AI, ethics, and digital humanities

David Berry

Publication date
01-12-2022

Document Version
Published version

CHAPTER FORTY-TWO

AI, Ethics, and Digital Humanities

DAVID M. BERRY (UNIVERSITY OF SUSSEX)

With the rise of artificial intelligence (AI) and machine learning (ML) it is inevitable that they become both a focus of method within the digital humanities but also a question of practice and ethics. Across a range of disciplines, the impact of the automation of methods of analysis that AI promises raises specific challenges in each field in terms of implementation and methodological reflection. However, even as they raise questions of biases and errors, there is also a more fundamental obscurity problem that is raised by the particular explanatory deficit in relation to the difficulty of understanding what it is that an AI is doing when it classifies and analyzes the underlying research data. How then can we be sure that the algorithms act ethically and legally? We should note that there is a growing literature on digital technology, software, and related areas that are pertinent to the questions raised in this chapter. For example, studies of algorithms such as software studies/cultural analytics (Manovich 2001, 2008; Dix et al. 2003; Fuller 2003; Galloway 2006; Chun 2008, 2017; Hayles 2012; Bratton 2016; Montfort 2016; Berry 2021) and critical code studies (Marino 2006, 2020; Wardrip-Fruin 2009). Other relevant research includes work on new media studies and machine learning (Manovich 2010; Alpaydin 2016; Pasquale 2016, 2020; Domingos 2017), as well as more recent work on ethics and algorithmic bias (Chun 2011; Floridi 2013; Gray 2014; Floridi and Taddeo 2016; Han 2017; Noble 2018; Benjamin 2019; Amoore 2020). There is also an important related area analyzing data and surveillance (Browne 2015; D'Ignazio and Klein 2019; Zuboff 2019).

In this chapter, I walk through some of the issues that pertain to the field of digital humanities, particularly in light of AI's likely impact on ways of doing digital humanities, automation more generally, and ethics of artificial intelligence. In section one, I look to debates about how the automation of humanities approaches to their research objects change the nature of humanities research and the role of the humanities scholar in relation to digital humanities. In section two, I concentrate on how digital humanities currently uses artificial intelligence, mostly machine learning, but also some of the potentials for expansion of its methods more widely in data analysis and forms of distant reading. Finally, in section three I look at the potential problems raised by artificial intelligence and machine learning for the practices of digital humanists, how we might think about the ethics issues that are raised and bring forward key elements of the debates currently taking place.

Berry, D. M. (2023) AI, Ethics, and Digital Humanities, in O'Sullivan, James (Ed.), The Bloomsbury Handbook to the Digital Humanities, Bloomsbury Academic.

AUTOMATION OF THE HUMANITIES

As far back as 1981, Steve Jobs, then CEO of Apple, famously called computers "Bicycles for the Mind," implying that they augmented the cognitive capacities of the user, making them faster, sharper, and more knowledgeable. He argued that when humans "created the bicycle, [they] created a tool that amplified an inherent ability … The Apple personal computer is a 21st century bicycle if you will, because it's a tool that can amplify a certain part of our inherent intelligence … [it] can distribute intelligence to where it's needed" (Jobs 1981, 8–9). Jobs was influenced by earlier work by Doug Engelbart and Vannevar Bush who were both concerned with creating "man-computer symbiosis" and "augmenting human intellect" using technology (Ekbia and Nardi 2017, 26, 7).[1] This has also been the major way in which digital humanities, particularly in its earlier incarnations as humanities computing, conceptualized creating digital technology to augment the capacities of the humanities researcher. This generally took the form of thinking about methods of performing text analysis using a number of different forms of calculation, particularly of large corpora or datasets. Augmentation was thought to extend the human capacity to do things, and in contrast automation replaced the human with an algorithm, both of which raise different ethical considerations.[2]

Much of the earlier work that discussed the question of automation and augmentation tended to assume rather basic forms of automation that could assist the humanist in order to bring order to large amounts of data.[3] This was through algorithms that could help sort, list, or classify in simple ways, such as through the application of statistical methods. For example, McCarty (2003, 1225) described the "methodological commons" the digital humanities shared, which he listed as database design, numerical analysis, imaging, text analysis, and music retrieval and analysis. He offered a high-level division of digital humanities labor in terms of three branches working respectively on algorithmic (e.g., "mechanical" data analysis), metalinguistic (e.g., encoding texts in TEI and related markup languages), and representational (e.g., visual and presentational techniques) aspects of digital humanities work.

To simplify, today we tend to group digital humanities work into the divisions of digital tools and digital archives. Digital tools are software methods for working with textual data, such as the infamous word cloud which provides a visualization based on the frequency of words in a text so that words that have greater frequency are visually represented in a larger size.[4] Digital tools augment the humanist to be able to select and manipulate datasets so that they can be grouped and re-sorted to look for patterns. In contrast, the creation of digital archives tends to be image-based or transcriptions of physical textual archives, often with diplomatic versions for each of reading and searching. In the case of digital archives, automation is embedded in databases that collect corpora together and allow various searches and filters to be applied. Some archives have attempted to use OCR (optical character recognition) to change the text-based images into searchable text, but generally the quality of automated OCR output has remained relatively unsatisfactory for unsupervised use and actually generated a requirement for a lot of human labor to correct the mistakes that it inevitably created. This drove a need for what has come to be called "crowd-sourcing" through the use of Internet technologies to mediate a relationship between a researcher and a purported public (Terras 2016, 420–39).[5] This has also raised the interesting idea of slow digitization that foregrounds the difficulty of producing digital archives (Prescott and Hughes 2018). Both of these generate questions over how humanists should undertake their practices and

the ethical considerations that might guide their action. In other words, what digital humanists ought to do and how they should think about their practices reflexively as a discipline.

