



Archaeology, Typology and Machine Epistemology

Gavin Lucas

Abstract

In this paper, I will explore some of the implications of machine learning for archaeological method and theory. Against a back-drop of the rise of Big Data and the Third Science Revolution, what lessons can be drawn from the use of new digital technologies and computational approaches as they are applied to archaeological typologies? In this paper, I explore two key aspects of these approaches - automation and epistemic novelty – and attempt to unravel their implications for archaeological practice. Furthermore, the paper will situate these topics within developments of the philosophy of science and technology and suggest an alternative way to think about machine learning that draws on re-thinking what we mean by machines and automation.

Introduction

Typological debates have periodically surfaced in the history of archaeology but in the Anglophone world the height of such discussion occurred in the middle of the twentieth century. Although it would be useful to situate these debates within what was happening elsewhere in Europe, given the limitations of my knowledge and experience, I will unfortunately confine myself to the anglophone sphere. It happened at quite a critical period, but one that

straddled the traditional divide between culture historical and processual archaeology, i.e. between the 1940s and 1970s. The various controversies and discussions around typology during this period, especially in North America, have been written about extensively (e.g. see Wylie 2002) so I will not discuss them in any depth. But in broad terms, they did initiate a new set of approaches to both typological method and the meaning of the type concept (i.e. what does a 'type' refer to?). The link between method and meaning was also significant as it was felt, especially in North America, that the earlier reigning taxonomic approach was wedded to a simplistic analogy of archaeological typology to biological classifications (eg. Hargrave 1932). Such analogies of course also have an old pedigree in European archaeology (e.g. Åberg 1929, Gorodozov 1933; also see Reide 2006). But what the new methodological approaches did, especially in the work of Spaulding, was to combine statistical methods of classification with a stress on making more explicit the cultural or social meanings of such classifications (Spaulding 1953).

It is perhaps not surprising that the meaning of the type-concept should have come under scrutiny at this time; the emergence of a more functionalist approach to the interpretation of the past in the 1940s which eventually morphed into New Archaeology and processualism in the 1960s and 1970s was in large part, a reaction to a perceived poverty of interpretation in North American archaeology (Wylie 2002). Retaining a concept like the 'type' in a context which was explicit in rejecting a normative view of culture, was to some extent problematic as manifest in the Ford-Spaulding debate which hinged on whether types were simply heuristic constructs of the archaeologist or whether they represented real, empirical distinctions in the material itself (Ford 1954; Spaulding 1953). To an extent, this boiled down to the emic-etic distinction, where emic categories are those used by the past society being studied while etic categories are those used by the archaeologist. In other words, is a type such as 'beaker pottery', purely a concept used by archaeologists (etic) or did the idea of beaker pottery have any meaningful significance for people in the past (emic)?

Nonetheless, it should be stressed that the meaning of a type is not exhausted by this emic-etic opposition. Artefact types may correspond to empirical differences which relate to social or cultural processes in the past without these having been mental categories held by people in the past. In other words, types can be both heuristic constructs *and* socially meaningful. Indeed, as Marie-Louise Sørensen has pointed out, the very fact that objects display a

repetition of form or appearance that enables us to construct a typology in the first place is, in itself, of significance regarding past practices of production and standardization (Sørensen 1997; 2014).

But what have all these debates – now nearly a century old – to do with current computational approaches to typology? I want to suggest that the trouble that the type-concept gave to a processual archaeology that was moving away from normative views of culture in the middle of the twentieth century is re-surfacing today in the wake of new materialisms, especially those which stress a relational ontology. In a theoretical context where the stress is on becoming over being, on change as a continual process rather than a punctuated sequence of static entities (Crellin 2020; Gosden & Malafouris 2015), the type-concept becomes again, a cause for concern, albeit framed in quite different ways. Indeed, typological ways of thinking seem to lie at the very heart of the way we construct block-time chronologies (i.e. periods) and explains why, despite the fact that we were supposed to have abandoned culture historical archaeology nearly a century ago now, its legacy remains as strong as ever in the various period divisions (e.g. Three Age System) and typologies that remain routinely in use by the discipline (Griffiths 2017). At the same time, new digital and computational technologies have started to be applied to typological studies in archaeology – but in stark contrast to the typological debates of the 1950s to 1970s where the use of statistics and computers was directly connected to the theoretical ‘revolution’ of processualism, today these two aspects appear completely divorced. As far as I am aware, there has been no dialogue between the new materialisms and big data/artificial intelligence in the context of typology. Indeed the latter seem more geared towards the theoretical status quo - if not even aimed at fulfilling the promises of the 1970s.

