

2 Genealogies of the Digital Humanities

In its early days, humanities computing tended to concentrate on supporting humanities-technical projects through the application of technical knowledge and experience. This sometimes meant that it seemed as if the ‘real’ humanities scholars drove the intellectual side of the project and the humanities computing technicians implemented it.¹ Even though the digital humanities is a relatively recent disciplinary formation and it draws its history from ‘computing in the humanities’, or ‘humanities computing’, it also incorporates a number of strands both theoretical and technically oriented from across a range of disciplinary areas (see Schreibman et al. 2004; Berry 2012a; Gold 2012).² Scheinfeldt (2008) argues that ‘the digital humanities family tree has two main trunks, one literary and one historical, that developed largely independently into the 1990s and then came together in the late-1990s and early-2000s with the emergence of the World Wide Web’.³ This is a useful way to understand the debates and arguments that are manifest in the field and, to a remarkable extent, map onto this division. In this chapter, although we focus mainly on digital humanities, it is important to remember the complementary development of and interconnectedness with work in digital history (see Gregory 2014).

The foundational work in digital humanities was generally concerned with the application of the computer to the textual materials of the humanities (see Feeney and Ross 1993): what has been described as treating the ‘machine’s efficiency as a servant’ rather than as a ‘participant enabling of criticism’ (McCarty 2009). The term ‘digital humanities’ is of much more recent origin, and the field’s previous incarnation was humanities computing, which was strongly associated with tools and archival work, to the extent that humanities computing soon became understood as a service department providing specialized computing support work and technical assistance to other humanities departments (see Flanders 2011; Nyhan 2012).

The workings of digital technologies, systems, encodings, processes and projects have been a key concern of digital humanities from its very beginning. This is broadly known within the digital humanities as knowledge representation (see Sowa 2000; Schreibman et al. 2004: xxv). This is also connected to the remediation of culture that digitization presents and that ‘the whole of our cultural inheritance has to be recurated and reedited in digital forms and institutional structures’ (McGann 2014: 1). Indeed, Schreibman et al. (2004: xxvi) argue that the process of knowledge representation ‘requires humanists to make explicit what they know about their material and to understand the ways in which that material exceeds or escapes representation’, and the fact that the process digital humanists follow ‘to develop, apply and compute these knowledge representations is unlike anything that humanities scholars, outside of philosophy, have ever been required to do’. Whilst we might take issue with the claim that other humanities scholars were not reflexive about their categories and knowledge frameworks, it is, however, true that the explicitly calculable requirements of much digital

humanities work on encoding encourages a very different relationship with knowledge and information. This calculability is linked to the discovery of ‘patterns and connections in the human record that we would never otherwise have found or examined’, making an implicit reference to the scaling-up in the quantity of data that computation makes possible (Schreibman et al. 2004: xxvi). As Ramsay (2004: 195) argues, ‘dealing with patterns necessarily implies the cultivation of certain habits of seeing . . . [and] of all the technologies in use among computing humanists, databases are perhaps the best suited to facilitating and exploiting such openness’ towards seeing.⁴ This is taken together with the aggregation, analysis and presentation of particularly visual forms of analysis, such as data visualization, charts and so forth (see Warwick 2004: 375; Drucker and Nowviskie 2008). This points to the key aspects of acquisition, encoding, processing and representation that make up what we might think of as the digital humanities workflow. It is this explicitly technical orientation that informs both digital humanities and the way in which it tends to reflect on its activities.

It is helpful to think about these issues in terms of a genealogy of the digital humanities that tends to shun its historical antecedents, and hence think through some of the ways in which it has been previously articulated – but also to think through analytical classifications, such as that introduced by Stephen Ramsay, who discussed digital humanities ‘type 1’ and ‘type 2’ (Ramsay 2013b), or by David Golumbia, who proposed ‘narrow’ or ‘broad’ digital humanities (Golumbia 2013). We also want to set up the context for the later argument that digital humanities needs to widen and deepen its cultural critique (Liu 2011), and introduce design and the medium-specific work of software studies and media studies (Berry 2014).⁵ Indeed, digital humanities has also developed a public engagement role and, as Liu argues, digital humanities is ‘ideally positioned to create, adapt, and disseminate new methods for communicating between the humanities and the public’, but also to ‘create technologies that fundamentally reimagine humanities advocacy’ (Liu 2012: 496–7). Today, the field is much more multi-faceted and, as we argue in this book, the potential for cultural critique of computational culture offers new possibilities for the field of digital humanities (Liu 2012: 496–7). In this book, we want to move beyond linear narratives of digital humanities development and make connections between digital humanities and other disciplinary formations, both historically and in terms of important influences. We also want to link these to the contemporary milieu in which digital humanities is expanding rapidly and connecting to other areas and approaches. But, first, we briefly turn to look at the foundation of the digital humanities, as computing in the humanities.