Today, many of these ways of working with text and other forms of humanities data have become standardized into software packages and websites. This allows them to be used by scholars who may not necessarily understand the underlying complexity of the algorithms, and who may be unaware of the ethical decisions that have been embedded in the code. This is similar to the way that computer scientists working on algorithmic systems bracket off complexity by studiously ignoring how the functions they depend on are implemented—that is, humanists "black box" the computational aspects of doing digital humanities. Programmers who create these systems are taught to construct and respect "walls of abstraction"—functional modules that can be invoked in standard and consistent ways, hiding complexities within. Within the digital humanities there is a growing move toward these black-boxed systems as their convenience has increased, such as with the Voyant Tools system which provides a number of different text analysis functions within one system.[6] Even though Voyant Tools is an open-source project and the code is available through GitHub, few digital humanists are likely to inquire deeply into the code itself or check the source code. This leaves open the possibility that digital humanists may not fully be aware what the code that operates on their data may be doing. That is to raise the issue of software mediation which becomes a potentially new and interesting research question for the digital humanist (Berry 2011b, 2012). It also gestures to some of the underlying ethical issues when using automated systems that may have norms embedded into the code which humanists may be unaware of.

This process of automation can also increase the tendency towards formalization made possible by "algorithmization" (for an example see Allison et al. 2011, 6). Digital humanists then run the risk of creating systems where few people know explicitly or understand how they work, or how they make decisions. This creates the danger that we rely on the automation and output of these systems and thereby are asked to trust the computer. This also raises the question of the creation and maintenance of digital humanities' intersubjective, historical, and philosophical concepts as manifested in and supported by these technologies (Liu 2012; Berry 2014). So, for example, assumptions can be made about gender or race, or about which books are more important than others, such as a literature canon (Benjamin 2019; Bordalejo and Risam 2019). Indeed, computational systems tend towards the generation of quantifiable hierarchies of entities, particularly when classifying, sorting, and filtering, and even the notion of such a hierarchy is, of course, an imposed framework of understanding (see Johnson 2018).[7] Similarly, because of the underlying programming logics of a particular software tool, certain design assumptions, such as objects, networks, collections, or even the presumed suitability for particular properties of entities to be valued over others, may have effects on the way in which those digital objects are subsequently manipulated by the software (Berry 2011a; Emerson 2014; Marino 2020; Chun 2021).

DIGITAL HUMANITIES AND ARTIFICIAL INTELLIGENCE

We might, therefore, think of the automation of the humanities as moving through a number of phases. The first involved bespoke software, often programmed by a humanist-coder, to solve a particular problem. The second was the generalization of these early software tools into packages

which provide a "one-stop shop" and which involve little or no programming requirement, such as Voyant Tools. In the latest phase, we see the beginning of much more sophisticated machine-learning tools, again by humanist-coders who are required to customize the machine-learning software they use for a specific digital humanities project, for example the MALLET (MAchine Learning for LanguagE Toolkit) software.[8] We can probably assume therefore that we are on the cusp of a new set of packages that will further democratize access to machine learning for non-programming humanists over the next few years, such as Orange, which offers a visual programming experience for using machine learning.[9] Each of these distinct phases raises different kinds of ethical conundrums for the humanist, from questions of use, sharing, openness, respect for others, privacy, and inclusion to the ethics of crowdsourcing, managing, and working with controversial or situated knowledges, and even with the issue of absences and decolonization of digital humanities more broadly.

Digital humanists have already begun to experiment with algorithms that draw on artificial intelligence and machine learning such as the widespread use of topic models. These are algorithms that use machine-learning techniques to sort and classify textual data into "buckets" that can represent themes or concepts drawn from the underlying data (Blei 2012). Machine learning relies on a process whereby the programmer does not necessarily instruct the system by writing an algorithm, rather the system is shown the input and the output and taught to "extract automatically the algorithm for this task" (Alpaydin 2016, 17). So, for example, a humanist might show an image to a machine-learning system to teach it to generate a text description in order to standardize the descriptive text associated in a database of images. This adds another layer of complexity to the use of digital tools because when one utilizes a machine-learning system, there remains the strong possibility that the source code, which is at least available with systems like Voyant Tools, is actually not legible in quite the same way. Indeed, due to the way in which machine-learning systems work, they tend to encode their classifying processes within a table of numbers which represent relationships between entities. These tables of numbers can be very large and daunting for a human researcher to understand. One can see how these automating processes might cover over ethical issues by transferring them into the hashtables of the machine-learning system. Machine learning also has the power to automate larger parts of the humanities, replacing research work that was previously thought could only be the preserve of the human researcher such as identifying key concepts. So, while the humanist might have access to the source code, this becomes merely a preliminary layer upstream from the later classification learning process which is encoded through either supervised or unsupervised learning processes. This downstream classificatory structure then becomes much more difficult for the researcher to cognitively map or check. In fact, as the number of parameters used in the classificatory processing increase, so does the complexity of the tables or networks of numbers. For example, in 2020, the OpenAI Generative Pre-trained Transformer 3 (GPT-3) is described as having a capacity of over 175 billion machine-learning parameters and the texts it classifies produce tables that are extremely difficult to understand or, indeed, explain at the level of the models generated by the software. In relation to this, the idea of explainability has emerged as the idea that artificial intelligence systems should be able to generate a sufficient explanation of how an automated decision was made, representing or explaining, in some sense, its technical processing. With concerns over biases in algorithms there is an idea that self-explanation by a machine-learning system would engender trust in these

systems.[10] Explanations are assumed to tell us how things work, thereby giving us the power to change our environment in order to meet our own ends and thereby raise interesting ethical questions (see Berry 2021).

