

**TARA ANDREWS**

---

## DIGITAL HUMANITIES

---

“Digital humanities what is that, then?” This is almost invariably the first question that greets anyone working in the field, once they have told a new acquaintance what they do for a living. For those who follow discussions within the digital humanities, it quickly becomes clear that there is not a single answer to the question – that any attempt to define digital humanities is also an attempt to delineate it, and in the process someone would invariably be excluded from a field that prides itself on its inclusive nature. As digital humanities has gained in prominence since the mid-2000s in the eyes of policymakers and research funding bodies, these attempts at a definition have become both more frequent and more politically charged.

Even so, there remains no single practice that can lay an exclusive claim to the term. The label “digital humanist” might be adopted by a faculty member, a student, an adjunct university lecturer, a librarian, an academic IT support employee, or an independent scholar. The work carried out under that rubric might, correspondingly, refer to many things:

- the use of cutting-edge digital methods to find a new avenue of approach to a more or less traditional question of scholarship or research;
- the pursuit of new questions that only arise in the context of digital technology;
- critical reflection on how digital tools are used in teaching and research within the humanities;
- methods for presentation and preservation of cultural heritage in the realm of digital media;
- deeply technical programming work that intermixes humanistic theory into the computer code;
- policy advocacy for models of open access to scholarship on the Internet;
- exploration of multi-modal forms of publication of research and the difference that these forms of publication make in the impact of that research.

A comprehensive introduction to the digital humanities is a rather broad remit, and this anthology pretends to be no such thing. Its purpose is to give the newcomer some orientation as to what the field is about—to give a workable, if not an exhaustive, answer to the question “What is Digital Humanities?” To this end the material is presented in three sections. The first section includes a number of articles that help to sketch out the history of how the use of computers by researchers in the humanities developed, first into a practice known as “humanities computing”, and then re-cast in the early 2000s as “digital humanities”. The second section comprises some of the more high-profile attempts at a definition of the field and discussions of its social and professional dynamics. In the third section are collected a few online resources for the interested student or scholar who would like to gain some of the skills necessary to bring the digital humanities to his or her work.

### **A brief sketch of the development of Digital Humanities**

The story of Digital Humanities is usually held to begin in 1949 with one Father Roberto Busa, a Jesuit scholar of patristics.<sup>1</sup> Busa wished to create a **concordance** of the works of Thomas Aquinas—essentially an index of every word used by Aquinas and where in his texts it appears. Busa ventured to speak to the head of IBM, Thomas J. Watson, about whether and how computers could help. A pilot project on the poetry of Aquinas was finished in 1951, and the first volume of the full concordance was published in 1974.

By the time the concordance began to appear, however, Busa was no longer working alone. By the 1960s, computer technology had already progressed substantially beyond the level used for Busa’s first project in the late 1940s, and several other concordancing projects had been undertaken. The last non-digital concordance was published in 1966; this could be said to be the first specialty that computers rendered obsolete.

Notwithstanding the place of Busa’s work as the foundational story of the digital humanities, some roots of mechanically-inspired methodology can be traced to a time well before Turing. For example, it had already been hypothesized in the 19<sup>th</sup> century that word frequency in a text might give clues about who wrote it. Until the 1960s, due to the tedious and error-prone nature of the experiment, the hypothesis had only rarely been put to any sort of test; now such methods suddenly became feasible on a large scale. The first use of computers for settling a question of disputed authorship came in 1962, and the most famous early use was to settle the disputed authorship of twelve of the Federalist Papers.

One of the perennial characteristics of the digital humanities is the way in which our affordances shape our interests—that which we can do often determines what we want to do. As digital methods continued to be adopted throughout the 1960s and the 1970s, and as the textual data came to be collected and catalogued through initiatives such as the Oxford Text Archive, vocabulary studies began to emerge on the basis of that data—a precursor, perhaps, of the “big data” approach that is currently in vogue. The marked increase in transcription and encoding of texts that accompanied the advent of the personal computer in the 1980s led to the launch of the Text Encoding Initiative around 1988; this has been one of the primary centers of activity ever since in the field that was then known as “humanities computing”. With the invention of the CD-ROM in 1989 and the World Wide Web in 1991, the humanities computing scene in the 1990s was dominated by electronic scholarly editions of

texts, and the accompanying debates about what form, and what functions, an electronic edition ought to have. It was around this time that libraries became more involved in the world of humanities computing, which in turn began to create tension between the needs of archivists to have clear standards for data formats and the needs of scholars to maintain a certain amount of flexibility and redefinition in encoding. This tension is still present in much of the discussion surrounding the Text Encoding Initiative in particular.