Digital humanities often refers back to a disciplinary ‘origin story’ associated with humanities computing and its ‘founding’, which has become something of a taken-for-granted explanation of where digital humanities emerged from. This tends to be articulated as stemming from the work of Roberto A. Busa, a Jesuit priest and professor of philosophy. Busa developed a list of concordances of St Thomas Aquinas’s work using IBM computers in 1949, resulting in the 56–printed-volume *Index Thomisticus* as an analysis of the work of Thomas Aquinas (Schreibman et al. 2004; see also [chapter 4](#) below).⁶ For this reason Busa is often seen as the starting point for thinking about the history of digital humanities, sometimes giving the impression that Busa was working in a vacuum. It is laudable therefore that Busa himself argues, ‘although some say

I am the pioneer of the computers in the humanities, such a title needs a good deal of nuancing . . . isn't it true that all new ideas arise out of a milieu when ripe, rather than from any one individual?' (Busa 1980: 84).⁷

In the case of Busa, there is a clear link made between the affordances of computation and humanities work.⁸ Thus, very early on, a connection was made between brute-force computation as a tool of humanities research and the transformation of historical or literary documents into a computerized archive or database (see Nyhan 2012). As McCarty documents, in 1965 the last handmade concordance was published (by Ione Dobson, about Byron), and in 1966 'the American scholar Joseph Raben founded the first professional journal in the field, *Computers and the Humanities* (CHum).' (McCarty 2003: 1226). The connections between the human-generated work and the move to automated systems for collecting and collating humanities materials is useful for thinking about the shift from a predominantly print culture to a digital one. This could be said to have necessitated a new conceptual language of tools and archives. This early conceptual distinction between tool and archive, program and data, continues as an important conceptual shorthand for the field. Indeed, computation was strongly linked to its new storage capacity and the possibility of comprehensive archives and databases as a result of digitization, together with the computational power of textual searching, filtering, concordance production and manipulation that was made possible. This led to the creation of support systems and semi-standardized software for the archiving and maintaining of textual repositories (Hockey 2004: 10).

Tools and archives continue to be key concepts for understanding the technical contributions of the field but they also point towards a disciplinary specificity of digital humanities through a sense of 'community' around these tools and archives. This has remained remarkably consistent in the years since Busa's early work and has marked the field in particular ways. For example, Koh (2014) argues that 'the social contract of the digital humanities is composed of two rules: 1) the practice of *civility*, or "niceness"; and 2) *possession of technical knowledge*, defined as knowledge of coding or computer programming. These rules are repeatedly cited within the digital humanities community, but they are at the same time contested and criticized' (2014: 94). These basic norms are key for creating the fundamental elements for social interaction and organization. The cultural norms of technical knowledge in terms of knowledge, representation and elements of programming (understanding, if not necessarily doing), combine with the shared meaning generated through collective interpretative schemas that inform and define their interaction. In many ways, as the taken for granted of 'humanities computing' has changed into digital humanities, it was inevitable that the expansion of disciplinary concerns would generate new challenges and sources of conflict and contestation. Indeed, there has been a great amount of internal debate and contestation in the field recently, for example. Scheinfeldt (2008) has argued that we are in '[a] moment of change right now, that we are entering a new phase of scholarship that will be dominated not by ideas, but once again by organizing activities, both in terms of organizing knowledge and organizing ourselves and our work . . . collaborative encyclopedism, tool building, librarianship'. In practice, however, digital humanities has not developed a static set of ideas and concepts which can enable organization building in this manner. Rather, it continues to be an extremely

flexible area for debates over knowledge, computation and the humanities, as evidenced by recent work such as Allington et al. (2016). In our opinion, this is a sign of growing strength in the digital humanities.

In digital humanities, the notion of scholarly work tends to be wider than in normalized forms of academic output, such as monographs and academic papers. Rather, digital humanities scholars are keen to stress their non-standard concerns with producing encodings, digital versions, databases, ontologies, metadata and visualizations as new forms of academic knowledge production made possible by computation. This aspect of digital humanities is much discussed in the field due to its implications for academic standing, tenure and promotion. This links very clearly to digital humanities' own self-perception as an 'outsider' field, and, whatever the merits of these claims, it is interesting that computation has, for the most part, been more influential at the encoding phase of knowledge production. Of course, this is changing as computation as a reading medium gains ground through both web technologies and mobile devices, and new forms of digital screens, and hence on the form of the reading public in relation to the monograph or academic paper. However, it cannot be denied that digital humanities has been at the forefront of identifying and problematizing the way in which the materiality of paper is both crucial to, and a problem for, the humanities as a whole.