As the capability of these systems to automatically classify even very large datasets expands, we see a growing use of computational systems to abstract, simplify, and visualize complex "Big Data." However, these often produce simplistic probabilistic or statistical models used to organize complex humanities data. Additionally, many of these systems allow researchers to analyze data without a hypothesis so they can "play" with the data until a preferred result is found. Humanists can throw their data into large-scale computing clusters and let machine-learning algorithms find patterns that might have evaded human capacities—they may also encourage a carelessness with the data and what it may mean to specific communities (Johnson 2018, 61). This ability to analyze large datasets can be combined with the desire to deepen or widen the corpus from which data is being extracted, meaning that often the entire set of texts on which the dataset rests is never read in total, or even partially, by a researcher. This is the "computational turn" towards what has been called "distant reading" and assumes the automation of the reading process to ascertain themes and concepts from within a set of texts (Berry 2011a; see Underwood 2019). In using these tools, it has been argued that a researcher is required to trust the output of the algorithm given to analyzing the data, and they may have little formal understanding of the meaning of the results calculated or their significance (see Da 2019; Eccles 2020 has an excellent summary of this debate).

The use of these digital tools is increasing across the humanities, but at different speeds depending on the discipline. Nonetheless, the rise in convenience and the ease of access to both digitized and born-digital content in online databases strongly creates affordances towards using these techniques (see Posner 2016, 2018). The application of searches, such as Boolean search, to the downloading of datasets into pre-packaged CSV files or Zip files and using artificial intelligence, has the potential to skew research because technology tends to foreground a selection which might hide absences in a dataset or assert a presumed neutrality (Johnson 2018, 68). It is certainly one of the reasons for the calls for a critical digital humanities that foregrounds reflexivity and social justice in research practice.[11] This approach argues that digital humanists should become more critical of working in a digital environment, and the ways in which automation and the digital can subtly affect research practices (Berry and Fagerjord 2017; Johnson 2018, 68–9; Dobson 2019).[12]

ETHICS OF ARTIFICIAL INTELLIGENCE AND DIGITAL HUMANITIES

In this final section I want to take these emergent issues around digital humanities research in terms of the automation and augmentation of humanities work to look at the ethical issues raised.[13] The literature concerned with ethics related to the digital is a huge and growing area and, of course, I can only provide a broad outline of some of the key issues for consideration. So, in this chapter I will be focusing on the specific issues raised by thinking about ethics and ethical practices in relation to the digital humanities (see Ess 2009; Berry 2012, 12; Floridi 2013; Bostrom and Yudkowsky 2014 for some of the wider issues; Mittelstadt et al. 2016). The question is: What ought to be the principles or normative frameworks for guiding decisions in digital humanities projects and research units?[14]

Within the digital humanities there are differing ways of thinking about the ethics which might attach to digital projects or indeed to practices of digital humanists and have only recently begun to be foregrounded within the field (see for example Kennedy 2019; Padilla 2019). There have been some important debates about the ethos of the digital humanities, such as detailed by Cecire (2011) who has criticized the supposed "'virtues' of digital humanities – which include collaboration, humility, and openness" and its tendency to self-narrate through its putative moral virtues rather than specific methodological or epistemological commitments (see also Spiro 2012). In this section I want to think about how ethics is seen in practice in relation to digital humanities through the lens of artificial intelligence. There has been a considerable amount of work on ethics in the related areas of digital media and Internet studies, together with areas like sociology, science and technology studies and in computer science itself, but much less in digital humanities (see Ess 2004: CPSR 2008, 134; Floridi 2010; AoIR 2021).[15]

When thinking about ethics, broadly speaking it is common to talk of three approaches to ethics, namely deontological, consequentialist, or virtue ethics (see van den Hoven 2010). Deontological approaches tend to bring forward the idea that choices cannot be made based on their result, rather the decision has to be in conformity with a moral norm or duties. This means that deontological approaches often are constructed in terms of maxims or principles that should be followed, in "compliance with the highest ethical principle: the Categorical Imperative," that is not to use people as means to ends (van den Hoven 2010, 69). This creates the sense of a shared set of rules, perhaps professional guidelines, that also allow others to complain or hold to account those who break these principles, for example the AoIR ethics guidelines for research (AoIR 2021). In contrast, consequentialist approaches hold that ethics should be assessed in relation to the consequences that actions bring about, one "ought to choose that course of action which maximizes utility compared to the alternatives" (van den Hoven 2010, 69). The more likely the actions that maximize that a given end is realized, that is that the moral definition of a good is achieved, the more right the decision (Anscombe 1958, 9). Virtue ethics, on the other hand, emphasizes the "virtues" or "moral character" of the actor. So, virtues and vices tend to be the key categories used to assess ethical action, and virtue itself is a matter of degree (Annas 2011).[16] So the "virtuous person also possesses a general intellectual capacity, practical wisdom" and this enables them to "identify the morally relevant features in every situation and determine the right course of action" (MacIntyre 2007; van den Hoven 2010, 68).[17]

Much of the existing literature on ethics and artificial intelligence tends to prioritize either the deontological or consequentialist approaches to thinking about the questions that the digital raises.[18] For example, in the case of data bias one of the key ways in which ethics tend to be thought about in relation to the digital is to locate it within the context of computational inequities or computational power. Injustice, for example, might be strongly linked to problems with the data, which some have argued can be addressed by more data, ethical data, or democratizing data sources (but see Presner 2015; Eubanks 2019, 299; D'Ignazio and Klein 2019, 60–1). This might raise deontological issues of how we prevent artificial intelligent systems using people as sources of data, or manipulating them, or alternatively to focus on the perceived consequences of particular machine-learning systems and the possible harms they might cause.

