Pre-peer review manuscript of Huggett, J. (2022) 'Archaeological Practice and Digital Automation', in: Watrall, E. and Goldstein, L. (eds.) *Digital Heritage and Archaeology in Practice: Data, Ethics, and Professionalism*, University Press of Florida: Florida, USA, pp. 275-304.

Archaeological Practice and Digital Automation

Jeremy Huggett, University of Glasgow, UK

1. Automation and archaeology

Archaeologists are accustomed to constructing narratives cast within periods frequently defined by the rise and ultimate replacement of distinctive technologies. Contemporary society has been characterized as being at the beginning of the Fourth Industrial Revolution, with the growth of technologies including artificial intelligence, machine learning, robotics, nanotechnology, 3D printing, and biotechnology (Leopold et al. 2016:5). Similarly, we are said to live in the Second Machine Age, characterized by sustained exponential improvements in computing, the dramatic growth in digital information, and accumulating recombinant innovation (Brynjolfsson and McAfee 2014). Some argue that automation today will be much more disruptive than previous technological interventions, in part because the traditional view associating computer-based automation with tasks that are tightly defined and clearly sequenced is beginning to break down. Digital devices are increasingly demonstrating the capacity to move into roles previously considered un-computable: for example, the development of complex algorithmic methodologies combining big data and deep learning has enabled computers to perform activities otherwise requiring cognitive ability and to improve themselves with little or no human intervention (Manyika et al. 2017:48). The smartphones in our pockets are powerful networked computers which act as digital assistants displaying apparent intelligence in ways that would have seemed like science-fiction barely twenty years ago. What are the implications for the future practice of archaeology?

Archaeologists and anthropologists might take comfort from the widely cited study by Frey and Osborne (2013; 2017) which analyzed the susceptibility of jobs to computerization. They estimated that 47% of total US employment has a high probability (0.7 or above) of computerization (Frey and Osborne 2017:269). However, they placed archaeology and anthropology 39th out of 702 occupations least likely to be susceptible to computerization, with a likelihood that algorithms will displace archaeologists/anthropologists of 0.007. By way of comparison, geographers are placed 222nd with a probability of 0.25, historians 283rd (0.44), and computer programmers 293rd (0.48) (see Frey and Osborne 2017:appendix A). This analysis is not uncontested: other studies reduce the proportion of jobs at risk of automation significantly by recognizing that technological change impacts on specific tasks within occupations rather than whole occupational groups. For example, a study for the Organization for Economic Co-operation and Development (OECD) suggested that only 9% of US jobs, rather than 47%, face high automatability (Arntz et al. 2016:25). A subsequent more extensive analysis for the OECD suggested around 10% of US jobs and 14% across 32 OECD countries are highly automatable (Nedelkoska and Quintini 2018:48). On the other hand, the McKinsey Global Institute estimates that about half of work activities could theoretically be automated using current technologies, and around 60% of occupations have the potential for at least one third of their constituent activities to be automated (Manyika et al. 2017:2). The situation is therefore not as straightforward as a simple human replacement and machine substitution story (Brynjolfsson and Mitchell 2017:1530). On balance it seems that, if anything, archaeology is even less susceptible to automation than Frey and Osborne calculate, especially in those areas requiring complex sensorimotor skills, but work on the

automation of repetitive recognition and identification skills in archaeology has been underway for some time.

2. Characterizing Digital Practice

Archaeology has been transitioning for some years towards a more automated, computerized form. This process is illustrated in Figure 1, building from a model of professional work developed by Susskind and Susskind (2015:195-210). Although the model could be seen to have an evolutionary aspect to it, in reality the categories overlap: some aspects of practice will remain at the same stage, others may only appear at a more 'advanced' stage (Susskind and Susskind 2015:197), still others may co-exist at several stages simultaneously (see Huggett 2021).

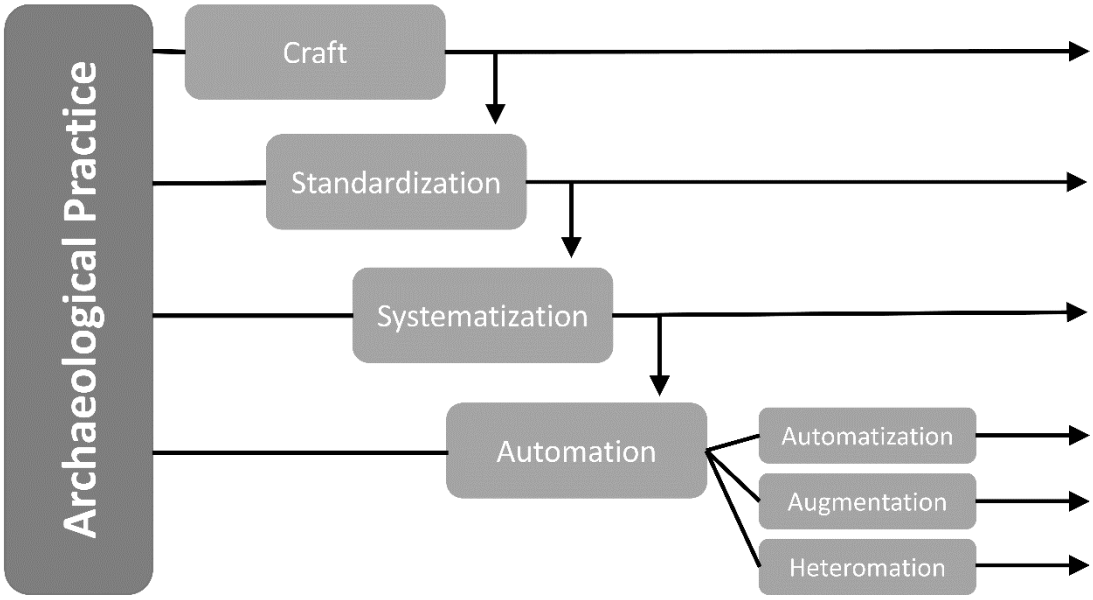

Figure 1: A model of archaeological practice

Archaeology is often conceptualized as a *craft* practice (Shanks and McGuire 1996; Olsen et al. 2012; see also Beale and Reilly 2017; Caraher 2016). For instance, there are technical craft aspects to a range of excavation and survey techniques (e.g., Poller 2018) which require the acquisition of a creative set of manual and visual skills (e.g., Shanks and McGuire 1996:78-83; Bradley 1997:70).

Standardization can be perceived as a move away from craft and consequently has often been resisted in archaeology (see Huggett 2012:540-545). While the adoption of standards may be viewed as restrictive to craft practice, the routinization of practice is seen to help prevent errors, ensure consistency, and reduce duplication of effort (Susskind and Susskind 2015:200; see also Yarrow 2008, for example). Standardization is therefore associated with enhanced productivity and efficiency in archaeology, although this too can be open to challenge (e.g., Huggett 2000:13-15). Crucially, standardization is an enabling feature upon which archaeological digital information infrastructures are predicated (Huggett 2012:542-543).

Systematization sees tools and technologies brought to bear in support of the archaeologist. These can range from typical office software (word processors, spreadsheets, etc.) to more complex tools such as geographical information systems and digital

photogrammetry, and to the digital infrastructures providing data search, retrieval, and archiving facilities.

With *automation*, archaeological tasks are delegated to mechanical, or software, or software-driven machines. This diverges from Susskind and Susskind's fourth stage, 'externalization' (2015:202), in which the practical expertise of professionals is made available online to non-specialists, a feature which is already present in the systematization of archaeology and is arguably a characteristic of the subject more generally. Different kinds of automation can be highlighted which overlap in different ways.

Most automation in archaeology is better characterized as *augmentation*, emphasizing technological assistance rather than outright replacement. Much automation in archaeology, as elsewhere, tends to support and reduce routinised and repetitive work, both in the archaeological office environment and in the field. For instance, robotized devices assist and augment the archaeologist, ranging from remotely operated to fully automated devices which perform tasks with increasing autonomy. These typically function in circumstances that might be inaccessible or unattractive to the human practitioner. For example, remotely operated vehicles and autonomous surface and underwater vehicles are increasingly commonplace in archaeology for sonar and optical survey (e.g., Buxton et al. 2016) to the extent that it has been suggested that underwater mapping and site photography will soon be delegated to robots, allowing humans to focus on more interesting and appropriate activities such as delicate excavation (Foley 2014:2081). Bio-mimetic or anthropomorphic robots such as Ocean One facilitate remote recovery and recording using stereoscopic cameras and haptic devices to enable the human pilot to feel the forces experienced by the robot's 'hands' (e.g., Khatib et al. 2016). Lab-based robotic devices have been employed in the analysis of tool use-wear (e.g., Pfleging et al. 2015) and for controlling rigs for automated photogrammetry and reflectance transformation imaging (e.g., Duffy et al. 2013:12), for example. Remotely controlled terrestrial and aerial drones are increasingly a feature of archaeological fieldwork (e.g., Prentiss 2016; Campana 2017), and autonomous characteristics such as collision avoidance are appearing in even basic models. At the furthest extreme, space archaeology has been described as "essentially a robotic frontier" (Gorman 2015:30).