The introduction of computing to the humanities has not traditionally been met with enthusiasm by researchers and scholars. There have been debates on the use of computers in the humanities ever since they were introduced and 'even today there are pockets of stubborn resistance against computing. At the same time, we can see that, although basic computing skills of word processing, e-mailing and web browsing are nowadays omnipresent among humanities scholars, their methodical and technical skills for computerised research are fairly limited' (Boonstra et al. 2004: 13). Disciplines such as linguistics have been somewhat more welcoming to computation, and have worked with computation for a considerable period of time from the 1950s, although until more recently it has tended to be more 'occasional' than methodologically central (see Feeney and Ross 1993; Kaltenbrunner 2015: 3).

Indeed, some other important milestones in the digital humanities were the electronic editions of Christian Latin texts produced by the CETEDOC research centre in Louvain in the 1970s, the founding of the previously mentioned professional journal *Computers and the Humanities* by Joseph Raben in 1966 (Gilmour-Bryson 1987), and the Philadelphia Social History Project (PSHP) (Hockey 2004: 5; Kladstrup 2015). But also the 'first international professional body, the Association for Literary and Linguistic Computing (ALLC), was founded in 1972, followed by the Association for Computers and the Humanities (ACH) in 1978, and the Association for History and Computing in 1987. The *ALLC Bulletin* began in 1973, becoming *Literary and Linguistic Computing* (LLC) in 1986' (McCarty 2003: 1227).

However, there was generally not a large amount of work done on computing from within the humanities themselves due to lack of interest and expertise – indeed, 'scholars in the humanities had to rely on help mainly from computer linguists. A great deal of the activity therefore centered around source-editing' (Boonstra et al. 2004: 26). This particularly textual dimension remains strong within the disciplinary focus of the digital humanities and makes

clear the links with the humanistic disciplinary object of analysis.⁹ As Hockey explains, ‘a glance through the various publications of this period shows a preponderance of papers based on vocabulary studies generated initially by computer programs . . . of interest for some kind of stylistic analysis or for linguistic applications’ (Hockey 2004: 9–10). This remained so with the foundation of the Association for Computers and the Humanities (ACH) in the 1970s, a major professional society for the digital humanities that supports and disseminates research and cultivates a professional ‘community’ through conferences, publications and so forth.

An early counter-example emerged in Germany where, in 1975, German historians and sociologists founded the Quantum-group in order to explore, in close collaboration, possibilities and problems with the use of historical and process-produced data. This was ‘driven by a feeling of uneasiness in empirical sciences with data based on surveys only, and by the turn of historians away from ideographic and narrative approaches. It was aimed at closing the gap between the German situation and the upswing of quantitative history elsewhere’ (Boonstra et al. 2004: 26). Its journal, *Quantum Information*, later changed its name to *Historical Social Research (Historische Sozialforschung)* and this grew into a broader publication that covered subjects which related to history and computing, whilst nonetheless focused on the computational aspects of historical research (see Boonstra et al. 2004: 26).

Nonetheless, the larger linguistic and literary influence on digital humanities can be felt in the field’s relation to the Text Encoding Initiative (TEI), published as a set of guidelines in 1994. TEI is essentially a markup language drawn from SGML (Standard Generalized Markup Language), similar to XML. TEI was the ‘first systematic attempt to categorize and define all the features within humanities text that might interest scholars’ (Hockey 2004: 12). Much of this work of humanities computing was focused on building archives, infrastructure and digital tools and included the idea of ‘comprehensive’ and ‘exhaustive’ archives using XML, usually as TEI markup, in textual works, and the creation of metadata around physical artifacts. With the rise of the internet and the World Wide Web in the 1990s, together with its own markup language HTML, there has been some divergence between the TEI and other encoding practitioners but in any case, as McGann argues, ‘organizing our received humanities materials as if they were simply information depositories, computer markup as currently imagined handicaps or even baffles altogether our moves to engage with the well-known dynamic functions of textual works’ (2004: 199–201).

The encoding aspect of digital humanities remains crucially important and, as Cordell has argued, ‘you may find encoding or archival metadata development boring or pedantic – certainly some do – but you cannot pretend that encoding is less a part of the digital humanities than coding’ (Cordell 2014). These encoding practices can also be traced back to earlier work in textual computation but also innovative email discussion lists like *The Humanist* set up in 1987 by Willard McCarty (McCarty 2003: 1227; Hockey 2004: 11). As McCarty documents, ‘conversing in groups such as *Humanist*, as well as one-to-one, provides an obvious means for establishing a loose, dynamic sense of the field and the relative importance of contributions to it’. For example, McCarty explored the possibility of a research agenda particular to humanities computing, first articulated as ‘methodology’, then later as a concern with ‘models’

and ‘simulation’ (see McCarty 2004, 2009, 2013a). Indeed, this has been mirrored in Kirschenbaum’s (2013) notion of ‘methodological outlook’ and his focus on forensic approaches to media (Kirschenbaum 2008). This is sometimes articulated as an ‘exploratory methodology, where the researcher or student is encouraged to explore materials, datasets or issues in an experimental fashion’ (Svennson 2010).¹⁰ This notion of experimentation is a key issue and relates also to the notion of a humanities lab that we will examine later.