Within the digital humanities as a field, thinking about the idea of a set of ethical principles that might serve as a professional guide to action has been discussed but as yet no specific digital humanities guidance has been drawn up (but see Berne 2013; Rehbein 2018). Considering the breadth and depth of digital humanities work that has been undertaken over the past twenty years,

it is notable that ethics has been under-theorized within the digital humanities itself to a remarkable extent (see Raley 2014 as an alternative way of formulating this).[19] When one wishes to undertake a digital humanities research project, a set of guidelines is missing, a serious lack when one might want to think through ethical concerns.[20] Instead, most ways of thinking about ethics in digital humanities tend towards a form of virtue ethics in terms of the kinds of character that digital humanists are thought to have, such as being "'nice': collegial, egalitarian, non-hierarchical, and able to work well with others" (Koh 2014, 95). While virtue ethics offers some means for thinking through the questions raised by challenging problems in undertaking digital humanities work, virtue ethics tends to articulate these questions in terms of aretaic concepts (that is, in terms of vices/virtues) or axiological concepts (that is, in terms of better/worse or good/bad) rather than deontological notions (such as the right/wrong action, duty, or obligation).[21] This emphasis on the "good will" of the individual researcher places responsibility on the good character of the individual to do the virtuous thing but is less clear on the professional expression of that requirement.[22] This lack of formalized standards of good practice is notable considering the close relationships digital humanities have with digital archives, for example (see eHumanities 2010), which tend to raise questions about decolonization, gender, race, class, and sexuality, not just in terms of representation but also in terms of the assumed categories inherited from physical archive practices or in using data scraped from the Internet (see Liu 2012; Johnson 2018; Berry 2019).[23] Considering that computer ethics itself has been thinking about questions of how computer technology should be used since at least the 1960s, and as Moor (1985) argues, a central task of such an ethics should be to formulate policies to guide our actions, the lack of these guidelines is puzzling.[24] Drawing from Rehbein (2018) we might outline three areas from which such a set of ethical guidelines for the digital humanities might be drawn,

1. **Moral issues in specific fields of research and in close relation to the objects of study** (e.g., personal data, attribution, publishing historical or web data that might have wider public consequences for individuals, neutrality, quantification).
2. **Moral aspects of digital humanities as a profession**, both in terms of members of the digital humanities community and externally for society or individuals outside this community. Examples include "do not fabricate, do not falsify, do not plagiarize, honour the work of others, give credit to all who supported you, name your co-authors, do not publish the same thing twice, and various other guidelines that many scholarly communities have given themselves" (Rehbein 2018, 23).
3. **The responsibility of an individual scholar as well as of the scholarly community at large within digital humanities to the larger society**, for example creating risk for future generations through the use of technologies, their environmental impact, or commitment to equality, diversity, and inclusion.

Within the context of artificial intelligence, a field which has itself only recently been subjected to some ethical critique (but see Agre 1997), its growing use by digital humanities raises the importance of developing a set of ethics, particularly for thinking about the use of these technologies within the field. From the ethics of creating digital tools and digital archives, the use of facial recognition software, automated processes of cataloging or text generation, to the use of crowdsourcing using the "artificial artificial intelligence" or "ghost work" of human task work systems, there is clearly an urgent need for digital humanities to be thinking seriously about ethics, both professionally and at the level of the individual scholar (see Bianco 2012, 97; Reid 2012, 363; Spiro 2012, 18; Smithies

2017, 208–9; Gray and Suri 2019, 7).[25] The time is indeed ripe for the professional bodies in digital humanities to begin to develop such a set of ethics to build a strong foundation for future research, and to encourage a wider discussion of ethics within the digital humanities community.

ACKNOWLEDGMENTS

Thanks to Natalia Cecire, Charles Ess, David Golumbia, Lauren F. Klein, Alan Liu, Willard McCarty, Miriam Posner, and Andrew Prescott for their helpful suggestions and comments on this chapter.

NOTES

1. Engelbart himself was deeply influenced by Vannevar Bush's article *As We May Think*, published in *The Atlantic* and his notion of the "memex," an enlarged intimate mechanical supplement to a person's memory (Bush 1945), and J. C. R. Licklider's (1960) work on *Man-Computer Symbiosis*. As Licklider wrote using a very revealing metaphor:

   > the fig tree is pollinated only by the insect *Blastaphaga grossorum*. The larva of the insect lives in the ovary of the fig tree, and there it gets its food. The tree and the insect are thus heavily interdependent: the tree cannot reproduce without the insect; the insect cannot eat without the tree; together, they constitute not only a viable but a productive and thriving partnership. This cooperative "living together in intimate association, or even close union, of two dissimilar organisms" is called symbiosis. "Man-computer symbiosis" is a subclass of "man-machine systems" … The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today.

