelling' (McCarty, 2005, p. 197).

[61] Thaller is professor of *Historisch-Kulturwissenschaftliche Informationsverarbeitung* which he translates as 'Computer Science for the Humanities' on his website http://www.hki.uni-koeln.de/manfred-thaller-dr-phil-prof [accessed 12 January 2013].

'informatica umanistica' (Orlandi, 2003) and which they jointly define as 'the canon (or set of tools) needed to increase the knowledge agreed to be proper to a particular field' ([Thaller], 1999, p. 25) leads them to a stronger identification with computer science than is currently acknowledged. The formalization of problems through algorithms and data representation by means of imposing structures on the data are identified by Thaller and Orlandi as central methods for computing in the humanities ([Thaller], 1999, p. 27; Thaller, 2001 and 2004; Orlandi, 2002).

This formalization of problems in particular has met with some criticism from text encoding and modelling theory. Since the humanities are not problem-oriented, Lou Burnard argues that their methodologies cannot be formalized. Instead, Burnard puts hermeneutics and text encoding at the centre of Humanities Computing, two methods that are not shared with computing or any other science. Hermeneutics is the study of interpretation that confers value on cultural objects (Burnard, 2001, p. 32). Burnard locates the starting point for the hermeneutic continuum in transcription and editing which are decisive and subjective acts of interpretation. The use of markup for the articulation and documentation of different semiotic systems in text[62] offers the humanities a single formalism that reduces 'the complexity inherent in representing the interconnectedness of all aspects of our hermeneutic analysis, and thus facilitate[s] a polyvalent analysis' (Burnard, 2001, p. 37). Text encoding in this sense is different from the industrial preparation of a text for scholarship but constitutes a new form of scholarship, as Sperberg-McQueen has argued (Sperberg-McQueen, 1991, p. 34), and which McCarty has called 'a kind of epistemological modelling' (McCarty, 2003a). The Text Encoding Initiative (TEI) provides the humanities with dedicated markup models for the articulation and documentation – that is, representation – of different interpretations of and on text, and makes explicit a theory of text in a formalization that is processable by computers. So Burnard does not argue for the formalization of the definition of problems, but for the formalization of texts and their interpretation into processable data structures.[63]

A second form of critique on what Orlandi and Thaller propose as central methods of Humanities Computing can be distilled from modelling theory which accepts a way of representing the full range of knowledge, even beyond what can be told explicitly and precisely (McCarty, 2004, p. 256) and thus beyond what can be formalized in algorithmic expressions. The problematizing purpose of modelling is furthered not only by failure

62 Burnard discusses three interlocking semiotic systems of text: text as image, text as linguistic construct, and text as information structure (Burnard, 2001, p. 33).

63 e.g. in the proposed formalization of texts in ordered hierarchies of content objects (OHCO-thesis), and the interpretation of meanings of markup.

but also by success, even when this is accidental and inexplicable from what is known at the time of constructing the model. The fluid status of modelling as experimentation away from formalization plays an important role in the perspective on scholarship as a process rather than a product.[64] The exponent of this is the model as tinkertoy[65] which denotes its playful, experimental character. Indeed 'the virtue of the noun "model" is that in computationally based research, in which the work is fundamentally experimental, it defaults to its present participle "modelling" and so to denoting that process' (McCarty, 2003b).

In their contribution to modelling theory for the humanities, Meurig Beynon, Steve Russ, and Willard McCarty pleaded for 'reappraising computing from a perspective in which experience rather than logic plays a privileged role' (Beynon et al., 2006, p. 145). In order to do so, they offered a perspective on computing they named Human Computing and defined as 'a joint collaborative activity in which devices, typically electronic, augment what is the essentially human activity of the making of meaning' (Beynon et al., 2006, p. 145). Instead of the discovery of meaning (heuristics) that Burnard proposed to realize by means of text encoding – an approach which perceives scholarship primarily as a product[66] – they shifted the central concern of computing in the humanities to the making of meaning by means of what they called Empirical Modelling.[67] By the perspective of Human Computing, computation transgresses its conventional functionality of executing algorithms that was introduced by the acceptance of the Turing machine as its primary model and in which the human and the computer are in an alternating relationship to each other. It does so by including the human in the computational activity which they described as the 'continuous engagement and negotiation of the human with the computer through the experience of the construction and behaviour of the computer model' (Beynon et al., 2006, p. 145). In this approach, the algorithmic formalization of problems as the central method of Humanities Computing is replaced by modelling of an empirical kind.

---

[64] McCarty (1999a), for instance, sees text-encoding, that is, 'rendering phenomena computable by addition of metadata that unambiguously state what it is' as fundamental to the perspective of scholarship as product. He seems to overlook here that the actual text-encoding is a transformative modelling activity which could produce a tinkertoy as well.

[65] Originally 'Tinkertoy' was a construction toy with which children could build whatever their imagination could dream up and 'learn by exercising what we now think of as "spatial intelligence"', http://www.toyhalloffame.org/toys/tinkertoy [accessed 12 January 2013]. As McCarty (2003b) noted, the term has been subsequently used 'to describe crude (or simply all physical) modelling techniques' (n. 1).

[66] Cf. McCarty (1999a).

[67] For more information on Empirical Modelling see the Empirical Modelling website at http://www2.warwick.ac.uk/fac/sci/dcs/research/em and the Empirical Modelling archive at http://www2.warwick.ac.uk/fac/sci/dcs/research/em/projects [accessed 12 January 2013].