The importance of knowledge representation within the digital humanities remains an absolutely key thread that continues to inform debate and contestations over what a digital humanities project is, for example regarding the rise in use of XML and the TEI. This notion of knowledge representation as the encoding of cultural forms in a computational wrapper, usually a form of metadata such as a Shakespeare play encoded into XML (see the Folger Shakespeare Library for exemplary versions of this), is very particular to the digital humanities. Even today, this interest in encoding objects remains a core part of the identity of digital humanities and its activities (Fitzpatrick 2012: 13). Indeed, on the importance of encoding, Cordell (2014) wrote, ‘any vision of digital humanities that excludes or dismisses the close and careful work of digital preservation, editing, and publication is simply false’ (see also Deegan and Tanner 2004: 488–93; Smith 2004: 577). We might also add that any history of the digital humanities that does not identify this element is also seriously misrepresentative of the field’s development. Indeed, as Renear argues, TEI is itself now a research community around the subject of encoding, and connects together many professions, disciplines and institutions across the globe. Its practitioners define themselves through a shared interest in encoding, concepts, tools and techniques. They tend to concentrate on textual discourse, with the focus on improving general theoretical understanding of textual representation, and the practical goal of using that understanding to develop further methods, tools and techniques to support practical applications in publishing, archives and libraries. TEI practitioners have sophisticated understanding of knowledge representation systems (such as formal semantics and ontology, object orientation methodologies, etc.), and offer new ways of theorizing them (non-hierarchical views of text, antirealism, hypertext and so forth) (see Renear 2004: 235).

Since 2001, humanities computing has become known as digital humanities, a relatively new term but one which draws on this older history of computing in the humanities. Unsworth remarked in 2001 that the reason for naming the first digital humanities programme at Université Laval, Québec, was because ‘the name of the program (“Digital Humanities”) is a concession to the fact that “Humanities Informatics” (which would probably be a more accurate name) sounds excessively technocratic, at least to American ears. The other obvious alternative – “Humanities Computing” – sounded too much like a computer support function’ (Unsworth 2001).

In this book, we accept this definition and understand the field as concerned with a certain computational ‘way’ of doing humanities research, which we survey and explore. Indeed, Kirschenbaum (2012a: 9) describes the term ‘digital humanities’ as a floating signifier ‘that increasingly serves to focus the anxiety and even outrage of individual scholars over their own lack of agency amid the turmoil in their institutions and profession’, but which is nonetheless a ‘scholarship and pedagogy that are bound up with infrastructure in ways that are deeper and

more explicit than we are generally accustomed to'.¹¹ For example, debates and critiques over the nature of the discipline of digital humanities circulate encoded as “‘DH” – variously also dh/DH/D_H/#dh as well as #transformdh and #dhpoco’ (Kirschenbaum 2014: 51; see also Lothian 2011).¹² Nonetheless, even as the term ‘digital humanities’ has solidified and entered into more general usage, we are keen to acknowledge that digital humanities is, and remains, a contested term. But the question of the digital humanities is not solely a symbolic or discursive one. Digital humanities is also symptomatic of wider changes taking place in the university as it increasingly informatizes its systems and begins to introduce computation into wider use. These infrastructural changes and requirements for undertaking research are key to deeper understanding of the positive benefits the digital humanities can bring to the culture of the humanities in research universities.

McCarty (2003) has usefully conceptualized digital humanities as a concern with the building of ‘methodological commons’ and, through these resources, enabling other possibly ‘less digital’ scholars to pick up and use the tools and archives to enable the search for ‘patterns’ and develop techniques of ‘modeling’ and transforming humanities into much more of an ‘experimental field’. By modelling, McCarty (2004) points to the ‘heuristic process of constructing and manipulating models, a “model” [he] takes to be either a representation of something for purposes of study, or a design for realizing something new’ (2004: 255, emphasis removed).

Using a metaphor of a mercantilist market, he argued that the ‘humanities computing specialist acts as merchant trader of these intellectual goods, seizing opportunities for importing and exporting them as the occasion warrants . . . From his or her perspective, the various disciplines serve as laboratories in which these goods are exercised, probed, and improved upon’ (McCarty 2003). McCarty categorizes the forms of approach used in digital humanities through the search for ‘patterns’ across digital humanities ‘branches’: (i) algorithmic: development of software for the analysis of source materials, to focus on mechanical elements in the analysis of data, as well as to allow large quantities of data to be processed and the specified patterns in them found; (ii) metalinguistic: computationally rigorous linguistic analysis; (iii) metalanguages by which computationally elusive entities may be tagged in texts and thus reliably processed; (iv) representational: focuses on arranging, formatting or otherwise transforming the appearance of data.