An extension of automation and augmentation is *automatization* (Dertouzos 1997:83-4) where the computer takes on human information work rather than human physical work in a combination of information and automation. This allows tasks to be computerized that could not previously be automated, particularly through the development of data-driven 'intelligence amplification' methods. This can be characterized as the use of digital 'cognitive artefacts', employed to assist in the performance of a cognitive task, and able to represent, store, retrieve, and manipulate information (Huggett 2017:section 2). Typical examples in archaeology include expert systems and other forms of artificial intelligence tools (e.g., Barceló 2009), data mining, agent-based and network modelling (e.g., Wurzer et al. 2015; Brughmans et al. 2016), automated or semi-automated classification of pottery (e.g., Gualandi et al. 2016) and the detection of features in airborne imagery (e.g., Bennett et al. 2014). Indeed, given its dependence on data and information, automatization has become thoroughly embedded in the current practice of archaeology, as seen in the growing reliance on archaeological data archives and their associated search tools.

A third variant of automation is *heteromation* (Ekbia and Nardi 2014; 2017) which focuses on the role of humans as free or low-cost labor, performing critical tasks in support of the technological devices in contrast to automation which keeps the human at arm's length. Heteromation combines elements of augmentation and automatization in archaeology with the availability of voluntary labor, as can be seen in participatory digital archaeology

(Bonacchi 2017), for example. Indeed, a number of the heteromated systems identified by Ekbia and Nardi (2014:section 1.2; 2017:93-157) have also been examined within archaeology, including video games (e.g., Reinhard 2018), social media (e.g., Perry and Beale 2015), crowdsourced applications (e.g., Ridge 2013), and microwork systems (e.g., Bonacchi et al. 2014).

Different aspects of digital archaeological practice operate at each of these stages and often several at the same time. As a broad generalization, archaeology has seen a period of standardization and systematization since the 1980s, and is now witness to increasing levels of automation, and more recently, to technological developments which supersede or augment classic forms of automation by incorporating 'intelligent' devices which can perform apparently cognitive tasks and modify their own behavior in the light of their experience.

3. The Automated Archaeologist?

The totality of automation, augmentation, automatization, and heteromation is not far removed from Barceló's proposals for a specialized automated archaeologist capable of learning through experience to associate archaeological observations with explanations and using them to solve archaeological problems (2009:352; see also Barceló 2007; 2010). The specter of an automated archaeologist had been raised earlier in relation to the development of expert systems (e.g., Huggett 1985, Huggett and Baker 1985), but Barceló argues that the rejection of archaeological expert systems at the time was a consequence of insufficient formalization of the subject and resistance as "a result of absurd prejudice and the weights of individual authority" (Barceló 2010:20). Since then, the development of deep learning, neural networks, and machine learning in general has led to a revival of artificial intelligence methodologies, albeit in a more highly developed form.

This resurgence of interest arises from a congruence of three key areas (Manyika et al. 2017:24; see also Cath et al. 2018; Schwab 2016:22-24). First is the development of machine-learning, in particular employing deep learning and reinforcement-learning in neural networks, with Google's DeepMind and IBM's Watson being well-known examples. Second is an exponential increase in computing capacity and a concomitant drop in costs which enable larger and more complex computational models to be implemented. Finally, there has been a significant growth in the availability of digital data which can be used as training data within these models. Associated with this is an increase in human computing capacity through the rise of micro-tasking marketplaces such as the Mechanical Turk, Clickworker, and Figure Eight, which provide organizations and researchers access to large, cheap bodies of human workers undertaking piecework such as categorizing images, translating and tagging text, undertaking surveys and questionnaires, and creating training datasets, for instance (for example, Buhrmester et al. 2018; Sheehan 2018). The combination of these three factors lies behind advances in the automation of cognitive tasks. However, these advances are achieved by lowering expectations and measures for success rather than major technical breakthroughs (e.g., Darwiche 2018). Developments remain focused on relatively restricted areas of expertise rather than demonstrating high-level machine intelligence (HLMI), defined as when unaided machines can accomplish every task better and more cheaply than human workers (Grace et al. 2018:731). Surveys of artificial intelligence and robotics experts suggest that there is a 50% chance of HLMI by around 2065 (Walsh 2018:641; Grace et al. 2018:731) although non-experts predict it will be reached much sooner (Walsh 2018:641). Expert prediction of HLMI at the 90% level was around 2114, compared to non-expert expectations of 2060.

The range of reasons proposed for this lag between current machine-learning capabilities and HLMI are instructive. For example: "They are not yet good at putting knowledge into context, let alone improvising, and they have little of the common sense that is the essence of human experience and emotion. They struggle to operate without a pre-defined methodology … much work remains to be done integrating different capabilities into holistic solutions in which everything works together seamlessly." (Manyika et al. 2017:24).

In their analysis, Frey and Osborne (2013:262) identified three key technological bottlenecks: limited perception abilities and consequent problems undertaking tasks in unstructured work environments, the challenge of creative intelligence (the ability to come up with novel and valuable ideas and artefacts), and the challenge of social intelligence and the ability to negotiate complex relationships and cultural sensitivities. Again, these are not uncontested, and all are active areas of current research. For example, Boden (2016:67-72) argues that creativity is already a feature of AI though not as much as might be expected, partly due to problems of relevance and partly because of disagreements over whether an outcome is truly creative. Similarly, social intelligence is a core aspect of social robotics research facilitating interaction with human agents (e.g., Dumochel and Damiano 2017) although they are currently susceptible to variants of the uncanny valley.

The present limits of machine learning are demonstrated by characterizing those tasks that are suitable for machine learning applications. For example, Bynjolfsson and Mitchell (2017:1532-1533) identified 21 task properties for determining suitability for machine learning and reduced these to eight key criteria. Foremost amongst these is the requirement that the phenomenon or function should be easily described, well-defined, with unambiguous end goals. There should be large digital datasets available as training data which are representative of future, yet to be incorporated data, and there should be straightforward empirical associations in the data rather than a requirement for long chains of reasoning or reliance on common sense. It should not be necessary to explain decisions to the human user, and errors should be tolerated. Finally, no specialized dexterity, mobility or other physical skills should be required. These criteria can be somewhat flexible--for instance, a machine learning system can operate in less well-defined areas using statistical and probabilistic tools--but there are costs in doing so, such as reducing the transparency of the system's inferential reasoning.

These criteria make machine learning techniques a challenge since the task areas which can be described in these terms are clearly limited, and, furthermore, even these restrictive criteria are ambiguous. For example, how large must a dataset be to realistically train a machine learning system? How can the representativeness of a dataset be gauged--not just from an archaeological perspective, problematic as that is, but as a measure of its reliability in relation to as-yet unknown data? Can the outputs or conclusions be accepted in the absence of any explanation or justification for them? What levels of error are acceptable? Behind the batteries of sampling statistics, training sets and test data, such decisions are often ultimately heuristics which appear to work in that specific circumstance rather than being truly generalizable.

Ultimately these characteristics of machine learning tasks and their associated bottlenecks underline that there is still some way to go before the basic components identified by Barceló (2005:224) for an automatic archaeologist are met: mobile robotics to enable physical interaction with archaeological spaces, decision-making tools to determine best outcomes, perceptual elements linked to knowledge, and a cognitive and explanatory component.

4. A Focus on Archaeological Classification and Machine Learning

Nevertheless, there has been considerable effort in archaeology in recent years to employ machine learning in a number of application areas--in particular for the identification and automation of artefact classification (especially pottery) and for the automated detection and identification of features from aerial or satellite imagery (see, for example, the overview in Davis 2020). In both areas, proponents generally see the tools as complementing and supporting archaeological practice (e.g., Bennett et al. 2014:897; Makridis and Daras 2012:3; Roman-Rangel, Jimenez-Badillo and Marchand-Maillet 2016:2; Wright and Gattiglia 2018:61) rather than replacing the role of the human expert. However, their ultimate use may result in the replacement of expertise: the ability of inexperienced practitioners to perform a task in the field rather than relying on specialist expertise (e.g., Brough and Parfitt 1984:49; Ennals and Brough 1982; Tyukin et al. 2018), for example. The very availability of a successful tool may inadvertently lead to its substitution for a human expert.