This Empirical Modelling is supported by tools that engage with human cognitive processes and that allow for the 'experimental identification of relevant observables associated with some phenomenon and of reliable patterns of dependency and agency among these observables' (Beynon et al., 2006, p. 146). Thus, it resembles the research methods humanities scholars develop in approaching their objects of study in the personal and subjective relationship that is established between them.

Thus far, it seems that modelling in general and data representation or text encoding in particular – at least in their heuristically and epistemologically qualified meaning – are crucial methods in Humanities Computing. As crude methods of practice they are not exclusive, that is identifying methods of Humanities Computing. Therefore it is essential to distinguish text encoding as a scholarly (modelling) practice from the industrial data representation by markup which is in use, for instance, in the publishing industry. Likewise, modelling which is computational by nature should be distinguished from the ancient art of model-building.

As noted earlier, the application of modelling and data representation may be specific to Humanities Computing but they link back to the two general methods of humanities research, that is, analysis and synthesis respectively. With respect to the issue of synthesis, McCarty (2005) discusses scholarly commentary in digital editions as the most promising instantiation of synthesis in the humanities in which a certain degree of data representation, heuristics and modelling are combined into the scholarly reference work *par excellence* (McCarty, 2005, pp. 73–113).[68] Paradoxically, the characteristic which may identify the application of these common methods of humanities research as belonging to the field of Humanities Computing is indeed the computational aspect that we tried to transcend. If McCarty's book *Humanities Computing* is an attempt to provide a theory of Humanities Computing that incorporates this transgression, then it has failed to provide us with a clear, citeable, or formalized articulation of the methods of Humanities Computing which appeals to the problem of self-representation.[69] Curiously, McCarty presents exactly this problem of communication in two points in his preliminary agenda for Humanities Computing that provides the final perspective of the method of the field in his book. Without a clear description of the formal method of Humanities

[68] The idealized structure of the commentary form McCarty discusses comprises the following major parts: '(a) a scholarly introduction to the work on which a commentary is offered; (b) an edited text of that work with textual apparatus; (c) the commentary itself, in the form of paragraph-length notes keyed to the text; and (d) the usual table of contents and index' (McCarty, 2005, p. 77).

[69] McCarty's erudite and philosophical idiolect and stylistic fingerprint produces a dense and sometimes enigmatic prosaic style which takes much effort on the part of the reader to apprehend.

Computing, however, the field of activity will never succeed in providing popularized explanations of every activity and project undertaken in Humanities Computing, and in explaining and justifying what it does.

10. Digital Humanities

The same is true for Digital Humanities. The term has definitely but not definitively replaced Humanities Computing as a name for the field. There seems to be a common understanding of the term as referring to humanities research in the digital era, as opposed to traditional humanities research. However, the popularization and socialization of this new name for the field entails the risk of trivialization. The popular qualification 'digital' only relates to the technological (instrumental?) element of computation without using jargon language such as 'computer', 'computing' or 'computational'. This, however, does not solve the field's defining question, and even obscures the problem. Although Humanities Computing was a more hermetic term than Digital Humanities, it had a clearer purview. Humanities Computing relates to the crossroads where informatics and information science met with the humanities and it had a history built on the early domains of Lexical Text Analysis and Machine Translation, as we have seen above. Digital Humanities as a term does not refer to such a specialized activity, but provides a big tent for all digital scholarship in the humanities. The editors of *A Companion to Digital Humanities* who introduced the term abruptly in 2004 as an expansion of what was commonly referred to as Humanities Computing, argue that the field 'redefined itself to embrace the full range of multimedia' (Schreibman et al., 2004, p. xxiii). It still remains the question, however, whether well established disciplines such as Computational Linguistics or Multimedia and Game Studies will want to live under the big tent of Digital Humanities.

Recently Fred Gibbs, the Director of Digital Scholarship at the Roy Rozenzweig Center for History and New Media, introduced his classification of digital humanities definitions by warning that '[i]f there are two things that academia doesn't need, they are another book about Darwin and another blog post about defining the digital humanities' (Gibbs, 2011; Chapter 21 in this volume). Indeed, since Blackwell's *Companion*, a wide range of defining statements about the aims and nature of Digital Humanities (and sometimes why it differs from Humanities Computing) have been voiced.

The current volume harvests such defining contributions from journal articles and blog posts which are informed by roundtable discussions, conference panels, papers and posters, mission statements of digital humanities centres and institutes, facebook walls, and tweets,[70] and demonstrates that defining essays already constitute a genre of their own (Kirschenbaum, 2010, p. 55; Chapter 9 in this volume).[71] Among the recent and most elaborated additions to this genre is a four-part series of essays by Patrik Svensson in *Digital Humanities Quarterly*. In these essays, Svensson attempts to chart and understand the emerging field of Digital Humanities by examining the discursive shift from Humanities Computing to Digital Humanities in the first essay (Svensson, 2009; Chapter 7 in this volume), by exploring the broader landscape of Digital Humanities in the second essay (Svensson, 2010), and by discussing the cyberinfrastructure for the humanities in general and for the Digital Humanities in particular in the third one (Svensson, 2011). In the fourth essay, Svensson presents 'a tentative visionary space for the future of the Digital Humanities' (Svensson, 2012).