In this construction, digital humanities ‘is an epistemological practice of its own, not needing the *imprimatur* of theory to proceed’, and ‘need not wait on the emergence of a theoretical framework, . . . its semidirected, semicoherent activities are no discredit, rather the norm for an experimental field’ (McCarty 2003; cf. Berry 2012a; Sterne 2015). This draws on Hacking’s notion of an epistemology by intervention (e.g. using a microscope, staining, injecting). But it also disavows a theoretically informed practice, whether historically or otherwise. As McCarty argues, ‘research questions for humanities computing arise from [one’s] involvement [in various disciplines] and so a research agenda and all that goes with it’ (2003). However, this formulation leaves the research questions *for* the digital humanities to be defined elsewhere. McCarty describes this as ‘serving client disciplines, which tend to initiate collaborations, set the agenda for research and take academic credit for the result’

(McCarty 2012: 117). Hence, digital humanities becomes a shared set of methods, tools, archives, across the humanities, without a disciplinary focus of its own, beyond that of building better tools and archives (creating tools for ‘intervening’). Thus, ‘digital humanities’, although a contested term, becomes used by a growing number of researchers to describe the broad kinds of research work they undertake, or even a tactical or discursive construction (Kirschenbaum 2013, 2014).¹³

Consequently, there remains a theoretical and disciplinary anxiety in digital humanities, expressed by McCarty (2010) as ‘our little and fragile field: [it has a] tendency to industrialization’. While digital humanists may develop tools, data and metadata critically e.g., debating the ‘ordered hierarchy of content objects’ principle; disputing whether computation is best used for truth finding or, as Samuels and McGann put it, ‘deformance’ (1999) – ‘rarely do they extend their critique to the full register of society, economics, politics or culture’ (Liu 2012), or indeed to the very tools with which they undertake their research activities. The key to the argument of this book is not that one should choose one or the other, but rather that the discipline is strengthened by having scholars that specialize in and can move between digital humanities projects and digital humanities theory.

Until quite recently, digital humanities projects have involved using the machine to support research practice rather than as a critical tool in its own right (McCarty 2009). The computer as archive or search tool has been relatively accepted by the humanities as a means to undertake research. But these tools and resources haven’t generally been accepted as scholarship in and of itself, and certainly there has not been enough work on absorbing and developing the approaches that enable the digital aspect of the digital humanities to be used as a means of critique, in and of itself. Indeed, when it comes to critical work in the humanities, the gold standard remains the monograph. Thus, there continues to be a boundary between the kinds of work produced between these worlds of humanities (academic) and the technical (tools). This division of academic knowledge from support echoes that found in industry where information management functions are often separated into a service function for the corporation. Among the oddities of this conception are that there is often already a computing support department in most, if not all, universities and that the digital humanities are usually inhabited by academics and research-oriented staff. This historical division may help explain some of the way in which certain activities were rewarded as academic (monographs/promotion), and others were seen as ‘merely’ academic-related or technical (tools, archives). But it is notable that computer science has successfully distinguished itself from other computing functions within the university and produces its own form of knowledge output that is understood as discipline-specific.

In digital humanities, both its recent and historical instantiations have tended to concentrate on text-focused work which has given the field a focus and a scholarly concern that makes its work recognizably digital humanities (see Schreibman et al. 2004: xxiii). That is not to say that digital humanities work is purely textual in nature – far from it, as many projects demonstrate – however, we could argue that text has a paradigmatic function in digital humanities work. Nonetheless, the discussion over the sustainability of these kinds of digital humanities has been a constant refrain, particularly in relation to the fragility of digital works, especially located on

webservers and cloud-based platforms (see Pitti 2004: 471–3).

In a previous work, Berry (2012a) argued that we might analytically divide the digital humanities: 1st Wave – Computing in the Humanities, computer archives and tools (1940–2001); 2nd Wave – Digital Humanities (DH), interfaces and the born digital (2002–9); 3rd Wave – digital humanities (dh), materiality and cultural critique (2009–). The aim of this division is not to create closed concepts; rather, the aim is to develop analytical time periodization that is paradigmatic, or exemplars around which digital humanities work can be clustered during specific moments. Indeed, there was a lot of cross-over between these different modes of digital humanities activities, and they continue to inform each other.

This division broadly follows Schnapp and Presner’s (2009) identification of the first two waves of digital humanities. The first, in the late 1990s and early 2000s they argue, tended to focus on large-scale digitization projects and the establishment of technological infrastructure. The second wave, which they inevitably call ‘Digital Humanities 2.0’, was, they argue, generative, creating the environments and tools for producing, curating and interacting with knowledge that is ‘born digital’ and lives in various digital contexts (see also Davidson 2012: 476). These are different models for thinking about the trajectory of digital humanities projects and for helping us to understand the temporal and methodological grounds for the various aspects of digital humanities projects that continue to be debated in digital humanities today.