4.1 Machine classification

Such tools and their associated controversies have a long gestation dating back over thirty years. For example, expert systems were proposed to allow a relatively inexperienced user to characterize beaker types (Bishop and Thomas 1984), to age horse teeth (Brough and Parfitt 1984), to classify Mediterranean wine amphorae (Bourelly and Chouraqui 1984), and to simulate reasoning around Seljukid and Greek iconography (Lagrange and Renaud 1985). All were based on structured data and established systems of classification, and generally employed production rules which modelled the problem domain and enabled the system to narrow down and classify a case put to it by a user in response to targeted questions. At the time, criticism focused on the inflexibility of the systems, their narrow and shallow focus, the inappropriateness of the abstraction and reduction required to model their inferential rule bases, and the overly optimistic assessment of the tools (e.g., Huggett 1985; Huggett and Baker 1985). More recently, there has been a growth in development of deep learning tools which seek to reduce a complex problem into a series of simple nested mappings, each described by a different layer of the model and generated automatically from data (e.g., Goodfellow et al. 2016:5-7), rather than attempting to manually create a set of formal rules which adequately described expert reasoning (see Figure 2). For example, several systems have focused on the automation of pottery classification (e.g., Makridis and Daras 2012; Teddy et al. 2015; Roman-Rangel et al. 2016; Hein et al. 2018; Rasheed and Nordin 2020; Wright and Gattiglia 2018). Others have taken related approaches to pottery, such as automating the creation of 3D models from 2D catalog drawings (Banterle et al. 2017), an intelligent search engine for pottery retrieval (Benhabiles and Tabia 2016), the automated classification of mineral inclusions in pottery (Aprile et al. 2014), automatic Munsell color characterization of sherds (Milotta et al. 2018), and shape- and decoration-based identification of pottery (Itkin et al. 2019), for example. Similar machine learning approaches have been applied to automated feature extraction from LiDAR and satellite imagery (see overviews in Bennett et al. 2014; Optiz and Herrmann 2018; Davis 2019; 2020).

4.2 Training and tuning

These machine learning approaches rely on the availability of very large datasets for training purposes as well as large amounts of computing power for generating the deep models, and this presents challenges for archaeology. In particular, most pottery-related systems have only relatively small datasets available, typically because of the time and expense involved in the creation of suitable datasets (e.g., Wright and Gattiglia 2018:62). For instance, the ArchAIDE project used 274 images of sherds from 49 out of 84 types for its appearance-based

recognition analysis (Gascón et al. 2019:14), and 381 images from 42 different types for its shape-based recognition (Gascón et al. 2019:22). Arch-I-Scan scanned ten vessels and created an unspecified number of images for testing (Tyukin et al. 2018). Elsewhere, Benhabiles and Tabia (2016:4) used a dataset consisting of 1012 digitized, manually modelled and semi-automatically generated 3D models, Cintas and colleagues (Cintas et al. 2019:108) used a set of 1133 profile images, while Aprile and colleagues (Aprile et al. 2014:263) used 14 thin sections to assess their methodology.

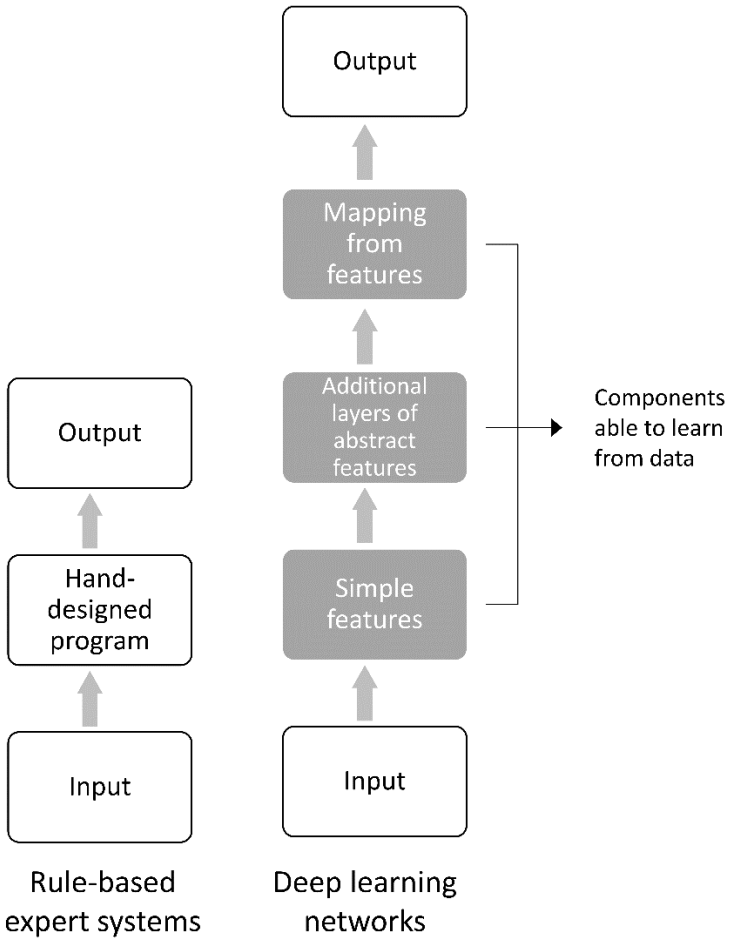

Figure 2: Flowchart comparing the different parts of expert systems and deep learning networks (adapted from Goodfellow et al. 2016:10).

The use of a small training dataset risks overfitting, where the resulting system is overly dependent on the characteristics of the training data and performs poorly on additional unseen data. On the other hand, if large datasets were available, training a deep network with many layers is computationally expensive and can take many hours, even days. Consequently, many deep learning applications in archaeology have employed pre-trained neural networks, commonly based on the generic ImageNet dataset (e.g., Deng et al. 2009) consisting of over 14 million annotated images across more than 20,000 categories. For example, ArchAIDE employed a pre-trained version of the ResNet-101 network based upon ImageNet (Itkin et al. 2019:9) for their decoration-based tool and a variant of the PointNet network (Qi et al. 2017) for their shape-based tool (Itkin et al. 2019:8). Benhabiles and Tabia (2019:3) employed AlexNet which was also created against ImageNet, while Roman-Rangel and colleagues (Roman-Rangel et al. 2016:12-13) used a subset of the SHREC'13 dataset of 3D models of generic objects (Li et al. 2013). Such pre-trained models have proved

successful in their original classification and identification tasks and have been subsequently applied in many other image recognition contexts.

Although the use of pre-trained models is widespread and can provide substantial savings, they can also give rise to problematic results: for instance, drawing inferences from features other than the subject of the images, becoming reliant on texture rather than shape, and suffering from selection bias (e.g., Ribani and Maregoni 2019:50), famously resulting in facial recognition systems displaying racial bias (e.g., Buolamwini and Gebru 2018), for example. This underlines that even the very largest datasets may reflect human and machine bias. Furthermore, it can be argued that training and test data should be related to each other to some extent (e.g., Pan and Yang 2010:1345-1347). Put simply, images of airplanes and animals such as found in ImageNet are arguably of little relevance in identifying pottery sherds since the features of interest are so different (see also Opitz and Herrmann 2018:30). However, in pre-trained models the uppermost layers of the network tend to relate to low-level functions such as edge detection and geometric shapes while lower layers are more closely related to the identification of the original training data. Hence in a new application the uppermost layers of the pre-trained network may be frozen and used as they stand while the lower layers are fine-tuned to adapt to the new training data (Ribani and Maregoni 2019:51), although the process for achieving this selection and tuning is not well-defined and is rarely discussed in published archaeological examples.

Related to the freezing and fine-tuning of layers is the question of the number of layers in the network in the first place, since the selection of a pre-trained model of a certain depth is rarely explained. For example, it is often assumed that the deeper the network (the more layers), the more complex and therefore more reliable the performance. However, while ResNet-101 with 101 layers may be more accurate in categorizing ImageNet images than ResNets with fewer layers (He et al. 2016:774-775), a ResNet with significantly more layers does not necessarily improve performance (He et al. 2016:777), and the marginal gains achieved by each additional layer diminishes with depth (Wu et al. 2019:119). It may be that wider rather than deeper networks are more effective through, for instance, increasing the resolution of the input images which can both reduce the depth of the network and enhance performance (Wu et al. 2019:122-125).

These issues underline the need for transparency about decision-making during the design and implementation phases of machine learning tools, but even in the most detailed accounts key information is frequently missing. This lack of transparency in the developmental phase is compounded by the black-boxing of the processes underlying the results, with the difficulty of understanding increasing with the depth of the network. Furthermore, machine learning systems do not fail gracefully but will tend to force any object into existing known categories rather than recognizing it as something not previously encountered and flagging it as distinctively new. While currently such tools do not carry the authority of an expert, this overall lack of clarity should be of concern (Huggett 2021).