Concluding that a broadly conceived Digital Humanities would necessarily include humanities computing with its focus on 'the instrumental, methodological, textual and digitalized' (Svensson, 2009, p. 56) in his first essay, Svensson also acknowledges that 'the epistemic commitments and conventions of a tradition', namely Humanities Computing, 'cannot easily be subsumed in another type of digital humanities' (Svensson, 2010, p. 4). In other words, Digital Humanities is claiming a larger territory (Svensson, 2009, p. 42).[72] Svensson thus argues that both terms are non-synonymous and that the discursive transition from Humanities Computing to Digital Humanities is not just a repackaging but a broadening of scope. He adds that the term is used by the Digital Humanities community as a collective name for activities and structures in between the humanities and information technology (Svensson, 2009, p. 42).

The identification of Humanities Computing as a mainly instrumental application of computation to the text-based humanities is also present in Tara McPherson's typology of Digital Humanities. In her 2008 HUMlab lecture 'Dynamic Vernaculars: Emergent Digital Forms in Contemporary Scholarship', McPherson distinguishes among the Computing Humanities, the Blogging Humanities, and the Multimodal Humanities and sketches a certain interdependency among them. While the Computing Humanities refer to the long-lasting tradition of Humanities Computing with its focus on tools, standards, and interoperability (McPherson, 2008, 0:10:00) and

[70] Assembled on the companion website to this volume, http://blogs.ucl.ac.uk/definingdh.

[71] Volumes which gather such essays and discussions have also constituted a genre of their own. Cf. Berry (2012), Gold (2012), Lunenfeld et al. (2012).

[72] The editors of *A Companion to Digital Humanities* argue that the field 'has redefined itself to embrace the full range of multimedia' (Schreibman et al., 2004, p. xxii) with the launch of the name Digital Humanities.

the Blogging Humanities refer to networked peer-to-peer writing, mainly by non-specialist computing humanists, the Multimodal Humanities combine these and investigate the computer as simultaneously a platform, a medium, and a display device. It is in this multimodal scholarship that McPherson sees an agenda for Digital Humanities. We notice the same interdependency in Svensson's analysis where he states that '[t]here are many humanities scholars involved in what may be called digital humanities who have no or little knowledge of humanities computing, and vice versa, many humanities computing representatives who do not engage much with current "new media" studies of matters such as platform studies, transmedia perspectives or database aesthetics' (Svensson, 2009, p. 7).

Defining statements about Digital Humanities as those discussed so far commonly take references to Humanities Computing methodologies and scope as their starting point but hardly come to a definition. According to Rafael Alvarado that is because there is no definition of digital humanities. 'Instead of a definition,' Alvarado argues, 'we have a genealogy, a network of family resemblances among provisional schools of thought, methodological interests, and preferred tools, a history of people who have chosen to call themselves digital humanists and who in the process of trying to define the term are creating that definition' (Alvarado, 2011).

Therefore, Alvarado calls Digital Humanities a *social category*, not an ontological one. He is supported by Matt Kirschenbaum, who defined Digital Humanities in the 2011 Day of Digital Humanities survey as 'a term of tactical convenience' (Taporwiki, 2011). Kirschenbaum in his essay 'What Is Digital Humanities and What's It Doing in English Departments?' (Kirschenbaum, 2010; Chapter 9 in this volume) reminds us that the affirmation of Digital Humanities as the common name for the field was facilitated by the publication of *A Companion to Digital Humanities* in 2004 (Schreibman et al., 2004), the establishment of the Alliance of Digital Humanities Organizations (ADHO)[73] in 2005, the launch of the Digital Humanities Initiative by the NEH in 2006, and the publication of *Digital Humanities Quarterly* from 2007 onwards. Only recently, the Association for Literary and Linguistic Computing joined this movement by changing its name to European Association for Digital Humanities (EADH) in 2013. In a recent essay, Kirschenbaum (2012) insists on the reality of circumstances in which the term is currently used to get things done:

> At a moment when the academy in general and the humanities in particular are the object of massive and wrenching changes, digital humanities emerges as a rare vector for jujitsu, simultaneously serving to position the humanities at the

[73] http://www.digitalhumanities.org [accessed 12 January 2013].

> very forefront of certain value-laden agendas—entrepreneurship, openness and public engagement, future-oriented thinking, collaboration, interdisciplinarity, big data, industry tie-ins, and distance or distributed education—while at the same time allowing for various forms of intrainstitutional mobility as new courses are mooted, new colleagues are hired, new resources are allotted, and old resources are reallocated. (Kirschenbaum, 2012)

Not only the current use of the term, but also its origin was a moment of tactical convenience, as we learn from Kirschenbaum's 'What is Digital Humanities' essay. Apparently, Blackwell's editorial and marketing people disliked the title *A Companion to Humanities Computing* and wanted to name the volume *A Companion to Digitized Humanities*. Even the Humanistic Informatics was mentioned to cover the field, but as a compromise and to shift the emphasis away from simple digitization and complicated computing, John Unsworth suggested *A Companion to Digital Humanities* (Kirschenbaum, 2010, pp. 56–7).