Davidson further argues that ‘Humanities 2.0 [or multimodal humanities] is distinguished from monumental, first-generation, data-based projects not just by its interactivity but also by openness about participation grounded in a different set of theoretical premises, which decenter knowledge and authority’ (Davidson 2012: 711). This trajectory ‘can be traced by comparing Unsworth’s 2002 “What Is Humanities Computing and What Is Not”, with Schnapp and Presner’s 2009 “Manifesto 2.0”. At the top of Unsworth’s value hierarchy are sites featuring powerful search algorithms that offer users the opportunity to reconfigure them to suit their needs . . . By contrast, the “Manifesto” consigns values such as Unsworth’s to the first wave, asserting that it has been succeeded by a second wave emphasising user experience rather than computational design’ (Hayles 2012: 44).

Berry has argued that digital humanities should orient itself towards a third wave of digital humanities that is critically reflexive about its engagement with computation / the digital as an object of research, as part of the problematic of humanities scholarship in the digital age (Berry 2012a, 2014). This is broadly in line with Alan Liu’s work, which suggests that digital humanities is ideally placed to work with new-media scholars and particularly with the questions raised by cultural criticism and engagement with wider publics. Berry has argued elsewhere this could represent a positive turn towards a more critical digital humanities. Drucker similarly argues that the ‘challenge is to take up . . . theoretical principles and engage them in the production of methods, ways of doing our work on an appropriate foundation. The question is not, does digital humanities need theory?, but rather, how will digital scholarship be humanistic without it?’ (Drucker 2012: 94).

In order to help explain the historical conjunction of digital humanities as different from cognate but nonetheless distinct fields, Ramsay (2013b) argued that there were two types of

digital humanities. Type 1 digital humanities saw digital humanities as archives/tools, ‘building’, ‘more hack, less yack’. The ‘hack’, as it is sometimes called, ‘refers to the concept of making: this might include developing a game, prioritizing a pedagogical practice that includes teaching students how to develop a tool, coding an archival project, soldering circuits together for an art installation, or developing a new tool for scholarly publishing’ (Ramsay 2013b). This is often contrasted, both in the literature and in everyday discussions of digital humanities, with the view that “‘yack” is perhaps . . . recognizable to most humanities scholars. It refers to the acts of theorization by writing or speaking, usually in solitary exercises, or at times on a conference panel or scholarly collaboration’ (Barnett 2014: 74).

Ramsay argues that Type 1 digital humanities (DH) is ‘united not by objects of study, per se, but by a set of practices that most regarded as intimately related: text encoding, archive creation, text analysis, historical gis, 3d modeling of archaeological sites, art historical cataloging, visualization’, and general meditation on what all of these new ‘affordances might mean for the study of the human record’ (Ramsay 2013b). He explained that ‘Type 1 DH is [a] community’ and ‘in early 2001, this community fatefully decided to call itself “digital humanities”’, as humanities computing sounded like a ‘campus technical support group’. Ramsay argues that ‘digital humanities’ is ‘useful because it distinguished our activity from media studies’.¹⁴

With Type 2, on the other hand, Ramsay argues that ‘I don’t know exactly how it happened . . . Media studies practitioners were digital humanists; people who had devoted several decades to digital pedagogy were digital humanists; cultural critics who were interested in Internet culture were digital humanists; and digital artists of a certain variety were digital humanists.’ The resultant confusion of disciplinary identity, for Ramsay ‘sounds like the recreation of the humanities itself after some technological event horizon’ (Ramsay 2013b). Type 2 digital humanities, then, is a more expansive notion, including media theory, cultural critique, media and communications, etc.

Ramsay has also reconceptualized the key approaches to digital humanities that are being informed by the notion of ‘more hack, less yack’ as ‘building things’ – sometimes standing in for the requirement that one ‘codes’ to be a digital humanist. Critiquing this distinction, Golumbia (2013) termed them the ‘narrow’ (DH) or ‘big tent’ (dh) digital humanities. He did this to draw attention to the limited way in which digital humanities tended to be used, but also to how it would be deployed differently depending on the circumstances of its use – thus, for funding applications the narrower definition would tend to be used, but when called upon to reflect on social and political questions, a wider notion might be deployed (see also Koh 2014: 103). Golumbia argued that there are important implications for digital humanities regarding widening of the research field to other disciplinary and critical research broadly understood.¹⁵

Without failing to acknowledge the excellent critical work that is going on in the field of Digital Humanities (see Schreibman et al. 2004; Gold 2012), we argue that more support for critical work in digital humanities from funders, and indeed from digital humanities projects, is needed. This could take the form of better support for critical/reflexive dimensions in digital

humanities-funded projects – not just connecting to interdisciplinary work in computer science/engineering and so on, but also with more critical scholarly academics (see Gold 2012). These are important issues because they help us, as digital humanists, to think through the nature of our work, particularly in relation to how that work is conceptualized and understood, but also the metaphors and analogies that are invoked in that understanding. The selfdefinition of digital humanists sometimes treads on the uneven lines between metaphor and problematic constructions of the future of the humanities – for example, Jockers argues: ‘what is needed now is the equivalent of open-pit mining or hydraulicking . . . Close reading, traditional searching, will continue to reveal nuggets, while the deeper veins lie buried beneath the mass of gravel layered above. What are required are the methods for aggregating and making sense out of both the nuggets and the tailings [. . .] [to] exploit the trammel of computation to process, condense, deform, and analyze the deeper strata from which these nuggets were born, to unearth, for the first time, what these corpora really contain’ (Jockers 2013: 9–10).¹⁶