Given this - and in tune with the general goals of this volume - perhaps it is time to ask some important questions again about typologies and the type-concept. But we cannot resurrect the same questions that were raised in the 1950s which revolved around an ambivalence concerning what types mean. Indeed, in the wake of non-representational or more-than-representational epistemologies, our question today should be what do – or can - types *do*, not what they represent or mean. In light of this general re-orientation, I want to examine more specifically on the way contemporary approaches to typology work, specifically those connected to a machine epistemology – i.e. typologies drawing on the use of artificial intelligence, machine learning, big data and digital technologies. What are the issues raised by their mode of operation?

In this paper, I will focus on two aspects that, to me at least, seem salient in this respect: automation and novelty. Automation is all about the delegation of work, normally done by people, to machines; as Hörr et al. (2014) point out, at a time when archaeology is facing ever greater strains on resources and personnel, automation is an obvious solution, especially for time-consuming tasks like documentation and classification. While automation has a more pragmatic bearing, in contrast novelty is about the ability of computers and digital technology to draw something out of our data that we could not have achieved otherwise; in the context of typology, an example would be how the use of artificial intelligence is improving upon older classifications (Hörr et al. 2014). In the rest of this chapter, I want to explore the theoretical implications of these two issues of automation and novelty and conclude by reflecting on how these issues intersect with recent archaeological theory as well as more generally raising questions over our understanding of computation within a broader ontological context. What conception of reality is implied in invoking the language of artificial intelligence and machine epistemology?

Automation

Automation is perhaps most famously entangled with the issue of time-saving; whether it is saving time for human labour which can be directed elsewhere or actually speeding-up a process which would otherwise have taken much longer. It is important to distinguish these two dimensions of time-saving though and their association with the Marxist labour theory of value. In the wider context of industrial capitalism, Marx's famous critique of automation in a capitalist society is based on the argument that it fed off the de-skilling of the workforce, turning humans into cogs in the machine (Marx 1976, ch. 15; also see MacKenzie 1984). The factory scene in Chaplin's famous film *Modern Times* (1936) has a worker desperately trying to keep up with the pace of the conveyor belt; this is automation as time-saving in the second sense (speeding-up) but not the first sense (freedom to do something else). Marx's vision of a socialist society was of course, one where the second sense was enabled by the first. This is also the assumption behind arguments for using automation in archaeology as well, although as should be clear from Marx's critique, one does not necessarily entail the other. In the context of machine epistemology, these issues have recently been articulated around the notion of slow science (Stengers 2017). I will take my cue from Caraher (2019), Marila

(2019) and Mol (2021), all of whom discuss this issue in terms of a slow archaeology and argue that acceleration of the archaeological process through automation can be detrimental to our understanding of the past. That the impetus for fast or speedy science as epitomized in the new computational and digital technologies creates 'Chaplin's factory effect' whereby we don't have time to properly reflect on what we are doing because we are too busy serving the demands of the technology. Marila especially defines the streamlining effects of 'fast science', where issues of standardization dominate, leaving little room for creative or alternative approaches. Related critiques have also been voiced by Ribeiro who argues for the importance of qualitative case studies to balance the increasing domination of quantitative research (Ribeiro 2019).

How plausible are such critiques? Taking your time might sound fine in an ideal world, but in the high-stakes of research grants and assessments or development-led archaeology, who can really afford to do a slow archaeology as advocated by Caraher, Marila and Mol? In response, one might counter that that is precisely the problem; that slow archaeology or indeed slow science is not a personal choice but a structural issue which cuts to the very heart of how knowledge production today is tied to the capitalist model of automation. In general, my sympathies lie more with the position of Caraher, Marila and Mol, and although the structural issues are important, they lie outside the scope of this chapter. Indeed they are bigger than archaeology as the concern some very broad issues about the nature of research and education in contemporary society (e.g. Sharma 2014; Stengers 2017). Instead, I will focus on the more specific issue of the relation between automation as time-saving vs. automation as labour-saving. Because most critiques essentially focus only on the former. In this light, countering automation with a slow science is not really the point; rather it is how we can use the acceleration of some aspects of archaeology facilitated by automation to put our limited time to other purposes. In other words, the issue of automation is not an either/or and by implication, neither are we faced with having to choose between a fast or slow archaeology; both are possible at the same time. The issue is not choosing between one or the other, but how best to partition them.