11. Conclusion

As stated before, the problem of self-presentation and self-representation remains with Digital Humanities. Willard McCarty, in his concluding chapter of *Humanities Computing*, defines a preliminary agenda for the field which shows kinship with McPherson's Multimodal Humanities. Most of the items in McCarty's agenda can be related or partly related to McPherson's three types and even the big tent idea is implicitly advocated in McCarty's argument for a rapprochement between scholars and practitioners. In fact, McCarty's book demonstrates nicely that Svensson's qualification of Humanities Computing as focused on 'the instrumental, methodological, textual and digitalized' is a reductionist perception. If one book has argued against an overemphasis of the instrumental use of the computer in the humanities and has promoted computing as a meaning-generating activity building on and bringing forth models of the world, it is McCarty's *Humanities Computing*.

The question 'what it is that we are doing in Digital Humanities and how does it relate to the world', is a question which should not be eschewed. Even if it opens a can of worms, or, for the purpose of this essay, the Gates of Hell.

For the moment, we know that Digital Humanities tries to model the world around us through success and failure in order to arrive at a better understanding of what we know and don't know about humankind, their activities, artefacts, and record. And this can maybe serve as a definition of the field.

Bibliography[74]

Adman, P. (1987). 'Computers and History', in S. Rahtz (ed.), *Information Technology in the Humanities: Tools, Techniques and Applications*, Chichester: Ellis Horwood Limited, pp. 92–103.

Ager, D.E., Knowles, F.E. and Smith, J. (eds) (1979). 'Advances in Computer-aided Literary and Linguistic Research'. Proceedings of the Fifth International Symposium on Computers in Literary and Linguistic Research held at the University of Aston in Birmingham, UK, 3–7 April 1978, Aston: AMLC.

Aitken, A.J., Bailey, R.W. and Hamilton-Smith, N. (eds) (1973). *The Computer and Literary Studies*, Edinburgh: Edinburgh University Press.

ALPAC (1966). 'Languages and Machines: Computers in Translation and Linguistics', A report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Academy of Sciences, National Research Council, Washington, DC: National Academy of Sciences, National Research Council (Publication 1416), http://www.nap.edu/books/ARC000005/html.

Alvarado, R. (2011). 'The Digital Humanities Situation', *The Transducer*, 11 May, http://transducer.ontoligent.com/?p=717.

Berkowitz, L. and Squiter, K.A. (1987). *Thesaurus Linguae Graecae, Canon of Greek Authors and Works*, New York/Oxford: Oxford University Press.

Berry, D.M. (ed.) (2012). *Understanding Digital Humanities*, London: Palgrave Macmillan.

Beynon, M., Russ, S. and McCarty, W. (2006). 'Human Computing? Modelling with Meaning', *Literary and Linguistic Computing*, 21 (2), pp. 141–57.

Booth, A.D. (1958). 'The History and Recent Progress of Machine Translation', in A.H. Smith (ed.) *Aspects of Translation. Studies in Communication* 2, London: Secker and Warburg, pp. 88–104.

Booth, A.D. (ed.) (1967). *Machine Translation*, Amsterdam: North-Holland Publishing Company.

Booth, A.D. (1980). 'Computers in the University of London, 1945–1962', in N. Metropolis, J. Holett and G-C. Rota (eds), *A History of Computing in the Twentieth Century. A Collection of Essays*, New York, London, etc.: Academic Press, pp. 551–61.

Booth, A.D., Brandwood, L. and Cleave, J.P. (1958). *Mechanical Resolution of Linguistic Problems*, London: Butterworths Scientific Publications.

Booth, A.D. and Booth, K.H.V. (1953). *Automatic Digital Calculators*, London: Butterworth.

[74] All URIs accessed on 12 January 2013.

Booth, A.D. and Booth, K.H.V. (2000). 'The beginnings of MT', in W.J. Hutchins (ed.), *Early Years in Machine Translation. Memoirs and Biographies of Pioneers*, Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 253–61.

Booth, A.D. and Locke, W.N. (1965). 'Historical Introduction', in W.N. Locke and A.D. Booth (eds), *Machine Translation of Languages: Fourteen Essays*, Cambridge, MA: The MIT Press, pp. 1–14. Original publication 1955.

Booth, A.D. (1997). 'Andrew D. Booth, an Autobiographical Sketch', *IEEE Annals of the History of Computing*, 19 (4), pp. 57–63.

Bowles, E.A. (ed.) (1967). *Computers in Humanistic Research: Readings and Perspectives*, London: Prentice Hall.

Burdick, A., Drucker, J., Lunenfeld, P., Pressner, T. and Schnapp, J. (2012). *Digital_Humanities*, Cambridge, MA: MIT Press, http://mitpress.mit.edu/sites/default/files/titles/content/9780262018470_Open_Access_Edition.pdf.

Burnard, L. (1988). 'Report of Workshop on Text Encoding Guidelines', *Literary and Linguistic Computing*, 3, pp. 131–3.

Burnard, L. (2001). 'On the Hermeneutic Implications of Text Encoding', in D. Fiormonte and J. Usher (eds), *New Media and the Humanities: Research and Applications*, Oxford: Humanities Computing Unit, pp. 31–8.

Busa, R. (1949). *La terminologia Tomistica dell'Interiorità: Saggi di metodo per un' interpretazione della metafisica della presenza*, Milano: Fratelli Bocca.

Busa, R. (1950). 'Complete Index Verborum of Works of St. Thomas', *Speculum: A Journal of Medieval Studies*, XXV/1 (January), pp. 424–5.

Busa, R. (1951). *S. Thomae Aquinatis Hymnorum Ritualium Varia Specimina Concordantiarum. Primo saggio di indici di parole automaticamente composti e stampati da macchine IBM a schede perforate*, Milano: Bocca.