5. Implications of Automation

The ubiquity and pervasiveness of machine learning algorithms across a broad range of tasks, coupled with their frequently hidden agency, emphasizes the importance of considering the conditions for their informed application. This is because the routinized use of such devices is associated with what Pasquale describes as ‘automation bias’: “an assumption that a machine-driven, software-enabled system is going to offer better results than human judgement” (Pasquale 2015:107). A taken-for-granted relationship with these devices as simply means to ends is problematic and gives rise to their unthinking prioritization. There is

a human susceptibility to fetishize, habituate, and be enchanted by digital tools (Smith 2018:7-11), but at the same time, the complexity, opacity, inscrutability of the tools themselves make it difficult to confirm their proper functioning other than through the management of their inputs and observation of their outputs. The challenge therefore is to reduce, if not resolve, the shortcomings in this relationship through a dual approach: an archaeology of the digital object, and an anthropology of the digital user (Huggett 2017; Seaver 2018). In the meantime, and developing the issues raised above, four interrelated key areas may be identified: the transparency, explainability, and authority associated with digital devices, together with the need for ethics development (see also Huggett 2021).

5.1 Transparency

Transparency is necessary to break out of the algorithmic black box and is a theme that has been extensively debated (e.g., Domingos 2015; Finn 2017; Fry 2018; O'Neil 2016; Pasquale 2015; Steiner 2013). Increasing transparency is seen to address concerns over algorithmic bias, responsibility, liability, and accountability, and consequently enables informed decisions about the deployment and use of such devices. However, the assumption that transparency can reveal hidden methods, inferences, and functions is limited by the complexities of the underlying systems. Lifting the lid of the black box can simply expose an impenetrable mass that even expert developers will have difficulty disentangling because of the depth and range of interrelationships and interdependencies. Access to underlying code (if available) is not in itself a means of increasing transparency given the size and complexity of the code and shortcomings of documentation. This becomes even more problematic where machine learning programs essentially reprogram themselves as they learn from data, rather than the software being written in its entirety by human programmers.

This does not mean that these devices and their operations are unknowable, but there needs to be a distinction drawn between seeing into or through such devices and understanding their functions. For example, Bucher prefers to see them as "neither black nor box, but a lot more gray, fluid, and entangled than we may think" (2016:94), and consequently proposes three steps to consider in researching algorithms: identifying those parts that can and cannot be known and understanding that they cannot be simply read as a text; focusing on the performative aspects of the device as it acts on the external world; and understanding the histories, evolutions, and contexts of the devices (Bucher 2016:85-93). Similar approaches are proposed by Kitchin (2017) and Ananny and Crawford (2018) who emphasize the benefits and shortcomings of different approaches to transparency. In particular, Ananny and Crawford argue that rather than looking inside the systems, we should look *across* them, "seeing them as sociotechnical systems that do not contain complexity but enact complexity by connecting to and intertwining with assemblages of humans and non-humans" (2018:974). In a similar vein, a layered series of requirements for transparency can be defined in terms of a combination of one or more of a critical reading of the code, a critical appreciation of the device or package, a critical engagement with the creative process of its design, and a critical understanding of its subsequent application within an archaeological context (Huggett 2017:section 8; see also Huggett 2021).

5.2 Explanation

Demands for the transparency of agential digital devices is frequently seen to require similar behavior to human experts, who are expected to be able to explain their actions and decisions as a means of checking and confirming their expertise, absence of bias, and overall performance. In certain respects, this can be relatively straightforward: the kind of rule-based expert systems developed in the 1980s could simply cite their chain of reasoning in terms of

the rules applied and the heuristic weights associated with each decision, but as the explanation moved up from the particular to the general, they ultimately ran out of rules to invoke (Huggett 1985:136). Machine learning and neural networks can be resistant to such approaches: the complexity of their decision trees rapidly become uninterpretable since the conceptualization of the problem space is machine-based rather than human-based. Consequently, the risk is that systems are developed and used that are not fully understood and that may be prone to unanticipated and unrecognized error, whereas it should be the case that "explanation is at the heart of a responsible, open data science, across multiple industry sectors and scientific disciplines" (Guidotti et al. 2018:2). Although some argue that explanations may be unnecessary where the decisions are not crucial or where there are no unacceptable consequences (e.g., Doshi-Velez and Kim 2017:3; Guidotti et al. 2018:5), for black boxes to be capable of being trusted, they require 'explainability': to have "the capacity to defend their actions, provide relevant responses to questions, and be audited" (Gilpin et al. 2018, 1). It seems improbable that the application of such systems within an archaeological context would not require a similar level of explainability, even if there are no critical or life-threatening consequences, since the inferences and assumptions underlying the system and the context of its application remain important to archaeological understanding and interpretation (c.f. Moses 2018).

At present, most of the methodologies for generating explanations rely on techniques such as simplification (decision tree pruning, for instance), the creation of proxy models, automatic rule-generation, or reverse-engineering (e.g., Gilpin et al. 2018:3-4; Lipton 2018:38-42). All are in different ways highly complex and yet still incomplete, not least because there is no agreement on what an explanation derived from such systems should consist of, or what properties it should have (Guidotti et al. 2018:56; Lipton 2018:37). Indeed, some suggest that the explanatory bar is set too high in the first place, that a double-standard is applied in that the requirement for explanation is based upon two false assumptions: "… that it is fair to impose a higher standard of transparency on such tools than would ordinarily be imposed on human decisionmakers ... or … that human decisions are comparatively more transparent than algorithmic decisions, because they can be inspected at a depth to which AI is not presently amenable" (Zerilli et al. 2019:680). They conclude that the kinds of explanations that we currently seek but cannot obtain from artificial intelligence are ones that we cannot obtain from humans either. What constitutes an appropriate machine-derived archaeological explanation and at what level it should apply therefore remains undetermined.

5.3 Authority

One of the key characteristics of artificial surrogates, whether expressed in hardware or software, is that they possess authority. Shirky provided an early definition of algorithmic authority: "the decision to regard as authoritative an unmanaged process of extracting value from diverse, untrustworthy sources, without any human standing beside the result saying 'Trust this because you trust me'" (Shirky 2009). More recently, Lustig and Nardi broadened out Shirky's definition by defining algorithmic authority as "the trust in algorithms to direct human action and to verify information, in place of trusting or preferring human authority" (2015:743), recognizing the authority of algorithms to direct human action as well as to decide which information is true.

As a source of authority, therefore, such devices are capable of exercising agency and provide opportunities for human authority to be delegated to them (Huggett 2021). This should make their transparency and explainability even more significant, but both daily news stories and more extensive studies suggest that instead they are frequently accepted and used

with relatively little consideration. For example, in a study looking at algorithm adoption it was found that simply knowing that other people were using the algorithm made it more than twice as likely to be adopted, even in the face of knowledge that it gave imperfect advice (Alexander et al. 2018). Furthermore, it was observed that using an algorithm reduced the cognitive load of the human user--a commonplace with satellite navigation systems, for instance--from which it was concluded that "participants did not monitor the algorithm when there was no information about it—precisely the situation in which monitoring is most warranted." (Alexander et al. 2018:287). So the authority of the system as a neutral, trustworthy tool is often uncritically adopted or, if considered at all, is adopted on the basis of peer use. The human practitioner accepts the authority of the tool in a form of deep automation bias (Strauß 2018:10), frequently ignorant of the restrictions and limitations hidden within and the corresponding implications for decision-making, both in terms of those made by the device itself but also those of the human user themselves, and their subsequent behavior. This can lead to what has been described as the 'de-responsibilization' of the human actors, a tendency to hide behind the machine and to assume that its results are correct by default (Mittelstadt et al. 2016:13).

5.4 Ethics and algorithms

The challenges associated with transparency, explainability, and authority of these devices requires a consideration of the ethics of their design, development, and application. What limits should be applied to their use? Can these agential devices contain ethical programming, and consequently act ethically? What moral responsibilities would such devices have towards us, and what responsibilities might be required of us towards them in turn? These kinds of questions have been debated within artificial intelligence and robotics for some years (e.g., Boddington 2017; Gunkel 2012; Wallach and Allen 2009), but the development of ethics within a machine environment remains in relative infancy (e.g., Trussell 2018). The fluid nature of algorithms makes an already complex problem even more difficult. As Dourish observes, there is a difference between what an algorithm makes possible and what an implementation of that algorithm makes feasible (Dourish 2017:213): while the essential algorithm remains the same, the technical infrastructure it is embedded within and the context of its application and use changes its capability and potential, suggesting that a case-by-case consideration is required for the development of a digital ethics of agential devices.

Digital ethics within archaeology are in a relatively early stage of development (see overviews by Colley (2015) and Dennis (2020), for example). There have been discussions concerning digital public archaeology (e.g., Richardson 2018), remote sensing (e.g., Myres 2010), 3D reconstruction (e.g., Kamash 2017; Khunti 2018), digital heritage (e.g., Stobiecka 2020), 3D models of human remains (e.g,. Ulguim 2018) and archaeogaming (e.g., Graham 2020), while a volume on digital ethics and practice in archaeology (Wilson and Edwards 2015) only obliquely addressed ethical issues. As yet, there are none concerning archaeological applications of machine learning and automation. In addition to ethical matters surrounding transparency, explainability, trust and authority, other challenges include issues of augmentation and the potential replacement of the human agent, control and oversight of the devices, and questions surrounding reproducibility since--like a human expert--machine learning systems adapt to new data and new inputs and so may draw different conclusions at different times.