Across the scholarly spectrum, as well as wider society, new levels of expertise are called for, usually grouped under the notion of ‘data science’, a thoroughly interdisciplinary approach sometimes understood as the movement from ‘search’ to ‘correlation’. The digital humanities can bring this expertise in humanities processing together with large data sets, such as those found in Big Data, or data lakes. Critical approaches can supplement these new digital methods and explore how technologies are organized into complex assemblages which create the conditions under which new knowledge and practices are developed – material digital culture. Across the multiple levels of their operation, basic principles of computation are applied, such as: modularity, iteration, abstraction, optimization, etc., to understand how normative decisions are delegated and prescribed. Digital humanities can also explore the governmentality of these structures, and how they connect to wider questions in society, and so on (Rouvroy 2009). Information has time value, and soon can lose its potency. This drives the growth of not just Big Data, but real-time analysis – particularly where real-time or archival databases can be compared and processed in real-time. Currently, real-time is a huge challenge for computational systems and pushes at the limits of existing computational systems and data analytic tools.

In a technical sense, this is the emergence of what we might call ‘really big data’ applications represented by ‘batch map/reduce’ – such as Hadoop and related computational systems – and ‘real-time map/reduce’, whereby real-time analytics are also made possible on data – represented currently by technologies like Google’s Dremel (Melnik et al. 2010), Caffeine (Higgenbotham 2010), Impala (Brust 2012), Apache Drill (Vaughan-Nichols 2013) and Spanner (Iqbal 2013), etc. This is the use of real-time stream processing combined with complex analytics and the ability to manage large historical data sets, for example audio-visual archives, massive text sets, web archives, etc. The challenges for digital humanities are considerable at this scale, requiring peta-scale computational architectures so that the data can be held in memory, but also the construction of huge distributed memory systems enabling in-memory analytics, combined with developing complex algorithms and technical skills that create the research infrastructures to support and develop these resources for humanities

scholars. With the requirement to build, manage and critique these systems, one can readily understand why a physical lab can be an important part of the supporting infrastructure of computational work in the humanities.

The digital humanities is described here as an historical, material and critical approach, understanding that the key organizing principles of our experience are produced by ideas developed within the array of social forces that human beings have themselves created. We agree with Moretti when he argues that ‘to make the humanities relevant you need something much bigger than the digital humanities. What the humanities need are large theories and bold concepts’ (Moretti 2016). We think that the digital humanities could have something important to contribute towards thinking about and developing our understanding of the role of the human in an increasingly digital present – it could and should develop large theories and bold concepts. This includes understanding computational conditions of possibility for developing and supporting research infrastructures, and appreciation of the complexity of describing and encoding knowledge representation, creating and maintaining tools and archives as a critical reflexive subject and as an agent dynamically contributing and responding to these structures and practices. How might one structure a humanities education that contributes to the personal development and intellectual maturity of humanity, educate citizens, etc.? What are the key questions and challenges in the computational age and what might a humanities undergraduate or postgraduate programme look like when informed by these insights?

We now turn to some of these questions, firstly in terms of what we might call the conditions of possibility for developing the knowledge and understanding for critical and reflexive work in the digital humanities – something we are calling computational thinking.

Notes

1. A series of monographs in the British Library Research series give a helpful overview of humanities computing up until the mid-1990s (Katzen 1990; Kenna and Ross 1995; Mullings 1996).
2. For the purposes of this discussion, digital history and digital humanities are somewhat conflated although they do have separate histories and trajectories in their own right (for example, see Thomas 2004; Gregory 2014). For a discussion of music and digital humanities, see Fujinaga and Weiss (2004). For performing arts, see Saltz (2004). See The Old Bailey Online (<https://www.oldbaileyonline.org> and <https://www.londonlives.org>, Hitchcock and Shoemaker 2015), The Atlantic Slave Trade and Slave Life in the Americas: A Visual Record project (<http://slaveryimages.org>) and Valley of the Shadow project (<http://valley.lib.virginia.edu>) for good examples of digital history, together with the History Harvest (<http://historyharvest.unl.edu>).
3. As Scheinfeldt (2014) explains, ‘digital history [is] a natural outgrowth of long-standing public and cultural historical activities rather than a belated inheritance of the quantitative history experiments of the 1960s and 1970s. It is a story that begins with people like Allan

Nevins of the Columbia Oral History Office and Alan Lomax of the Library of Congress's Archive of American Folk-Song, especially with the man-on-the-street interviews Lomax coordinated in the aftermath of the Pearl Harbor attacks.'