Busa, R. (1980). 'The Annals of Humanities Computing: The Index Thomisticus', *Computers and the Humanities*, 14, pp. 83–90.

Busa, R. (2002). *Hermeneutika e kompiuterizar. Pas gjashtëdhjetë vjetësh – L'ermeneutica computerizzata. Sessant'anni dopo – Computerized hermeneutics. Sixty years on*, Tiranë: albin.

Busa, R. (2004a). 'Analysis of scientific and philosophical texts. What differentiates them and what they have in common', in D. Buzzetti, G. Pancaldi and H. Short (eds), *Augmenting Comprehension. Digital Tools and the History of Ideas. Proceedings of a conference at Bologna, 22–23 September 2002*, London: Office for Humanities Publication, pp. 15–17.

Busa, R. (2004b). 'Foreword: Perspectives on the Digital Humanities', in S. Schreibman, R. Siemens, and J. Unsworth (eds), *A Companion to*

Digital Humanities, Malden, MA/Oxford/Carlton, Victoria: Blackwell Publishing, pp. xvi–xxi.

Doornbusch, P. (2004). 'Computer Sound Synthesis in 1951: The Music of CSIRAC', *Computer Music Journal*, 28 (1), pp. 10–25.

Doornbusch, P. (2005). *The Music of CSIRAC, Australia's first computer music*, Altona, Vic.: Common Ground Publishing.

Duro, A. (1968). 'Humanities Computing Activities in Italy', *Computers and the Humanities,* 3 (1) (September), pp. 49–52.

Eckstein, P. (1996). 'J. Presper Eckert', *IEEE Annals of the History of Computing*, 18 (1) (March), pp. 25–44.

Ellison, J.W. (1965). 'Computers and the Testaments', in G.W. Pierson (ed.), 'Computers for the Humanities?' A Record of the Conference Sponsored by Yale University on a Grant from IBM January 22–23, New Haven, Connecticut: Yale University Press, pp. 64–74.

Genet, J.-P. and Zampolli, A. (eds) (1992). *Computers and the Humanities*, Aldershot: Dartmouth/European Science Foundation.

Gibbs, F. (2011). 'Digital Humanities Definitions by Type', http://fredgibbs.net/blog/teaching/digital-humanities-definitions-by-type (Chapter 21 in this volume).

Gold, M.K. (ed.) (2012). *Debates in the Digital Humanities*, Minneapolis: University of Minnesota Press.

Goldfarb, C.E. (1990). *The SGML Handbook*, Oxford: Clarendon Press.

Heller, J. and Logemann, G.W. (1966). 'PL/I: A Programming Language for Humanities Research', *Computers and the Humanities*, 1 (2) (November), pp. 19–27.

Hewlett, W.B. and Selfridge-Field, E. (1991). 'Computing in Musicology', *Computers and the Humanities*, 25 (6), pp. 381–92.

Hockey, S. (1980a). *A Guide to Computer Applications in the Humanities*, London: Duckworth.

Hockey, S. (1980b). *Oxford Concordance Program Users' Manual*, Oxford: Oxford University Computing Service.

Hockey, S. (1987). 'An Historical Perspective', in S. Rahtz (ed.), *Information Technology in the Humanities: Tools, Techniques and Applications*, Chichester: Ellis Horwood Limited, pp. 20–30.

Hockey, S. (1990). 'Tools for Literary and Linguistic Computing: What is there? what is emerging? and what is needed?', in Y. Choueka (ed.), *Computers in Literary and Linguistic Research. Literary and Linguistic Computing 1988*, Proceedings of the Fifteenth International Conference, Jerusalem, 5–9 June 1988, Paris/Genève: Champion Slatkine, pp. 29–35.

Howard-Hill, T.H. (1979). *Literary Concordances: A Complete Handbook to the Preparation of Manual and Computer Concordances*, Oxford: Pergamon Press.

Hutchins, W.J. (1986). *Machine Translation: Past, Present, Future*, Chichester: Ellis Horwood Limited, http://ourworld.compuserve.com/homepages/WJHutchins/PPF-TOC.htm.

Hutchins, W.J. (ed.) (2000a). *Early Years in Machine Translation. Memoirs and Biographies of Pioneers*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

Hutchins, W.J. (2000b). 'Warren Weaver and the launching of MT. Brief biographical note', in W.J. Hutchins (ed.), *Early Years in Machine Translation. Memoirs and Biographies of Pioneers*, Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 17–20.

IBM (1971a). *Introduction to Computers in the Humanities*, White Plains, New York: IBM.

IBM (1971b). *Literary Data Processing*, White Plains, New York: IBM.

IBM (1971c). *Computers in Anthropology and Archeology*, White Plains, New York: IBM.

Ide, N.M. (1987). 'Computers and the Humanities Courses: Philosophical Base and Approach', in N.M. Ide (ed.), *Special Issue on Teaching Computing to Humanists. Computers and the Humanities*, 21 (4), pp. 209–15.

Ide, N.M. and Sperberg-McQueen, C.M. (1988). 'Development of a Standard for Encoding Literary and Linguistic Materials', Cologne Computer Conference 1988. Uses of the Computer in the Humanities and Social Sciences. Cologne University, 7–10 September 1988, Volume of Abstracts, pp. E.6–3–4.

Ide, N. and Sperberg-McQueen, C.M. (1995). 'The TEI: History, Goals, and Future.' *Computers and the Humanities*, 29/1: 5–15.