A Participatory Turn

Floridi (2014:25) observes that a characteristic of technology is its 'in-betweenness': the way in which it mediates between the human actor and the world about them (see Figure 3). He defines first-order technologies as mediating between people and nature (an axe enables people to chop wood, for instance). Second-order technologies mediate between people and other technologies (a spanner enables people to tighten bolts to build an engine, for example). Third-order technologies mediate between one technology and another, leaving the human component a relatively peripheral beneficiary (as, for example, with the Internet of Things) (Floridi 2014:25-34).

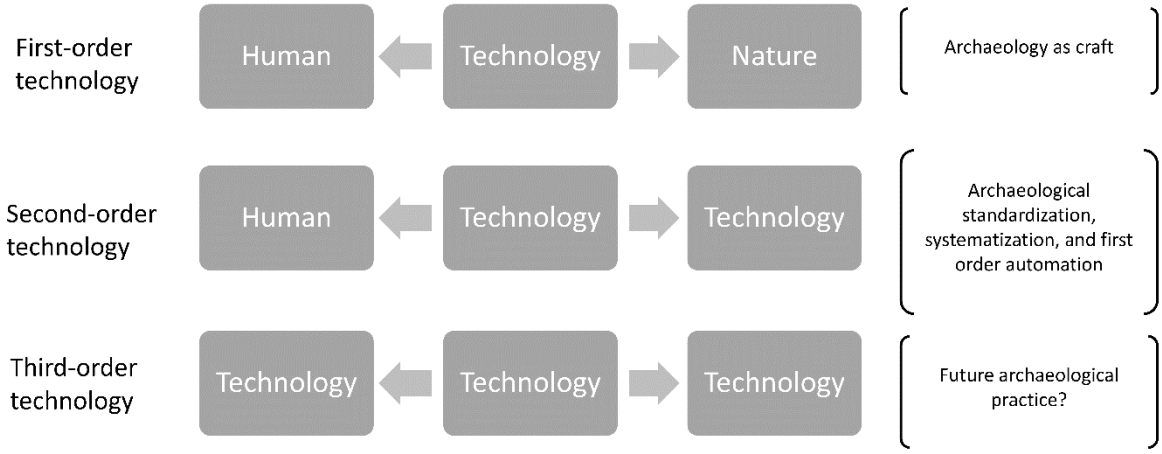

Figure 3: Human-technology relationships (adapted from Floridi 2014:26-29).

The craft of archaeology might be conceived as operating at a first-order level, with the direct, literally hands-on relationship between the archaeologist and physical, material remains mediated by a trowel or similar. A growing proportion of archaeological practice operates at a second-order level, with tools inserting themselves into archaeological tasks and standardization and systematization facilitating the automation of an increasing amount of repetitive work. As these devices spread across a broader range of tasks and acquire apparent intelligence through deep learning and other techniques, archaeology potentially moves into a third-order level of practice, with highly automated cognitive devices set loose and the human archaeologist increasingly relegated to observer status.

Archaeology is not at the third-order stage yet, although the predictions of business analysts, futurists, and other scholars suggest that this is largely a matter of time. For instance, if there is a 50% chance of high-level machine intelligence within the next 40 years (Walsh 2018:641; Grace et al. 2018:731), significant developments in and applications of a broad range of more focused cognitive devices across archaeology might be anticipated by that stage. Nor can archaeologists claim that specific aspects of the archaeological taskscape are in some way un-computable or un-automatable because experience suggests that this again is primarily a question of time: for example, the kinds of robotic devices used in archaeology today would have been seen as science fiction some forty years ago. Although we may--and likely will--resist the introduction of cognitive robotic devices and intelligent software tools in the future, the economics of archaeological practice will often determine their introduction into normalized workstreams, as is increasingly evident in the growth of Structure from Motion imagery used in archaeological recording, for instance.

It would be easy to adopt an overly utopian/dystopian or deterministic approach to technological futures in archaeology or indulge in what Broussard calls 'technochauvinism' (2018:7-8), a variant of technological determinism linking the perceived inevitability of technological advancement with a range of other socio-political beliefs such as techno-libertarianism, neoliberalism, and globalization. However, the critical issue is to turn the analytical gaze onto the centrality of the human in digital archaeological practice now and in the future. For example, as described above in relation to machine learning, digital technologies may be able to handle the mundane but they are unable to take care of the edge cases (Broussard 2018:176-7): the unexpected, unpredictable diversions from the norm which require human intervention to resolve. Systems which seek to resolve such edge cases are called *human-in-the-loop* (HITL) designs which assign the human agent participant rather than observer status. For example, HITL can be employed in machine learning where, when a confidence level is below a certain value, the system refers to a human expert and then incorporates the human judgement into the knowledge base. This form of incremental learning is employed in a wide range of areas such as medical informatics, autonomous vehicles, and image recognition, for example. However, the role of the human expert is primarily one of training, testing, annotating, and tuning the system in what is ultimately still an asymmetric relationship between cognitive digital object and human subject since the ultimate aim is the full automation of the process.

What is proposed here is an *archaeologist-in-the-loop* approach to intelligent software and hardware systems, reflecting a digital participatory turn. This borrows from concepts of participatory research in geography and anthropology (e.g., Gubrium et al. 2015; Harper and Gubrium 2017) as well as in archaeology (e.g., Harris 2012; Kiddey 2020) which seek to reconfigure who produced knowledge and for whom, raising issues of power, trust, and ownership, and adopting an explicitly collaborative approach. Here, though, the participatory model is flipped to one which seeks to reconfigure the digital divide between human archaeologist and machine, ensuring the human practitioner retains a critical influential and strategic oversight relative to the machine, rather than adopting a subservient, compliant, acquiescent role. This kind of critical engagement is not straightforward, but it is necessary to reduce the risks of fetishization, habituation, and enchantment. Focusing on the human archaeologist-in-the-loop in this participatory way should also help ensure that the issues of authority, transparency, explainability, and the ethics associated with our cognitive devices remain foregrounded and prioritized into the future. Furthermore, retaining an emphasis on the human practitioner at the center of the digital archaeological engagement means it should be possible to avoid an overly scientistic, positivistic, or instrumentalist perspective. By recognizing that these devices are not neutral but carry with them inscriptions and delimiting assumptions which affect their application and outcomes for both archaeologists and archaeology more generally, we can retain space for inspiration, intuition, subjectivity, and the primacy of human cognition within future archaeological practice.

References Cited

Alexander, Veronika, Collin Blinder and Paul J. Zak

2018 Why trust an algorithm? Performance, cognition, and neurophysiology. *Computers in Human Behavior* 89: 279-288. DOI: 10.1016/j.chb.2018.07.026

Ananny, Mike and Kate Crawford

2018 Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society* 20(3):973-989. DOI: 10.1177/1461444816676645

Aprile, Anna, Giovanna Castellano and Giacomo Eramo

2014 Combining image analysis and modular neural networks for classification of mineral inclusions and pores in archaeological potsherds. *Journal of Archaeological Science* 50:262-272. DOI: 10.1016/j.jas.2014.07.017

Arntz, Melanie, Terry Gregory and Ulrich Zierahn

2016 The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. *OECD Social, Employment and Migration Working Papers No. 189*. OECD Publishing, Paris. DOI: 10.1787/5jlz9h56dvq7-en

Banterle, Francesco, Barak Itkin, Matteo Dellepiane, Lior Wolf, Marco Callieri, Nachum Dershowitz and Roberto Scopigno

2017 VASESKETCH: Automatic 3D Representation of Pottery from Paper Catalog Drawings. *14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto*, pp. 683-690. Institute of Electrical and Electronics Engineers, New York NY. DOI: 10.1109/ICDAR.2017.117

Barceló, Juan A.

2007 A Science Fiction Tale? A robot called archaeologist. In *The World is in your Eyes. Proceedings of the 33rd Computer Applications and Quantitative Methods in Archaeology Conference (CAA2005), Tomar, Portugal*, edited by Alexandra Figueiredo and Gonçalo Leite Velho, pp. 221–229. Instituto Politécnico de Tomar, Tomar, Portugal.

2009 *Computational Intelligence in Archaeology.* Information Science Reference, Hershey PA.

2010 From Computable Archaeology to Computational Intelligence: New Prospects for Archaeological Reasoning. In *Cyber-Archaeology*, edited by Maurizio Forte, pp. 15-29. British Archaeological Reports International Series 2177. Archaeopress, Oxford.