   Licklider also makes the important distinction between automation and a form of augmentation, although he doesn't use the term, when he argues

   > "Mechanical extension" has given way to replacement of men, to automation, and the men who remain are there more to help than to be helped … [Whereas] to think in interaction with a computer in the same way that you think with a colleague whose competence supplements your own will require much tighter coupling between man and machine than is suggested by the example than is possible today. (Licklider 1960)

2. As Ceruzzi observes,

   > the term "automation" was coined at the Ford Motor Company in 1947 and popularized by John Diebold in a 1952 book of that title. Diebold defined the word as the application of "feedback" mechanisms to business and industrial practice, with the computer as the principal tool. He spoke of the 1950s as a time when "the push-button age is already obsolete; the buttons now push themselves." (Ceruzzi 2000, 32)

3. It is notable that the automation through artificial intelligence of formerly critical and cognitive practices by the humanities scholar raises particular ethical issues in terms of understanding and responsibility.
4. We should also note that different digital tools may have differing kinds of opacity associated with them. Their development and usage may require them to be critically analyzed depending on how this opacity is manifested.
5. Crowdsourcing is based on principles of outsourcing, raising related issues of labor inequality and exploitation.

6. See Voyant Tools, https://voyant-tools.org.
7. See Chapter 12, Critical Digital Humanities in this volume.
8. http://mallet.cs.umass.edu/index.php.
9. https://orangedatamining.com.
10. Within the field of AI there is now a growing awareness of this problem of opaque systems and a sub-discipline of "explainable AI" (XAI) has emerged and begun to address these very complex issues—although mainly through a technical approach.
11. I argue that it is important that "ethics" and "social justice" should not be seen as antithetical, and both can contribute towards an idea of a future digital humanities that is inclusive, diverse, ethical, and equitable (see Chapter 12, Critical Digital Humanities in this volume, but also see Metcalf et al. 2019, 471).
12. For example, the work of the #transformDH collective, which formed in 2011, and "offered a recent challenge to [digital humanities]. Members of the collective argued that the consolidation of digital humanities as a field presented university-affiliated laborers with an opportunity to embark on an explicitly antiracist, anti-ableist, radical and inclusive academic project" (Johnson 2018, 68).
13. Zuboff (1988) has argued that we might need to distinguish between *automation*, which relieves humans of physical effort, and *informating*, which integrates people into a higher level of intellectual involvement and productivity. Today, we are more likely to use the term *augmentation* than informating, but the idea of it being a better way of automating that "realizes human potential of the workforce as well as the technical potentialities of the computer" remains key.
14. Indeed, I argue that ethics is inseparable from reflexivity (Berry 2004, see Keane 2015, 23).
15. There is also an interesting attempt to reconceptualize the arts and humanities as "deep humanities" where one of the key contributions would be an attempt to engage with issues of ethics related to technology, sustainability and augmentation. This includes Ethics, understood as "ethical ways of conceiving and connecting with the Other all its planetary diversity; integrating ethics into STEM/STEM education, business, politics, planning, and policy;" Bias, "social, cultural, and technological, including algorithmic bias; common sense and critical thinking in ML/AI" and Design, "ethical, affective, inclusive, and sustainable but not human centered; instead of placing human at the center/top, Deep Humanities proposes to place the human in proportion to and in right relations with the nonhuman (animal, plant, machine)" amongst other topics (Krishnaswamy 2019). Indeed, a "deep digital humanities" might consider proposals from deep humanities and other critical attempts to think about technology and the humanities to develop its own disciplinary matters of concern.
16. Although related, due to space considerations the area of "information ethics" is not able to be included, but see Doyle (2010) for a critique of information ethics. Additionally, the area of "contractarianism" and ethics which stems from the work of the social contract philosophers, either through the work of Hobbes or Kant, and argues that ethics draws its force from mutual agreement or an idea of contract, is beyond the scope of the chapter (Rawls 1971; Gauthier 1986).
17. Due to limitations on space, it is impossible to do justice to MacIntyre's work on virtue ethics, but see Anscombe (1958) and MacIntyre (2007). There is a good summary of the position in the Stanford *Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/ethics-virtue/.
18. See also the ethics of care, an important way of reconceptualizing ethics in relation to feminist thought. Held (2006) states that an ethics of care "builds relations of care and concern and mutual responsiveness to need on both the personal and wider social levels. Within social relations in which we care enough about one another to form a social entity, we may agree on various ways to deal with one another." See also D'Ignazio and Klein (2019, 257, fn39) for a discussion of this in relation to what they call "Data Feminism."
19. There is also an interesting question about the relationship between open access and the ethics which unfortunately I don't have space to discuss in this article. For example, are the technical guidelines and requirements of open access rooted in ethical practices within the digital humanities? Or is there an alignment of open access and computational capital which needs to be carefully considered in relation to the neoliberal university (see Golumbia 2016)?
20. For example, see Lennon (2014) for a thoughtful discussion about digital humanities and the ethics of funding.