4. 'Recognizing a pattern implies remaining open to gatherings, groupings, clusters, repetitions, and responding to the internal and external relations they set up' (L. Hunter, quoted in Ramsay 2004: 195).
5. Kirschenbaum and Werner observe that 'software studies, critical code studies, and platform studies are each varyingly inflected methodologies for cultivating both the critical sensibility and the technical acumen necessary to swim deep into the cultural reservoirs of contemporary digital production' (Kirschenbaum and Werner 2014: 435).
6. Barnett (2014: 74) usefully asks what origin story of digital humanities would include the #AntiJemimas project, for example. For the project, Kismet Nuñez (#AntiJemimas Founder and Creative Director) 'deploys 21st century forms of art, autobiography, and performance against the discursive terrain of race, sex and personality. With the help of new media, [Nuñez] breaks herself into pieces to become more than her parts in a revolutionary act of defiance, affirmation & self-care.'
7. In a more complete history of digital humanities than we have space to outline here, we would include thinkers like G. W. Leibniz, who developed a method of calculation called the calculus ratiocinator, an innovation his successors George Boole and Gottlob Frege extended, and, with the subsequent work of Georg Cantor, David Hilbert, Kurt Gödel, and Alan Turing, contributed to the computer that we know today (Davies 2000; see also McCarty 2013b). But we might also think of including Ada Lovelace, Charles Babbage, Christopher S. Strachey, Joseph Weizenbaum, Grace Murray Hopper, John von Neumann, Adele Goldberg and many others. We would also include systems like Strachey's Love Poetry Generator developed in 1953, and Weizenbaum's ELIZA software program in 1966, which demonstrated how text could be manipulated in and through computational systems.
8. The theological beginnings of digital humanities is also interesting in light of its conceptualization of world and words, as Busa himself commented, 'I am full of amazement at the developments [in digital humanities] since [1949]; they are enormously greater and better than I could then imagine. *Digitus Dei est hic!* The finger of God is here!' (Busa 2004).
9. However, one would note that art historical projects within digital humanities are also fundamentally concerned with visuality and the image.
10. Golumbia notes, however: 'despite its insistence on being only a set of methods, DH can be productively understood as a political intervention within literary studies, one of whose functions is to challenge the authority of "non-DH" literature scholars regarding our own discipline, in particular via a tendentious deployment of both the terms digital and humanities' (Golumbia 2014: 157).

11. Key journals within digital humanities include: *Digital Humanities Quarterly*, which is published by the Alliance of Digital Humanities Organizations as a peer review and open access journal; *Digital Scholarship in the Humanities* (DSH) – originally titled *Literary and Linguistic Computing* (LLC), the journal has changed its name to reflect better the discipline of digital humanities and the content of publications; *Vectors: Journal of Culture and Technology in a Dynamic Vernacular*, an experimental journal that brings together multimedia texts; *Digital Studies / Le champ numérique*, sponsored by the Canadian Society for Digital Humanities (Société pour l'Étude des Médias Interactifs) (see also Kirschenbaum 2012b: 4).
12. It is also interesting to compare these debates with postcolonial critiques of the archive, such as that articulated by Mbembe (2015).
13. Kirschenbaum writes: 'To assert that digital humanities is a "tactical" coinage is not simply to indulge in neopragmatic relativism. Rather, it is to insist on the reality of circumstances in which it is unabashedly deployed to get things done – "things" that might include getting a faculty line or funding a staff position, establishing a curriculum, revamping a lab, or launching a center' (Kirschenbaum 2012b: 415). He continues this strategic sense of the term 'digital humanities' by arguing that 'DH is a means and not an end' (2012b: 427).
14. Sometimes this notion of community can be rather defensively deployed, for example Kirschenbaum argues 'so it is with digital humanities: you are a digital humanist if you are listened to by those who are already listened to as digital humanists, and they themselves got to be digital humanists by being listened to by others. Jobs, grant funding, fellowships, publishing contracts, speaking invitations – these things do not make one a digital humanist' (Kirschenbaum 2014: 55).
15. Koh argues, 'in all of these debates, technical knowledge is consistently emphasized or dismissed as an important condition of entry into the digital humanities community. Technical knowledge is linked to the concept of *building* and is cast as a means by which community members perform their civic duties' (Koh 2014: 99).
16. Hydraulic mining, also referred to as hydraulicking, is a technique used in mining that uses high-pressure jets of water to dislodge rock material or move sediment. In mining gold, for example, the resulting water–sediment slurry is filtered through sluice boxes that remove the gold ore. We are not sure that this is an appropriate metaphor for handling sensitive cultural artifacts as a form of knowledge-fracking, which appears exploitative and treats cultural history as a form of standing reserve.