Beale, Gareth and Paul Reilly

2017 Digital Practice as Meaning Making in Archaeology. *Internet Archaeology* 44. DOI: 10.11141/ia.44.13

Benhabiles, Halim and Hedi Tabia

2016 Convolutional neural network for pottery retrieval. *Journal of Electronic Imaging* 26(1): 011005. DOI: 10.1117/1.JEI.26.1.011005

Bennett, Rebecca, Dave Cowley and Véronique De Laet

2014 The data explosion: tackling the taboo of automatic feature recognition in airborne survey data. *Antiquity* 88(341):896–905. DOI: 10.1017/S0003598X00050766

Bishop, M.C. and J. Thomas

1984 'BEAKER'--An Expert System for the BBC Micro. In *Computer Applications in Archaeology 1984. Conference Proceedings*, edited by Sue Laflin, pp. 56-62. Centre for Computing and Computer Science, University of Birmingham, Birmingham. https://proceedings.caaconference.org/paper/08_bishop_thomas_caa_1984/

Boddington, Paula

2017 *Towards a Code of Ethics for Artificial Intelligence.* Springer, Cham. DOI: 10.1007/978-3-319-60648-4

Boden, Margaret

2016 *AI: Its Nature and Future*. Oxford University Press, Oxford.

Bonacchi, Chiara

2017 Digital Media in Public Archaeology. In *Key Concepts in Public Archaeology,* edited by Gabriel Moshenka, pp. 60-72. UCL Press, London. DOI: 10.14324/111. 9781911576419

Bonacchi, Chiara, Andrew Bevan, Daniel Pett, Adi Keinan-Schoonbaert, Rachel Sparks, Jennifer Wexler and Neil Wilkin

2014 Crowd-sourced Archaeological Research: The MicroPasts Project. *Archaeology International* 17:61-68. DOI: 10.5334/ai.1705

Bourrelly, Louis and Eugène Chouraqui

1984 Systeme Expert et Simulation d'un Raisonnement en Archeologie. *Informatique et Gestion* 151:46-51.

Bradley, Richard

1997 To see is to have seen: craft traditions in British field archaeology. In *The Cultural Life of Images: visual representation in archaeology,* edited by Brian L. Molyneaux, pp. 62-72. Routledge, London.

Brough, D.R. and N. Parfitt

1984 An Expert System for the Ageing of a Domestic Animal. In *Computer Applications in Archaeology 1984. Conference Proceedings*, edited by Sue Laflin, pp. 49-55. Centre for Computing and Computer Science, University of Birmingham, Birmingham. https://proceedings.caaconference.org/paper/07_brough_parfitt_caa_1984/

Broussard, Meredith

2018 *Artificial Unintelligence: How Computers Misunderstand the World*. MIT Press, Cambridge MA.

Brughmans, Tom, Anna Collar and Fiona Coward

2016 *The Connected Past: Challenges to Network Studies in Archaeology and History*. Oxford University Press, Oxford.

Brynjolfsson, Erik and Andrew McAfee

2014 *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W.W. Norton, New York.

Brynjolfsson, Erik and Tom Mitchell

2017 What can machine learning do? Workforce implications. *Science* 358(6370):1530-1534. DOI: 10.1126/science.aap8062

Bucher, Tania

2016 Neither Black Nor Box: Ways of Knowing Algorithms. In *Innovative Methods in Media and Communication Research*, edited by Sebastian Kubitschko and Anne Kaun, pp. 81-98. Palgrave Macmillan, Cham. DOI: 10.1007/978-3-319-40700-5_5

Buhrmester, Michael D., Sanaz Talaifar and Samuel D. Gosling

2018 An Evaluation of Amazon's Mechanical Turk, Its Rapid Rise, and Its Effective Use. *Perspectives on Psychological Science* 13(2):149-154. DOI: 10.1177/1745691617706516

Buolamwini, Joy and Timnit Gebru

2018 Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research* 81:77-91. http://proceedings.mlr.press/v81/buolamwini18a.html

Buxton, Bridget, Jacob Sharvit, Dror Planer, Nikola Mišković and John Hale

2016 An ASV (Autonomous Surface Vehicle) for Archaeology: The Pladypos at Caesarea Maritima, Israel. In *Mobilizing the Past for a Digital Future: The Potential of Digital Archaeology*, edited by Erin Walcek Averett, Jody Michael Gordon and Derek B. Counts, pp. 279-315. The Digital Press at the University of North Dakota, Grand Forks ND.

Campana, Stefano

2017 Drones in Archaeology. State-of-the-art and Future Perspectives. *Archaeological Prospection* 24:275-296. DOI: 10.1002/arp.1569

Caraher, William

2016 Slow Archaeology: Technology, Efficiency, and Archaeological Work. In *Mobilizing the Past for a Digital Future: The Potential of Digital Archaeology*, edited by Erin Walcek Averett, Jody Michael Gordon and Derek B. Counts, pp. 421–41. The Digital Press at the University of North Dakota, Grand Forks ND.

Cath, Corinne, Sandra Wachter, Brent Mittelstadt, Mariarosaria Taddeo and Luciano Floridi

2018 Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. *Science and Engineering Ethics* 24(2):505-528. DOI: 10.1007/s11948-017-9901-7

Cintas, Celia, Manuel Lucena, José Manuel Fuertes, Claudio Delrieux, Pablo Navarro, Rolando González-José and Manuel Molinos

2019 Automatic feature extraction and classification of Iberian ceramics based on deep convolutional networks. *Journal of Cultural Heritage* 41(1):106-112. DOI: 10.1016/j.culher.2019.06.005

Colley, Sarah

2015 Ethics and Digital Heritage. In *The Ethics of Cultural Heritage*, edited by Tracy Ireland and John Schofield, pp. 13-32. Springer New York, New York NY. DOI: 10.1007/978-1-4939-1649-8_2

Darwiche, Adnan

2018 Human-Level Intelligence or Animal-Like Abilities? *Communications of the ACM* 61(10):56-67. DOI: 10.1145/3271625

Davis, Dylan S.

2019 Object-Based Image Analysis: A Review of Developments and Future Directions of Automated Feature Detection in Landscape Archaeology. *Archaeological Prospection* 26:155-163. DOI: 10.1002/arp.1730

2020 Defining what we study: The contribution of machine automation in archaeological research. *Digital Applications in Archaeology and Cultural Heritage* 18:e00152. DOI: 10.1016/j.daach.2020.e00152

Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li and Li Fei-Fei

2009 ImageNet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition, Miami FL*, pp. 248-255. Institute of Electrical and Electronics Engineers, New York NY. DOI: 10.1109/CVPR.2009.5206848

Dennis, L. Meghan

2020 Digital Archaeological Ethics: Successes and Failures in Disciplinary Attention. *Journal of Computer Applications in Archaeology* 3(1):210-218. DOI: 10.5334/jcaa.24

Dertouzos, Michael

1997 *What Will Be: how the new world of information will change our lives.* Piatkus, London.

Domingos, Pedro

2015 *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Penguin Random House, London.

Doshi-Velez, Finale and Been Kim

2017 Towards a Rigorous Science of Interpretable Machine Learning. arXiv:1702.08608v2 [stat.ML] https://arxiv.org/abs/1702.08608v2

Dourish, Paul

2017 *The Stuff of Bits: An Essay on the Materialities of Information*. MIT Press, Cambridge MA.

Duffy, Sarah, Paul Bryan, Graeme Earl, Gareth Beale, Hembo Pagi and Eleni Kotouala

2013 *Multi-light Imaging for Heritage Applications.* English Heritage, Swindon.

Dumouchel, Paul and Luisa Damiano

2017 *Living with Robots*. Harvard University Press, Cambridge MA.

Ekbia, Hamid R. and Bonnie A. Nardi

2014 Heteromation and its (dis)contents: The invisible division of labor between humans and machines. *First Monday* 19(6). DOI: 10.5210/fm.v19i6.5331

2017 *Heteromation, and Other Stories of Computing and Capitalism.* MIT Press, Cambridge MA.

Ennals, R. and D. Brough

1982 Representing the Knowledge of the Expert Archaeologist. In *Computer Applications in Archaeology 1982. Conference Proceedings*, edited by Sue Laflin, pp. 56-62. Centre for Computing and Computer Science, University of Birmingham, Birmingham. https://proceedings.caaconference.org/paper/05_ennals_brough_caa_1982/

Finn, Ed

2017 *What Algorithms Want: Imagination in the Age of Computing*. MIT Press, Cambridge MA.

Floridi, Luciano

2014 *The 4th Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press, Oxford.

Foley, Brendan

2014 Deep-Water Archaeology. In *Encyclopedia of Global Archaeology,* edited by Claire Smith, pp. 2079-2081. Springer, New York. DOI: 10.1007/978-1-4419-0465-2_613

Frey, Carl Benedikt and Michael A. Osborne

2013 The Future of Employment: How Susceptible are Jobs to Computerisation? Oxford Martin Programme on Technology and Employment, University of Oxford. https://www.oxfordmartin.ox.ac.uk/publications/view/1314

2017 The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting & Social Change* 114:254-280. DOI: 10.1016/j.techfore.2016.08.019

Fry, Hannah

2018 *Hello World: How to Be Human in the Age of the Machine*. Doubleday, London.