21. Aretaic concepts are concerned with *arête* (excellence or virtue), *phronesis* (practical or moral wisdom), and *eudaimonia* (usually translated as happiness or flourishing) (see Annas 2011, 3). Axiological concepts are concerned with value, particularly intrinsic and extrinsic value, based on comparative analysis of, for example, which activity is better or worse than another (see Scheler 1973). In contrast, deontological concepts are concerned with moral rules or professional codes of practice, particularly in terms of duty and obligations and particularly associated with the work of Immanuel Kant and his conception of the "good will" which acts out of respect for the moral law (see Kant 1997).
22. Virtue ethics tends to derive from Christian notions of a "good will" and the reliance on virtue ethics within digital humanities is interesting in relation to its early historical links to biblical and theological scholarship (particularly philology and textual criticism). Early digital humanists, such as Father Roberto Busa, an Italian Jesuit who studied Thomas Aquinas's texts, were extremely influential on the early use of computers in humanities work and this connection perhaps raises suggestive questions about an under-theorized Christian ethical foundation to the discipline of digital humanities (see Fiormonte 2017 for an overview of this history, but also Rockwell and Sinclair 2020). It is notable that Aquinas has also been a very influential thinker on the development of some versions of virtue ethics (see MacIntyre 2007, x).
23. See also Chapter 12, *Critical Digital Humanities*, in this volume.
24. Although there are examples of good practice, such as https://www.colorado.edu/lab/dsrl/statement-ethics and http://mcpress.media-commons.org/offthetracks/part-one-models-for-collaboration-career-paths-acquiring-institutional-support-and-transformation-in-the-field/a-collaboration/collaborators'-bill-of-rights/.
25. Crowdsourcing has become increasingly mediated by platforms in the past decade, for example, Amazon's Mechanical Turk, https://www.mturk.com, Microworkers, https://www.microworkers.com, Clickworker, https://www.clickworker.com/clickworker/, Appen, https://appen.com/jobs/contributor/, and Upwork, https://www.upwork.com/i/how-it-works/client/.

REFERENCES

Agre, Philip E. 1997. "Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI." In *Social Science, Technical Systems, and Cooperative Work: Bridging the Great Divide*, edited by Geoffrey Bowker, Les Gasser, Susan Leigh Star, and Bill Turner, 131–58. Hillsdale, NJ: Erlbaum.

Allison, Sarah, Ryan Heuser, Matthew Jockers, Franco Moretti, and Michael Witmore. 2011. *Quantitative Formalism: An Experiment*. Literary Lab, Pamphlet #1, January 15.

Alpaydin, Ethem. 2016. *Machine Learning: The New AI*. Cambridge, MA: MIT Press.

Amoore, Louise. 2020. *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Durham, NC: Duke University Press

Annas, Julia. 2011. *Intelligent Virtue*. New York, NY: Oxford University Press.

Anscombe, G. E. M. 1958. "Modern Moral Philosophy." *Philosophy* 33 (124) (January): 1–19.

AoIR. 2021. Ethics. https://aoir.org/ethics/.

Benjamin, Ruha. 2019. *Race after Technology*. Cambridge: Polity.

Berne. 2013. Berne DHSS Declaration of Research Ethics in Digital Humanities. Written collaboratively at the Digital Humanities Summer School 2013. https://docs.google.com/document/d/1A4MJ05qS0WhNlLdlozFV3q3Sjc2kum5GQ4lhFoNKcYU/.

Berry, David M. 2004. "Internet Research: Privacy, Ethics and Alienation: An Open Source Approach." *Internet Research* 14 (4): 323–32.

Berry, David M. 2011a. "The Computational Turn: Thinking about the Digital Humanities." *Culture Machine* 12.

Berry, David M. 2011b. *The Philosophy of Software: Code and Mediation in the Digital Age*. London: Palgrave Macmillan.

Berry, David M. 2012. *Understanding Digital Humanities*. London: Palgrave Macmillan.

Berry, David M. and Anders Fagerjord. 2017. *Digital Humanities: Knowledge and Critique in a Digital Age*. Cambridge: Polity.