However, the scholarly debates surrounding the form and function of the electronic scholarly edition, and the advantages or disadvantages of clear standards for humanistic data were also significant in that they helped to stimulate the emergence of formal theory within this so far ad-hoc field. What was humanities computing and what was it not? How might practitioners set themselves apart from scholars in the traditional humanities; conversely, how might they anchor themselves within the traditional fields?<sup>2</sup> That is to say, where and how well-defined is the intersection between academic humanities and computational affordances? This is a question that the field continues to struggle with; those who would use digital methods in their research need the necessary skills, which are usually outsourced to one or more members of a project team; at the same time, it is exceedingly rare that the leader of the project possesses or wishes to gain those skills, as they are not seen as central to the theory or methodology of the humanistic field.

In 2002 John Unsworth proposed a possible solution by defining humanities computing as "a practice of representation, a form of modeling or [ ] mimicry. It is also [ ] a way of reasoning and a set of ontological commitments, and its representational practice is shaped by the need for efficient computation on the one hand, and for human communication on the other."<sup>3</sup> His definition was an early example of a view that has become markedly more widespread in the last three to four years, but at the time, as we shall see, it was obscured by a focus on new media studies that accompanied the shift from "Humanities Computing" to "Digital Humanities".

A reflection of this attitude was easy to find in the usual name of the field, "Humanities Computing", which hinted too much at the idea of computer programmers performing a service role for "real" research. Sometime before 2003, as the publication that would become the *Blackwell Companion to the Digital Humanities* was taking shape, a new name was proposed that had lasting implications for how the field is regarded by itself and others.<sup>4</sup> The name "Digital Humanities" reflected a shift that was occurring at the time away from emphasis on computing as analysis, toward a focus on digital publication, digital media, and critical reflection on digital culture that has its roots in the post-modern cultural studies of the prior decades.<sup>5</sup>

This shift in emphasis has been labelled a "second wave",<sup>6</sup> and Katherine Hayles was perhaps the most high-profile proponent for the grounding of this wave in the field of comparative media studies.<sup>7</sup> Her argument arises from the recognition that the computer is itself a medium of communication, which by its very nature alters and influences the messages that we use it to convey. This is clear enough from the very history of the field: once texts were conceived of as discrete and computable sequences of words, it became interesting to study the decontextualized vocabulary that the words constituted; once the conversion of text into digital data became a widespread practice with the availability of personal computers, scholarly interest arose in how these conversions should be expressed in markup. As hierarchical structures lent themselves relatively easily to the capabilities of computing, so

texts came to be conceived as hierarchical structures, expressible through standards such as the XML that is now the core of the work of the Text Encoding Initiative. A result of the second wave is that it led to the perception that Digital Humanities has as much to do with scholarly communication and public engagement with scholarship as it does with scholarship itself.

The second wave of Digital Humanities has risen in tandem with the #alt-ac, or “alternative academic” movement: a push for recognition and a secure place within academia for those who are not themselves in (or pursuing) a traditional tenure-track career path, but without whom any sort of digital practice in the humanities would be next to impossible. Lack of scholarly recognition for computational work has been a handicap to those in the field ever since the early days of humanities computing. There is a widespread *de facto* perception that a person is either a scholar or a programmer; that a scholar with tenure-track aspirations in the humanities simply cannot afford the time to learn to code, or to remain overly conversant with the technological state of the art. Code does not count as publication in the eyes of funding agencies, tenure commissions, or professorial hiring committees. As a result an entire class of academic hackers has sprung up, usually without tenure, surviving on fixed-term contracts funded by grant money. There is an argument to be made that the second-wave phenomenon, arising though it did in the field of media studies, is inseparable from the rise of the #alt-ac identity and the work that the bearers of that identity have done to define themselves and their role in the academic ecosystem.

### **The Definitional Debates of the Digital Humanities**

The very shift in nomenclature, from “humanities computing” to “digital humanities”, together with the increasing diversity in academic roles that are occupied by self-identified digital humanists, has given a certain sense of urgency to the matter of a definition for the field.<sup>8</sup> Any such definition is elusive nevertheless, and perhaps not even necessary—many a department of English literature, history, and sociology has existed without shared definitive subject boundaries.