Gascón, Eva Miguel, Gabriele Gattiglia, Llorenç Vila, Miguel Ángel Hervás, Nevio Dubbini, and Luis Alejandro García

2019 *D 8.1 Report of final recommendation for WP8 version: 1.1 (08/07/2019)* http://www.archaide.eu/documents/20181/128971/D8.1/

Gilpin, Leilani H., David Bau, Ben Z. Yuan, Ayesha Bajwa, Michael Specter and Lalana Kagal

2018 Explaining Explanations: An Approach to Evaluating Interpretability of Machine Learning. arXiv:1806.00069v2 [cs.AI] https://arxiv.org/abs/1806.00069v2

Goodfellow, Ian, Yoshua Bengio and Aaron Courville

2016 *Deep Learning* (Adaptive Computation and Machine Learning Series) MIT Press, Cambridge MA.

Gorman, Alice

2015 Robot Avatars: The Material Culture of Human Activity in Earth Orbit. In *Archaeology and Heritage of the Human Movement into Space,* edited by Beth Laura O'Leary and P.J. Capelotti, pp. 29-47. Springer, Cham. DOI: 10.1007/978-3-319-07866-3_3

Grace, Katja, John Salvatier, Allan Dafoe, Baobao Zhang and Owain Evans

2018 When Will AI Exceed Human Performance? Evidence from AI Experts. *Journal of Artificial Intelligence Research* 62:729-754. DOI: 10.1613/jair.1.11222

Graham, Shawn

2020 An Approach to the Ethics of Archaeogaming. *Internet Archaeology* 55. DOI:10.11141/ia.55.2

Gualandi, Maria Letizia, Roberto Scopigno, Lior Wolf, Julian Richards, Jaume Buxeda i Garrigos, Michael Heinzelmann, Miguel Angel Hervas, Llorenc Vila and Massimo Zallocco

2016 ArchAIDE--Archaeological Automatic Interpretation and Documentation of cEramics. In *GCH 2016 Eurographics Workshop on Graphics and Cultural Heritage*, edited by Chiara Eva Catalana and Livio De Luca, pp. 203-206. The Eurographics Association, Geneva. DOI: 10.2312/gch.20161408

Gubrium, Aline C., Krista Harper and Marty Otañez (editors)

2015 *Participatory Visual and Digital Research in Action*. Routledge, New York.

Guidotti, Riccardo, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti and Dino Pedreschi

2018 A Survey of Methods for Explaining Black Box Models, *ACM Computing Surveys* 51(5):Article 93. DOI: 10.1145/3236009

Gunkel, David J.

2012 *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*. MIT Press, Cambridge MA.

Harper, Krista and Aline Gubrium

2017 Visual and Multimodal Approaches in Anthropological Participatory Action Research, *General Anthropology* 24:1-8. DOI: 10.1111/gena.12028

Harris, Trevor M.

2012 Interfacing archaeology and the world of citizen sensors: exploring the impact of neogeography and volunteered geographic information on an authenticated archaeology. *World Archaeology* 44(4):580–591. DOI:10.1080/00438243.2012.736273.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren and Jian Sun

2016 Deep Residual Learning for Image Recognition. *29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, Las Vegas NV, pp. 770-778. Institute of Electrical and Electronics Engineers, New York NY. DOI: 10.1109/CVPR.2016.90

Hein, Irmgard, Alfonso Rojas-Domínguez, Manuel Ornelas, Giulia D'Ercole and Lisa Peloschek

2018 Automated classification of archaeological ceramic materials by means of texture measures. *Journal of Archaeological Science: Reports* 21:921–928. DOI: 10.1016/j.jasrep.2017.12.032

Huggett, Jeremy

1985 Expert systems in archaeology. In *Current Issues in Archaeological Computing*, edited by Julian D. Richards and Malcolm A. Cooper, pp. 123-142. British Archaeological Reports International Series 271, Oxford.

2000 Computers and Archaeological Culture Change. In *On the Theory and Practice of Archaeological Computing,* edited by Gary Lock and Kayt Brown, pp. 5–22. Oxford University Committee for Archaeology Monograph 51, Oxford.

2012 Lost in information? Ways of knowing and modes of representation in e-archaeology. *World Archaeology* 44(4):538-552. DOI: 10.1080/00438243.2012.736274

2017 The Apparatus of Digital Archaeology. *Internet Archaeology* 44. DOI: 10.11141/ia.44.7

2021 Algorithmic Agency and Autonomy in Archaeological Practice. *Open Archaeology* (in press).

Huggett, Jeremy and Katharine Baker

1985 The computerised archaeologist: the development of Expert Systems. *Science and Archaeology* 27:3-5.

Itkin, Barak, Lior Wolf and Nachum Dershowitz

2019 Computational Ceramicology. https://arxiv.org/abs/1911.09960v1

Kamash, Zena

2017 'Postcard to Palmyra': bringing the public into debates over post-conflict reconstruction in the Middle East. *World Archaeology* 49(5):608-622. DOI: 10.1080/00438243.2017.1406399

Khatib, Ossama, Xiyan Yeh, Gerald Brantner, Brian Soe, Boyeon Kim, Shameek Ganguly, Hannah Stuart, Shiquan Wang, Mark Cutkosky, Aaron Edsinger, Phillip Mullins, Michael Barham, Christian R. Voolstra, Khaled Nabil Salama, Michel L'Hour and Vincent Creuze

2016 Ocean One: A Robotic Avatar for Oceanic Discovery. *IEEE Robotics & Automation Magazine* 23(4):20-29. DOI: 10.1109/MRA.2016.2613281

Khunti, Roshni

2018 The Problem with Printing Palmyra: Exploring the Ethics of Using 3D Printing Technology to Reconstruct Heritage. *Studies in Digital Heritage* 2(1), 1-12. DOI: 10.14434/sdh.v2i1.24590

Kiddey, Rachael

2020 I'll Tell You What I Want, What I Really, Really Want! Open Archaeology that Is Collaborative, Participatory, Public, and Feminist. *Norwegian Archaeological Review* 53(1):23-40. DOI:10.1080/00293652.2020.1749877, accessed 24 June 2020.

Kitchin, Rob

2017 Thinking critically about and researching algorithms. *Information, Communication & Society* 20(1):14-29. DOI: 10.1080/1369118X.2016.1154087

Lagrange, Marie-Salomé and Monique Renaud

1985 Intelligent Knowledge-Based Systems in Archaeology: A Computerized Simulation of Reasoning by Means of an Expert System. *Computers and the Humanities* 19(1): 37-52. DOI: 10.1007/BF02259616

Leopold, Till Alexander, Saadia Zahidi, and Vesselina Ratcheva

2016 *The Future of Jobs: Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution*. Global Challenge Insight Report. World Economic Forum, Geneva, Switzerland.

Li, B. Y. Lu, A. Godil, T. Schreck, M. Aono, H. Johan, J. Saavedra and S. Tashiro

2013 SHREC'13 Track: Large Scale Sketch-Based 3D Shape Retrieval. In *Eurographics Workshop on 3D Object Retrieval*, edited by Umberto Castellani, Tobias Schreck, Silvia Biasotti, Ioannis Pratikakis, Afzal Godil and Remco Veltkamp, pp. 89-96. The Eurographics Association, Geneva. DOI: 10.2312/3DOR/3DOR13/089-096

Lipton, Zachary C.

2018 The Mythos of Model Interpretability. *Communications of the ACM* 61(10):36-43. DOI: 10.1145/3233231

Lustig, Caitlin and Bonnie Nardi

2015 Algorithmic Authority: The Case of Bitcoin. *48th Hawaii International Conference on System Sciences (HICSS)*, HI, USA, pp. 743-752. DOI: 10.1109/HICSS.2015.95

Makridis, Michael and Petros Daras

2012 Automatic Classification of Archaeological Pottery Sherds. *ACM Journal on Computing and Cultural Heritage* 5(4): Article 15. DOI: 10.1145/2399180.2399183

Manyika, James, Susan Lund, Michael Chui, Jacques Bughin, Jonathan Woetzel, Parul Batra, Ryan Ko and Saurabh Sanghvi

2017 *Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation*. McKinsey Global Institute. https://www.mckinsey.com/featured-insights/future-of-organizations-and-work/Jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages

Milotta, Filippo, Filippo Stanco, Davide Tanasi and Anna Gueli

2018 Munsell Color Specification using ARCA (Automatic Recognition of Color for Archaeology). *ACM Journal of Computing and Cultural Heritage* 11(4):article 17. DOI: 10.1145/3216463

Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter and Luciano Floridi

2016 The ethics of algorithms: Mapping the debate. *Big Data & Society* 3(2):1-21. DOI: 10.1177/2053951716679679

Moses, Lyria Bennett

2018 Is Your Algorithm Dangerous? *IEEE Technology and Society Magazine* 37(3):20-21. DOI: 10.1109/MTS.2018.2857618

Myres, Adrian

2010 Camp Delta, Google Earth and the ethics of remote sensing in archaeology. *World Archaeology* 42(3):455-467. DOI: 10.1080/00438243.2010.498640

Nedelkoska, Ljubica and Glenda Quintini

2018 Automation, skills use and training. *OECD Social, Employment and Migration Working Papers No. 202*. OECD Publishing, Paris. DOI: 10.1787/2e2f4eea-en

Olsen, Bjørnar, Michael Shanks, Timothy Webmoor and Christopher Witmore

2012 *Archaeology: The Discipline of Things*. University of California Press, Berkley CA.