Berry, David M. 2014. *Critical Theory and the Digital*. New York, NY: Bloomsbury.
Berry, David M. 2019. "Critical Digital Humanities." *Conditiohumana.io*. https://conditiohumana.io/critical-digital-humanities/.
Berry, David M. 2021. "Explanatory Publics: Explainability and Democratic Thought." In *Fabricating Publics: The Dissemination of Culture in the Post-truth Era*, edited by Bill Balaskas and Carolina Rito. London: Open Humanities Press.
Bianco, Jamie "Skye." 2012. "This Digital Humanities Which Is Not One." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 96–112. Minneapolis, MN: University of Minnesota Press.
Blei, David M. 2012. "Topic Modelling and Digital Humanities." *Journal of Digital Humanities* 2 (1) (Winter).
Bordalejo, Barbara, and Roopika Risam, eds. 2019. *Intersectionality in Digital Humanities*. York: Arc University Press.
Bostrom, Nick and Eliezer Yudkowsky. 2014. "The Ethics of Artificial Intelligence." In *The Cambridge Handbook of Artificial Intelligence*, edited by Keith Frankish and William Ramsey, 316–34. Cambridge: Cambridge University Press.
Bratton, Benjamin. 2016. *The Stack: On Software and Sovereignty*. Cambridge, MA: MIT Press.
Browne, Simone. 2015. *Dark Matters: On the Surveillance of Blackness*. London: Duke University Press.
Bush, Vannevar. 1945. "As We May Think." *The Atlantic*, July.
Cecire, Natalia. 2011. "Introduction: Theory and the Virtues of Digital Humanities." *Journal of Digital Humanities* 1 (1) (Winter).
Ceruzzi, Paul E. 2000. *A History of Modern Computing*. Cambridge, MA: MIT Press.
Chun, Wendy H. K. 2008. "On 'Sourcery,' or Code as Fetish." *Configurations* 16: 299–324.
Chun, Wendy H. K. 2011. *Programmed Visions: Software and Memory*. Cambridge, MA: MIT Press.
Chun, Wendy H. K. 2017. *Updating to Remain the Same: Habitual New Media*. Cambridge, MA: MIT Press.
Chun, Wendy H. K. 2021. *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*. Cambridge, MA: MIT Press.
CPSR. 2008. Computer Professionals for Social Responsibility. http://cpsr.org.
Da, Nan Z. 2019. "The Computational Case against Computational Literary Studies." *Critical Inquiry* 45 (3) (Spring).
D'Ignazio, Catherine and Lauren F. Klein. 2019. *Data Feminism*. Cambridge, MA: MIT Press.
Dix, Alan, Janet E. Finlay, Gregory D. Gregory Abowd, and Russell Beale. 2003. *Human Computer Interaction*. London: Prentice Hall.
Dobson, James E. 2019. *Critical Digital Humanities: The Search for a Methodology*. Urbana, IL: University of Illinois Press.
Domingos, Pedro. 2017. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. London: Penguin.
Doyle, Tony. 2010. "A Critique of Information Ethics." *Knowledge, Technology and Policy* 23: 163–75.
Eccles, Kathryn. 2020. "Digital Humanities." *The Year's Work in Critical and Cultural Theory* 28 (1): 86–101.
eHumanities. 2010. "Ethical Issues in Digital Data Archiving and Sharing." *eHumanities NL*. https://www.ehumanities.nl/ethical-issues-in-digital-data-archiving-and-sharing/.
Ekbia, Hamid R. and Bonnie Nardi. 2017. *Heteromation, and Other Stories of Computing and Capitalism*. Cambridge, MA: MIT Press.
Emerson, Lori 2014. *Reading Writing Interfaces: From the Digital to the Bookbound*. Minneapolis, MN: University of Minnesota Press.
Ess, Charles. 2004. "Revolution? What Revolution? Successes and Limits of Computing Technologies in Philosophy and Religion." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 132–42. London: Wiley-Blackwell.
Ess, Charles. 2009. *Digital Media Ethics*. Cambridge: Polity.
Eubanks, Virginia. 2019. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY: St Martin's Press.

Fiormonte, Domenico 2017. *The Digital Humanities from Father Busa to Edward Snowden*. The Centre for the Internet and Society. https://cis-india.org/raw/the-digital-humanities-from-father-busa-to-edward-snowden.
Floridi, Luciano. 2010. *The Cambridge Handbook of Information and Computer Ethics*. Cambridge: Cambridge University Press.
Floridi, Luciano. 2013. *The Ethics of Information*. Oxford: Oxford University Press
Floridi, Luciano and M. Taddeo. 2016. "What is Data Ethics?" *Philosophical Transactions of the Royal Society A* 374: 20160360.
Fuller, Matthew. 2003. *Behind the Blip: Essays on the Culture of Software*. London: Autonomedia.
Galloway, Alexander. 2006. *Protocol: How Control Exists after Decentralization*. London: MIT Press.
Gauthier, David. 1986. *Morals by Agreement*. Oxford: Oxford University Press.
Golumbia, David. 2016. "Marxism and Open Access in the Humanities: Turning Academic Labor against Itself." *Workplace* 28: 74–114.
Gray, Kishonna L. 2014. *Race, Gender, and Deviance in Xbox Live: Theoretical Perspectives from the Virtual Margins*. New York, NY: Routledge.
Gray, Mary and Siddharth Suri. 2019. *Ghost Work: How Amazon, Google, and Uber Are Creating a New Global Underclass*. Boston, MA: Houghton Mifflin Harcourt Publishing Company.
Han, Byung-Chul. 2017. *Psychopolitics: Neoliberalism and New Technologies of Power*. New York, NY: Verso.
Hayles, N. Katherine. 2012. *How We Think: Digital Media and Contemporary Technogenesis*. Chicago, IL: University of Chicago Press.
Held, Virginia. 2006. *The Ethics of Care: Personal, Political, and Global*. Oxford: Oxford University Press.
van den Hoven, Jeroen. 2010. "The Use of Normative Theories in Computer Ethics." In *The Cambridge Handbook of Information and Computer Ethics*, edited by Luciano Floridi, 59–76. Cambridge: Cambridge University Press.
Jobs, Steve. 1981. "When We Invented the Personal Computer…." *Computers and People Magazine*, July–August.
Johnson, Jessica Marie. 2018. "Markup Bodies: Black [Life] Studies and Slavery [Death] Studies at the Digital Crossroads." *Social Text* 36 (4): 57–79.
Kant, Immanuel. 1997. *Groundwork of the Metaphysic of Morals*. Cambridge: Cambridge University Press.
Keane, Webb. 2015. *Ethical Life: Its Natural and Social Histories*. Princeton, NJ: Princeton University Press.
Kennedy, Mary Lee. 2019. "What Do Artificial Intelligence (AI) and Ethics of AI Mean in the Context of Research Libraries?" *Research Library Issues* (299): 3–13. https://doi.org/10.29242/rli.299.1.
Koh, Adeline. 2014. "Niceness, Building, and Opening the Genealogy of the Digital Humanities: Beyond the Social Contract of Humanities Computing." *differences* 25 (1): 93–106.
Krishnaswamy, Revathi. 2019. "What is Deep Humanities & Arts?" San José State University, https://web.archive.org/web/20210425063845/https://www.sjsu.edu/humanitiesandarts/Faculty_Development_Resources/What_Is_Deep_Humanities.html.
Lennon, Brian. 2014. "The Digital Humanities and National Security." *differences* 25: 132–55.
Licklider, J. C. R. 1960. "Man Computer Symbiosis." *IRE Transactions on Human Factors in Electronics*, March.
Liu, Alan. 2012. "The State of the Digital Humanities: A Report and a Critique." *Arts & Humanities in Higher Education* 11 (1–2): 8–41.
MacIntyre, Alasdair. 2007. *After Virtue: A Study In Moral Theory*. Notre Dame, IN: University of Notre Dame Press.
Mackenzie, Adrian. 2017. *Machine Learners: Archaeology of a Data Practice*, London: MIT Press.
Manovich, Lev. 2001. *The Language of New Media*. London: MIT Press.
Manovich, Lev. 2008. *Software Takes Command*. https://web.archive.org/web/20110411164341/http://lab.softwarestudies.com/2008/11/softbook.html.
Manovich, Lev. 2010. *Info-Aesthetics*. London: Bloomsbury.
Marino, Mark C. 2006. *Critical Code Studies*. https://web.archive.org/web/20130804160545/http://www.electronicbookreview.com/thread/electropoetics/codology.
Marino, Mark C. 2020. *Critical Code Studies*. Cambridge, MA: MIT Press.