Jones, A. and Churchhouse, R.F. (eds) (1976). *The Computer in Literary and Linguistic Studies. (Proceedings of the Third International Symposium)*, Cardiff: The University of Wales Press.

Kay, M. (1967). 'Standards for Encoding Data in a Natural Language', *Computers and the Humanities,* 1 (5), pp. 170–7.

Kirschenbaum, M.G. (2010). 'What Is Digital Humanities and What's It Doing in English Departments?' *ADE Bulletin*, 150, pp. 55–61 (Chapter 9 in this volume).

Kirschenbaum, M.G. (2011). Comment to Alvarado (2011) on http://transducer.ontoligent.com/?p=717.

Kirschenbaum, M.G. (2012). 'Digital Humanities as/is a Tactical Term', in M.K. Gold (ed.), *Debates in the Digital Humanities*, Minneapolis MN: University of Minnesota Press, http://dhdebates.gc.cuny.edu/debates/text/48.

Lancashire I. (1990). 'Back to the Future: Literary and Linguistic Computing 1968–1988.' In Choueka, Yaacov (ed.), *Computers in Literary and Linguistic Research: Literary and Linguistic Computing,* *1988: Proceedings of the Fifteenth International Conference, Jerusalem, 5–9 June 1988*, Paris: Champion-Slatkine, pp. 36–47.

Lancashire, I. (ed.) (1991). *The Humanities Computing Yearbook 1989–1990*, Oxford: Clarendon Press.

Lancashire, I. and McCarty W. (eds) (1988). *The Humanities Computing Yearbook 1988*, Oxford: Clarendon Press.

Lancashire, I., Bradley, J., McCarty, W., Stairs, M. and Woolridge, T.R. (1996). *Using TACT with Electronic Texts*, New York: Modern Language Association of America.

Lessard, G. and Levison, M. (1998). 'Introduction: Quo Vadimus?', *Computers and the Humanities*, 31 (4), pp. 261–9.

Levison, M. (1962). 'The Mechanical Analysis of Language', *NPL*, pp. 562–74.

Levison, M. (1967). 'The Computer in Literary Studies', in A.D. Booth (ed.), *Machine Translation*, Amsterdam: North-Holland Publishing Company, pp. 173–94.

Locke, W.N. and Booth, A.D. (eds) (1965). *Machine Translation of Languages. Fourteen Essays*, Cambridge, MA: The MIT Press. Original publication 1955.

Lovelace, A.A., Countess of (1961). 'Notes by the Translator', in P. Morrison and E. Morrison (eds), *Charles Babbage and his Calculating Engines. Selected Writings by Charles Babbage and Others*, New York: Dover Publications, pp. 245–95. Original publication in R. Taylor (ed.), *Scientific Memoirs, Selections from The Transactions of Foreign Academies and Learned Societies and from Foreign Journals*, 1843.

Lunenfeld, P., Burdick A., Drucker, J., Presner, T. and Schnapp, J. (2012). *Digital_Humanities*, Cambridge (MA): The MIT Press.

Lusignan, S. and North, J. (eds) (1977). *Computing in the Humanities*, Waterloo: University of Waterloo Press.

Mauchly, K.R. (1984). 'John Mauchly's Early Years', *Annals of the History of Computing*, 6 (2), pp. 116–38.

McCarty, W. (1999a). 'We would know how we know what we know: Responding to the computational transformation of the humanities', Paper, The Transformation of Science – Research between Printed Information and the Challenges of Electronic Networks, Schloss Elmau, Max Planck Gesellschaft, 31 May to 2 June.

McCarty, W. (1999b). 'Humanities Computing as Interdiscipline', a seminar in the series 'Is Humanities Computing an Academic Discipline?' IATH, University of Virginia, 5 November, http://www.iath.virginia.edu/hcs/mccarty.html.

McCarty, W. (2003a). '"Knowing true things by what their mockeries be": Modelling in the Humanities', *Computing in the Humanities Working*

Papers, A.24, jointly published with *TEXT Technology*, 12 (1), http://www.chass.utoronto.ca/epc/chwp/CHC2003/McCarty2.htm.

McCarty, W. (2003b). 'Epistemic Tinkertoys', Unpublished manuscript.

McCarty, W. (2004). 'Modeling: A Study in Words and Meanings', in S. Schreibman, R. Siemens and J. Unsworth (eds), *A Companion to Digital Humanities*, Malden, MA/Oxford/Carlton, Victoria: Blackwell Publishing, pp. 254–70, http://www.digitalhumanities.org/companion.

McCarty, W. (2005). *Humanities Computing*, London: Palgrave.

McCarty, W. (2006). 'Tree, Turf, Centre, Archipelago – or Wild Acre? Metaphors and Stories for Humanities Computing', *Literary and Linguistic Computing*, 21 (1), pp. 1–13 (Chapter 5 in this volume).

McCarty, W. (2012). 'A Telescope for the mind?' in M.K. Gold (ed.), *Debates in the Digital Humanities*, Minneapolis MN: University of Minnesota Press, http://dhdebates.gc.cuny.edu/debates/text/37.

McCarty, W. and Kirschenbaum, M. (2003). 'Institutional Models for Humanities Computing', *Literary and Linguistic Computing*, 18 (4), pp. 465–89, http://www.allc.org/publications/institutional-models-humanities-computing.