That said, it may be helpful to group the practice of digital humanities as it is today roughly into three overlapping categories. The first of these might be labeled “digital analysis”. This is the natural home of practitioners who might regard themselves as first-wave, but it has grown well beyond the concordances, authorship attribution, stylistics, and statistics that are usually (and somewhat dismissively) associated with the first wave. Authorship attribution and literary stylistics in particular are flourishing, with methods greatly refined since the first appearances of these sub-disciplines; in recent years other methods such as social networking analysis, agent-based modeling, and data mining have been used to great effect in pursuit of research goals within the humanities. For better or worse, the quantification of evidence looms large in this category of digital humanities, which has provided the substance for a great deal of criticism of its methods both by those of the second wave and by humanities scholars that do not identify themselves as “digital”—that this quantification simplifies and distorts the available evidence, removing interpretative ambiguity and closing the door to critical reflection.

A second category might be called “the digital downstream”. This incorporates many of the so-called “second-wave” digital humanists, with roots both in media studies and in the more general practice within the humanities of critical reflection on cultural phenomena. It also

incorporates digital librarians, archivists, and those whose interest is in the curation and preservation of digital culture and its artifacts. The practical aspects of this category include the creation and maintenance of online archives, digital libraries, digital exhibitions in museums, and the like. One of its concerns is to ensure the robustness and longevity of the digital data that is created; in this sense it draws in the “big infrastructure” projects both national and international that have been set up to support digital research in the humanities. It also encompasses discussions of open access, peer review, tenure credit, and the democratization and accessibility of knowledge.

There is a flourishing theoretical aspect within this category, which considers questions such as how the digital medium affects the scholarly message or the ways in which digital affordances affect the culture around us. It is also the home of the meta-reflective stance concerning digital methods in the humanities—how “evidence” becomes “data” and what is lost when it occurs, the uses and dangers of visualization as a tool of scholarly argument, and so on. Tying itself back to media studies and even sociology, this branch of the Digital Humanities also looks at how we communicate, how effective our communications are, and how much of an impact they have in this digitized world.

A third aspect of digital humanities, much smaller than the other two, arises naturally out of the meta-reflective stance toward research in the humanities that emerged from the “second wave” of Digital Humanities, but also has its roots in the “first wave”. It treats questions concerning the very nature of humanistic enquiry from a digital and scientific point of view – how we approach problem formalization, data modeling, scientific falsifiability, and empirical investigation within the humanities, and how we reconcile these methods usually found in the natural sciences with the strong hermeneutical tradition in the humanities. Indeed this may be on its way to becoming the “third wave”, as it asks, how do digital affordances change how we do research, how we treat information and evidence, or even how we argue? Whereas in the “first wave” empirics and hermeneutics were generally kept safely separated, this approach to Digital Humanities makes more of an attempt to bring them together. It concerns the aspect of humanities that is most foreign to the sciences that revolve around empirics and problem solving—the way in which scholarly interpretation of cultural evidence has a cultural value in itself, and the differing standards of argument, evidence, or even research goal that exist in different branches of the humanities.

### **Where to Start in the Digital Humanities**

A field as diverse and elusive to define as the Digital Humanities has, of course, a great many potential points of entry. For students and scholars in the humanities or staff of cultural heritage institutions who wish to become practitioners, the most direct approach is a practice-based one: to become familiar with the technology. This often means learning the basics of a programming language; although the bar to entry into the ranks of the code-literate has usually been perceived as very high, lately there has been a concerted effort through groups such as Software Carpentry<sup>9</sup> to lower this bar, and resources such as the Programming Historian for humanities scholars in particular and Codecademy<sup>10</sup> for the general public have sprung up. Many others begin with text encoding and the TEI, either through training schools that are organized periodically, or through resources such as [TEI By Example](#).

Other scholars, particularly those with roots in critical theory or comparative media studies, may come to Digital Humanities from a more theoretical approach. There are a growing number of resources that follow in the footsteps of the *Blackwell Companion to the Digital Humanities* to provide a more or less thorough grounding in the field; these include the online collection [Debates in the Digital Humanities](#) and books such as *Defining Digital Humanities*<sup>11</sup> or *Digital Humanities*<sup>12</sup> that present historically- or theoretically-grounded visions of the field.