O'Neil, Cathy

2016 *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Allen Lane, London.

Opitz, Rachel and Jason Herrmann

2018 Recent Trends and Long-Standing Problems in Archaeological Remote Sensing. *Journal of Computer Applications in Archaeology* 1(1):19-41. DOI: 10.5334/jcaa.11

Pan, Sinno Jialin and Qiang Yang

2010 A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* 22(10):1345-1359. DOI: 10.1109/TKDE.2009.191

Pasquale, Frank

2015 *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, Cambridge MA.

Perry, Sara. and Nicole Beale

2015 The Social Web and Archaeology's Restructuring: Impact, Exploitation, Disciplinary Change. *Open Archaeology* 1(1):153-165. DOI: 10.1515/opar-2015-0009

Pfleging, Johannes, Marius Stücheli, Radu Iovita and Jonas Buchli

2015 Dynamic Monitoring Reveals Motor Task Characteristics in Prehistoric Technical Gestures. *PLoS ONE* 10(8):e0134570. DOI: 10.1371/journal.pone.0134570

Poller, Tessa

2018 The craft of earthwork survey. In *Re-Mapping Archaeology: Critical Perspectives, Alternative Mappings,* edited by Mark Gillings, Piraye Hacıgüzeller and Gary Lock, pp. 117-140. Routledge, Abingdon.

Prentiss, Anna Marie (editor)

2016 Drones in Archaeology. *The SAA Archaeological Record* 14(2).

Qi, Charles, Hao Su, Kaichun Mo and Leonidas Guibas

2017 PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *30th IEEE Conference on Computer Vision and Pattern Recognition CVPR 2017, Honolulu, HI*, pp. 77-85. Institute of Electrical and Electronics Engineers, New York NY. DOI: 10.1109/CVPR.2017.16

Rasheed, Nada and Md Jan Nordin

2020 Classification and reconstruction algorithms for the archaeological fragments. *Journal of King Saud University--Computer and Information Sciences* 32(8):883-894. DOI: 10.1016/j.jksuci.2018.09.019

Reinhard, Andrew

2018 *Archaeogaming: An Introduction to the Archaeology in and of Videogames*. Berghahn, Oxford.

Ribani, Ricardo and Mauricio Marengoni

2019 A Survey of Transfer Learning for Convolutional Neural Networks. 32nd Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), Rio de Janeiro, pp. 47-57. Institute of Electrical and Electronics Engineers, New York NY. DOI: 10.1109/SIBGRAPI-T.2019.00010

Richardson, Lorna-Jane

2018 Ethical Challenges in Digital Public Archaeology. *Journal of Computer Applications in Archaeology* 1(1):64-73. DOI: 10.5334/jcaa.13

Ridge, Mia

2013 From Tagging to Theorizing: Deepening Engagement with Cultural Heritage through Crowdsourcing. *Curator: The Museum Journal* 56(4):435-450. DOI: 10.1111/cura.12046

Roman-Rangel, Edgar, Diego Jimenez-Badillo and Stephane Marchand-Maillet

2016 Classification and retrieval of archaeological potsherds using histograms of spherical orientations. *ACM Journal on Computing and Cultural Heritage* 9(3): Article 17. DOI: 10.1145/2948069

Schwab, Klaus

2016 *The Fourth Industrial Revolution*. World Economic Forum, Geneva.

Seaver, Nick

2018 What Should An Anthropology of Algorithms Do? *Cultural Anthropology* 33(3):375-385. DOI: 10.14506/ca33.3.04

Shanks, Michael and Randall H. McGuire

1996 The Craft of Archaeology. *American Antiquity* 61(1):75–88.

Sheehan, Kim Bartel

2018 Crowdsourcing research: Data collection with Amazon's Mechanical Turk. *Communication Monographs* 85(1):140-156. DOI: 10.1080/03637751.2017.1342043

Shirky, Clay

2009 A speculative post on the idea of algorithmic authority. http://www.shirky.com/weblog/2009/11/a-speculative-post-on-the-idea-of-algorithmic-authority/

Smith, Gavin

2018 Data doxa: The affective consequences of data practices. *Big Data & Society* 5(1):1-15. DOI: 10.1177/2053951717751551

Steiner, Christopher

2013 *Automate This: How Algorithms Took Over Our Markets, Our Jobs, and the World*. Portfolio/Penguin, New York.

Stobiecka, Monika

2020 Archaeological heritage in the age of digital colonialism. *Archaeological Dialogues* 27(2):113–125. DOI:10.1017/S1380203820000239.

Strauß, Stefan

2018 From Big Data to Deep Learning: a Leap Towards Strong AI or 'Intelligentia Obscura'? *Big Data and Cognitive Computing* 2(3):16. DOI: 10.3390/bdcc2030016

Susskind, Richard and Daniel Susskind

2015 *The Future of the Professions: How Technology will Transform the Work of Human Experts*. Oxford University Press, Oxford.

Teddy, Debroutelle, Janvier Romain, Chetouani Aladine, Treuillet Sylvie, Exbrayat Matthieu, Martin Lionel and Jesset Sebastien

2015 Automatic pattern recognition on archaeological ceramic by 2D and 3D image analysis: A feasibility study. In *Image Processing, Theory, Tools and Applications. International Conference, IPTA 2015, Orléans,* edited by Rachid Jennane, pp. 224-228. Institute of Electrical and Electronics Engineers, New York NY. DOI: 10.1109/IPTA.2015.7367133

Trussell, H. Joel

2018 Why a Special Issue on Machine Ethics. *Proceedings of the IEEE* 106(10):1774-1776. DOI: 10.1109/JPROC.2018.2868336

Tyukin, Ivan, Konstantin Sofeikov, Jeremy Levesley, Alexander Gorban, Penelope Allison and Nicholas Cooper

2018 Exploring Automated Pottery Identification [Arch-I-Scan]. *Internet Archaeology* 50. DOI: 10.11141/ia.50.11

Ulguim, Priscilla

2018 Models and Metadata: The Ethics of Sharing Bioarchaeological 3D Models Online. *Archaeologies: Journal of the World Archaeological Congress* 14(2):189–228. DOI:10.1007/s11759-018-9346-x.

Wallach, Wendall and Colin Allen

2009 *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press, Oxford.

Walsh, Toby

2018 Expert and Non-expert Opinion About Technological Unemployment. *International Journal of Automation and Computing* 15(5):637-642. DOI: 10.1007/s11633-018-1127-x

Wilson, Andrew T. and Ben Edwards (editors)

2015 *Open Source Archaeology: Ethics and Practice*. De Gruyter Open, Berlin. DOI: 10.1515/9783110440171

Wright, Holly and Gabriele Gattiglia

2018 ArchAIDE: Archaeological Automatic Interpretation and Documentation of cEramics. In *Cultural Informatics 2018: Proceedings of the Workshop on Cultural Informatics co-located with the EUROMED International Conference on Digital Heritage 2018 (EUROMED 2018), Nicosia, Cyprus*, edited by Angeliki Antoniou and Manolis Wallace, pp. 60-65. urn:nbn:de:0074-2235-7. http://ceur-ws.org/Vol-2235/

Wu, Zifeng, Chunhua Shen and Anton van den Hengel

2019 Wider or Deeper: Revisiting the ResNet Model for Visual Recognition. *Pattern Recognition* 90:119-133. DOI: 10.1016/j.patcog.2019.01.006

Wurzer, Gabriel, Kerstin Kowarik and Hans Reschreiter (editors)

2015 *Agent-based Modeling and Simulation in Archaeology*. Springer, Cham. DOI: 10.1007/978-3-319-00008-4

Yarrow, Thomas

2008 In Context: Meaning, Materiality and Agency in the Process of Archaeological Recording. In *Material Agency: Towards a Non-Anthropocentric Approach,* edited by Carl Knappett and Lambros Malafouris, pp. 121-137. Springer, Boston MA. DOI: 10.1007/978-0-387-74711-8_7

Zerilli, John, Alistair Knott, James Maclaurin and Colin Gavaghan

2019 Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard? *Philosophy & Technology* 32(4):661-683. DOI: 10.1007/s13347-018-0330-6