McPherson, T. (2008). 'Dynamic Vernaculars: Emergent Digital Forms in Contemporary Scholarship', Lecture presented to HUMLab Seminar, Umeå University, 4 March, http://stream.humlab.umu.se/index.php?streamName=dynamicVernaculars.

Menabrea, L.F. (1961). 'Sketch of the Analytical Engine Invented by Charles Babbage. Translated by Ada Augusta, Countess of Lovelace', in P. Morrison and E. Morrison (eds), *Charles Babbage and his Calculating Engines. Selected Writings by Charles Babbage and Others*, New York: Dover Publications, pp. 225–45. Original publication, *Bibliothèque Universelle de Genève*, 82 (October 1842); original translation by Ada Augusta, Countess of Lovelace, in R. Taylor (ed.), *Scientific Memoirs, Selections from The Transactions of Foreign Academies and Learned Societies and from Foreign Journals*, 1843.

Miall, D.S. (ed.) (1990). *Humanities and the Computer. New Directions*, Oxford: Clarendon Press.

Mitchell, J.L. (ed.) (1974). *Computers in the Humanities*, Edinburgh: Edinburgh University Press.

Oakman, R.L. (1980). *Computer Methods for Literary Research*, Columbia: University of South Carolina Press.

Oakman, R.L. (1984). *Computer Methods for Literary Research*, 2nd edition, Columbia: University of South Carolina Press.

Oettinger, A.G. (1954). *A Study for the Design of an Automatic Dictionary*, PhD Thesis, Harvard University.

Orlandi, T. (2002). 'Is Humanities Computing a Discipline?' in G. Braungart, P. Gendolla and F. Jannidis (eds), *Jahrbuch für Computerphilologie*, 4, pp. 51–8. Also published in *Jahrbuch für Computerphilologie – online:* http://computerphilologie.uni-muenchen.de/jg02/orlandi.html.

Orlandi, T. (2003). 'Per un curriculum europeo di informatica umanistica', in D. Fiormonte (ed.), *Informatica umanistica dalla ricerca all'insegnamento*, Roma: Bulzoni Editore, pp. 19–25.

Parrish, S.M. (1962). 'Problems in the Making of Computer Concordances', *Studies in Bibliography*, 15, pp. 1–14.

Pierson, G.W. (ed.) (1965). 'Computers for the Humanities?' A Record of the Conference Sponsored by Yale University on a Grant from IBM January 22–23, New Haven, Connecticut: Yale University Press.

Polachek, H. (1997). 'Before the ENIAC', *IEEE Annals of the History of Computing*, April, pp. 25–30.

Potter, R.G. (1991). 'Statistical Analysis of Literature: A Retrospective on Computers and the Humanities, 1966–1990', *Computers and the Humanities*, 25 (6), pp. 401–29.

Raben J. (1973). 'The humanist in the Computer Lab: Thoughts on Technology in the Study of Literature', *Bulletin of the Association for Literary and Linguistic Computing*, 1 (1), pp. 3–9.

Richens, R.H. and Booth, A.D. (1952). 'Some Methods of Mechanized Translation', Presented at the Conference on Mechanical Translation, June, Massachusetts Institute of Technology. [Corrected reprint in: W.N. Locke and A.D. Booth (eds) *Machine Translation of Languages: Fourteen Essays*, Cambridge, Mass.: Technology Press of the Massachusetts Institute of Technology, 1955, pp. 24–46], http://www.mt-archive.info/MIT-1952-Richens.pdf.

Russell, D.B. (1967). *COCOA: A Word Count and Concordance Generator for Atlas*, Chilton: Atlas Computer Laboratory.

Schreibman, S., Siemens, R. and Unsworth, J. (eds) (2004). *A Companion to Digital Humanities,* Malden, MA/Oxford/Carlton, Victoria: Blackwell Publishing, http://www.digitalhumanities.org/companion.

Sparck Jones, K. (2000). 'R.H. Richens. Translation in the nude', in W.J. Hutchins (ed.), *Early Years in Machine Translation. Memoirs and Biographies of Pioneers*, Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 264–78.

Sperberg-McQueen, C.M. (1991). 'Text in the Electronic Age: Textual Study and Text Encoding with examples from Medieval Texts', *Literary and Linguistic Computing,* 6 (1), pp. 34–46.

Sperberg-McQueen, C.M. and Burnard, L. (eds) (1990). *TEI P1: Guidelines for the Encoding and Interchange of Machine Readable Texts*, Chicago/Oxford: ACH-ALLC-ACL Text Encoding Initiative, http://www.tei-c.org/Vault/Vault-GL.html.

Sperberg-McQueen, C.M. and Burnard, L. (eds) (1992–1993). *TEI P2 Guidelines for the Encoding and Interchange of Machine Readable*

Texts Draft P2 (published serially 1992–1993). Draft Version 2 of April 1993: 19 chapters, http://www.tei-c.org/Vault/Vault-GL.html.

Sperberg-McQueen, C.M. and Burnard, L. (eds) (1994). *Guidelines for Electronic Text Encoding and Interchange. TEI P3*, Oxford, Providence, Charlottesville, Bergen: Text Encoding Initiative.

Sperberg-McQueen, C.M. and Burnard L. (eds) (1999). *Guidelines for Electronic Text Encoding and Interchange. TEI P3. Revised reprint*, Oxford, Providence, Charlottesville, Bergen: Text Encoding Initiative.