As the field becomes established in humanities faculties around the world, degree programs on both the bachelor and master level have sprung into existence; these are excellent resources for those who encounter the digital humanities as a student. Yet perhaps the most effective point of entry, for those who are already humanists and who wish to “become digital”, is to experiment – make a database and use it, create a visual representation of a network, and most importantly: share your work with others. It is through experimentation, and it is in the adaptation of state-of-the-art techniques for new purposes, that digital humanities grows.

<sup>1</sup> For the history of Busa and the beginnings of humanities computing, see in the anthology Winter, Thomas: Roberto Busa, S.J., and the Invention of the Machine-Generated Concordance, in: Faculty Publications, Classics and Religious Studies Department, 01.01.1999. Online: DigitalCommons@University of Nebraska - Lincoln, <<http://digitalcommons.unl.edu/classicsfacpub/70>>, Stand: 17.02.2015.; for a fuller history of humanities computing up to the early 2000s, see Hockey, Susan: The History of Humanities Computing, in: Schreibmann, Susan; Siemens, Ray; Unsworth, John (Hg.): Companion to Digital Humanities, Oxford 2004. Online: Alliance of Digital Humanities Organizations, <[http://digitalhumanities.org:3030/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-2-1&toc.depth=1&toc.id=ss1-2-1&brand=9781405103213\\_brand](http://digitalhumanities.org:3030/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-2-1&toc.depth=1&toc.id=ss1-2-1&brand=9781405103213_brand)>.

<sup>2</sup> McCarty, Willard: Humanities Computing, in: Encyclopedia of Library and Information Science, 2003. Online: Maccarty.org.uk, <<http://www.mccarty.org.uk/essays/McCarty,%20Humanities%20computing.pdf>>.

<sup>3</sup> Unsworth, John: What Is Humanities Computing and What Is Not?, Forum Computerphilologie, 2002, <<http://computerphilologie.uni-muenchen.de/jg02/unsworth.htm>>, Stand: 17.02.2015.

<sup>4</sup> Kirschbaum, Matthew: What Is Digital Humanities and What's It Doing in English Departments?, in: Debates in the Digital Humanities, Minneapolis 2012. Online: The Graduate Center, City University of New York, <<http://dhdebates.gc.cuny.edu/debates/text/38>>, Stand: 08.10.2015.

<sup>5</sup> For complementary views on this phenomenon in this anthology: Burdick, Anne; Drucker, Johanna; Lunefeld, Peter u. a.: A Short Guide to the Digital Humanities, in: Burdick, Anne; Drucker, Johanna; Lunefeld, Peter u. a. (Hg.): Digital humanities, Cambridge, MA 2012, S. 121–135. Online: Library of Congress, <[http://jeffreyschnapp.com/wp-content/uploads/2013/01/D\\_H\\_ShortGuide.pdf](http://jeffreyschnapp.com/wp-content/uploads/2013/01/D_H_ShortGuide.pdf)>, Stand: 08.10.2015.

<sup>6</sup> See in this anthology: Jones, Steven E.: The Emergence of the Digital Humanities, New York 2014. Online: Tumblr, <<http://emergenceofdhbook.tumblr.com/>>, Stand: 17.02.2015.

<sup>7</sup> Hayles, N. Katherine: How we think. Digital media and contemporary technogenesis, Chicago 2012.

<sup>8</sup> See for example in this anthology: Nowviskie, Bethany: Eternal September of the Digital Humanities, nowviskie.org, 15.10.2010, <<http://nowviskie.org/2010/eternal-september-of-the-digital-humanities/>>, Stand: 17.02.2015.; Ramsay, Stephen: DH Types One and Two, <<http://stephenramsay.us/2013/05/03/dh-one-and-two/>>, Stand: 13.10.2015.

<sup>9</sup> Software Carpentry, <<https://software-carpentry.org/>>, Stand: 13.10.2015.

<sup>10</sup> Web Developer Skills, Codecademy, <<https://www.codecademy.com/learn>>, Stand: 13.10.2015.

<sup>11</sup> Terras, Melissa; Nyhan, Julianne; Vanhoutte, Edward (Hg.): Defining Digital Humanities. A reader, Farnham 2013.

<sup>12</sup> Burdick, Anne; Drucker, Johanna; Lunenfeld, Peter u. a.: Digital\_Humanities, Cambridge, MA 2012.