McCarty, Willard. 2003. "Humanities Computing." In *Encyclopedia of Library and Information Science*, 2nd edn. New York, NY: Marcel Dekker.
Metcalf, Jacob, Emanuel Moss, and danah boyd. 2019. "Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics." *Social Research: An International Quarterly* 82 (2) (Summer): 449–76.
Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. "The Ethics of Algorithms: Mapping the Debate." *Big Data & Society* (December 2016). https://doi.org/10.1177/2053951716679679.
Montfort, Nick. 2016. *Exploratory Programming for the Arts and Humanities*. Cambridge, MA: MIT Press.
Moor, James H. 1985. 'What is Computer Ethics?' *Metaphilosophy* 16 (4): 266–75.
Noble, Safiya Umoja. 2018. *Algorithms of Oppression*. New York, NY: New York University Press.
Padilla, Thomas. 2019. *Responsible Operations: Data Science, Machine Learning, and AI in Libraries*. Dublin, OH: OCLC Research. https://doi.org/10.25333/xk7z-9g97.
Pasquale, Frank. 2016. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.
Pasquale, Frank. 2020. *New Laws of Robotics: Defending Human Expertise in the Age of AI*. Cambridge, MA: Harvard University Press.
Posner, Miriam. 2016. "What's Next: The Radical, Unrealized Potential of Digital Humanities." In *Debates in the Digital Humanities*, edited by Matthew K. Gold and Lauren Klein, 32–41. Minneapolis, MN: University of Minnesota.
Posner, Miriam. 2018. "Digital Humanities." In *The Craft of Criticism: Critical Media Studies in Practice*, edited by Mary Celeste Kearney and Michael Kackman, 331–46. New York, NY: Taylor & Francis.
Prescott, Andrew and Lorna Hughes. 2018. "Why Do We Digitize? The Case for Slow Digitization." *Archive Journal*. http://www.archivejournal.net/essays/why-do-we-digitize-the-case-for-slow-digitization/.
Presner, Todd. 2015. "The Ethics of the Algorithm: Close and Distant Listening to the Shoah Foundation Visual History Archive." In *History Unlimited: Probing the Ethics of Holocaust Culture*, edited by Claudio Fogu, Wulf Kansteiner, and Todd Presner, 175–202. Cambridge, MA: Harvard University Press.
Raley, Rita. 2014. "Digital Humanities for the Next Five Minutes." *differences* 25 (1): 26–45.
Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
Rehbein, Malte. 2018. "It's our department: On Ethical Issues of Digital Humanities." *Core.ac.uk*. https://doi.org/10.25366/2018.41.
Reid, Alexander. 2012. "Graduate Education and the Ethics of the Digital Humanities." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 350–67. Minneapolis, MN: University of Minnesota Press.
Rockwell, Geoffrey and Stéfan Sinclair. 2020. "Tremendous Mechanical Labor: Father Busa's Algorithm." *Digital Humanities Quarterly*. http://www.digitalhumanities.org/dhq/vol/14/3/000456/000456.html.
Scheler, Max. 1973. *Formalism in Ethics and Non-Formal Ethics of Values: A New Attempt toward the Foundation of an Ethical Personalism*. Evanston, IL: Northwestern University Press.
Smithies, James. 2017. *The Digital Humanities and the Digital Modern*. London: Palgrave.
Spiro, Lisa. 2012. "'This Is Why We Fight': Defining the Values of the Digital Humanities." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 16–35. Minneapolis, MN: University of Minnesota Press.
Terras, Melissa. 2016. "Crowdsourcing in the Digital Humanities." In *A New Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 420–39. London: Wiley-Blackwell.
Underwood, Ted. 2019. *Distant Horizons – Digital Evidence and Literary Change*. Chicago, IL: University of Chicago Press.
Wardrip-Fruin, Noah. 2009. *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*. Cambridge, MA: MIT Press.
Zuboff, Shoshana. 1988. *In the Age of the Smart Machine*. New York, NY: Basic Books.
Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.