Sperberg-McQueen, C.M. and Burnard, L. (eds) (2002). *TEI P4: Guidelines for Electronic Text Encoding and Interchange*, *XML-compatible edition*. XML conversion by S. Bauman, L. Burnard, S. DeRose and S. Rahtz, Oxford, Providence, Charlottesville, Bergen: Text Encoding Initiative Consortium, http://www.tei-c.org/P4X.

Stern, N. (1980). 'John von Neumann's Influence on Electronic Digital Computing, 1944–1946', *Annals of the History of Computing*, 2 (4), pp. 349–62.

Svensson, P. (2009). 'Humanities Computing as Digital Humanities', *Digital Humanities Quarterly*, 3 (3) (Summer),
http://digitalhumanities.org/dhq/vol/3/3/000065/000065.html, (Chapter 7 in this volume)

Svensson, P. (2010). 'The Landscape of Digital Humanities', *Digital Humanities Quarterly,* 4 (1) (Summer), http://digitalhumanities.org/dhq/vol/4/1/000080/000080.html.

Svensson, P. (2011). 'From Optical Fiber To Conceptual Cyberinfrastructure', *Digital Humanities Quarterly,* 5 (1) (Winter), http://digitalhumanities.org/dhq/vol/5/1/000090/000090.html.

Svensson, P. (2012). 'Envisioning the Digital Humanities', *Digital Humanities Quarterly*, 6 (1), http://www.digitalhumanities.org/dhq/vol/6/1/000112/000112.html.

Tanner, S. (1975). 'How to bring the dead language to life (Report on the ALLC International Meeting. 1974)', *Dataweek*, 22 January. Reprinted in *ALLC Bulletin*, 3 (1) (Lent), pp. 52–4.

Taporwiki (2011). 'How do you define Humanities Computing/Digital Humanities?' http://tapor.ualberta.ca/taporwiki/index.php/How_do_you_define_Humanities_Computing_/_Digital_Humanities%3F.

TEI Consortium (eds) (2007). *TEI P5: Guidelines for Electronic Text Encoding and Interchange.* Oxford, Providence, Charlottesville, Nancy: TEI Consortium, http://www.tei-c.org/Guidelines/P5.

Terras, M. (2011a). 'Peering Inside the Big Tent: Digital Humanities and the Crisis of Inclusion', http://melissaterras.blogspot.com/2011/07/peering-inside-big-tent-digital.html.

Terras, M. (2011b). 'Present, not voting: Digital Humanities in the Panopticon', Closing plenary speech, Digital Humanities 2010. *LLC,* *The Journal of Digital Scholarship in the Humanities*, 26 (3), pp. 257–69, doi:10.1093/llc/fqr016.

Thaller, M. (1999). 'Defining Humanities Computing Methodology', in K. de Smedt, H. Gardiner, E. Ore, T. Orlandi, H. Short, J. Souillot and W. Vaughan (eds), *Computing in Humanities Education. A European Perspective*, Bergen: University of Bergen, pp. 13–62.

Thaller, M. (2001). 'Bridging the Gap; Splitting the Bridge? Studying Humanities Computer Science in Cologne.' Abstract. CLiP. Computers-Literature-Philosophy, Duisburg.

Thaller, M. (2004). 'Texts, Databases, Kleio. A note on the architecture of computer systems for the humanities', in D. Buzetti, G. Pancaldi and H. Short (eds), *Augmenting Comprehension. Digital Tools and the History of Ideas. Proceedings of a conference at Bologna, 22–23 September 2002*, London: Office for Humanities Communication, pp. 49–76.

Thaller, M. (2006). 'Waiting for the Next Wave: Humanities Computing in 2006', CliP, London: King's College London, 30 June.

Toole, B.A. (1996). 'Ada Byron, Lady Lovelace, an analyst and metaphysician', *IEEE Annals of the History of Computing, 18 (3), pp. 4–12.*

Toole, B.A. (1998). *Ada, The Enchantress of Numbers. A Selection from the Letters of Lord Byron's Daughter and Her Description of the First Computer*, Mill Valley, CA: Strawberry Press.

Warwick, C., Terras, M. and Nyhan, J. (eds) (2012). *Digital Humanities in Practice*, London: Facet Publishing/UCL.

Weaver, W. (1965). 'Translation', in W.N. Locke and A.D. Booth (eds) *Machine Translation of Languages. Fourteen Essays*, Cambridge, MA: The MIT Press, pp. 15–23. Original publication 1949.

Weaver, W. (1970). *Scene of Change. A Lifetime in American Science*, New York: Scribner.

Wilkes, M.V. (1995). 'A Tribute to Presper Eckert', *Communications of the ACM*, 38 (9) (September), pp. 20–2.

Winter, T.N. (1999). 'Roberto Busa, S.J., and the Invention of the Machine-Generated Concordance', *The Classical Bulletin*, 75 (1), pp. 3–20.

Wisbey, R.A. (ed.) (1971). *The Computer in Literary and Linguistic Research. Papers from a Cambridge symposium*, Cambridge: Cambridge University Press.

Yngve, V.H. (2000). 'Early research at M.I.T. In search of adequate theory', in W.J. Hutchins (ed.) *Early Years in Machine Translation. Memoirs and Biographies of Pioneers*, Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 37–72.

Zampolli, A. (1989). 'Introduction to the Special Section on Machine Translation', *Literary and Linguistic Computing*, 4 (3), pp. 182–4.