With that, let me now focus on the specific topic addressed in this volume: typology. In this context, when is automation and speed a good thing? My sense is, automation works best when it is applied to reproducing an accepted way of doing something; that is, to accelerate a pre-determined output. In terms of typology, it is all about using artificial intelligence and other

technologies to speed up the documentation or recording process; that is, to sort things into already accepted types rather than create a typology from these things (the latter, I will discuss in the next section). Projects like ArchAIDE or Arci-I-Scan which, although still in the trial stage, use AI to help with the identification of ceramics in part, by-pass the need for specialists (Meghini et al. 2017; Gualandi et al. 2021; Tyukin et al. 2018; van Helden et al. 2022). Although it does not make identification any faster than a specialist would, by turning anyone who uses the app into a specialist, it short-circuits the need for years of training and in that sense, is time-saving at a collective/disciplinary level. Moreover, in some cases, identification may even improve upon human eye, as a recent trial study on ceramics from the Southwest of the USA has suggested (Pawlowicz & Downum 2021). Indeed, the potential of using big data and machine learning to assist with these kinds of processes is valuable and its importance should not be under-estimated. Automation, when used in what is called ‘supervised learning’, where it is primarily directed to assign artefacts to a pre-determined output such as an accepted typology, is surely advantageous. If it can even speed up a process beyond what a specialist is capable of, then it should be embraced. If hundreds of thousands of pottery sherds or animal bones can be rapidly scanned and identified using automated technologies instead of a person sifting through these items at table, and performed at a rate much faster, is this not beneficial? Just as electronic and digital surveying instruments replaced much of the work done by surveyors reliant on theodolites, tapes or plane tables, can we imagine a future where AI can do the same for our pottery specialist or zooarchaeologist? In neither case is this about technology replacing the specialist; we still need surveyors just as we will always need zooarchaeologists; the question is not one of technologies replacing these people but simply taking over some of the more routine and replicable tasks, freeing them up to spend more time on analysis and interpretation.

At the same time, we should not be so naïve as to think any technology is neutral; in automating some aspects of archaeological work, especially those that operate at the coalface – that is, those in direct contact with the stuff of archaeology (e.g. layers, structures, finds), there is the danger that we lose something in the process. This has been most cogently argued in the context of digital recording practices on excavation, specifically the contrast between planning features using digital instruments versus traditional pencil, paper and tapes (see Morgan & Wright 2018; Morgan et al. 2021). It has been shown that critical cognitive work is sacrificed when recording goes digital – a loss of attentiveness to the nature of features in the ground that manual drawing

enhances but that automated planning obscures. It is very simple to extend these debates to the use of AI in identifying artefact types. At the same time, we need to come back to the point that this does not have to be an either/or scenario; using digital surveying does not have to replace hand-drawn plans, but rather it should free up time to explore new possibilities for different forms of manual recording which retain its benefits. The same should apply to the use of AI as an aid to artefact identification.

To recap: the argument proposed here is that automation works best when it is made to work on the more routine tasks, one where the output is pre-determined; where supervised learning is made to work with pre-existing typologies but where human interaction with artefacts is retained but directed in new ways. The implication here, is that automation is restricted to reproducing knowledge but not really creating new knowledge. This remains the preserve of the human archaeologist. But as we know, many archaeologists have advocated for a more active role in automation, one where AI is used to create new knowledge through semi-supervised and unsupervised learning which leads me to the second issue.

Novelty

One of the larger claims made for big data and the use of AI is its potential to say something new, not simply automate and speed-up a pre-existing process. In this context, novelty means new *kinds* of data rather than simply more of the same. Any new excavation will produce new finds, but if these finds are classified using conventional typologies, then all we have are more examples of the same type – not new types or even a new typology. Of course, even conventional types can be used to generate new knowledge if they elucidate new patterns or sets of relations within a wider context (e.g. distributions of pottery type X), but this only serves to further clarify what is implied by novelty. New kinds of data means new data patterns or relations which in the context of typology, means new types or classification systems. In the world of computation, such novelty is linked to a process called data mining.

The concept of data mining essentially refers to the task of pattern recognition in large quantities of data, a task which is either fully or semi-automated. It is considered a key step in KDD (knowledge discovery in databases) or KDP

(knowledge discovery process). In the field of archaeological typology, this is about using computers and digital technology to assist in the *construction* of a classification, not simply as an aid to sorting out objects into pre-existing types. As a result, the method of working is very different; the latter is characterized largely by supervised learning but with data mining, this is the last or final stage of the process which begins with unsupervised and semi-supervised learning (e.g. see Hörr et al. 2014 for an example). With supervised learning, the classification is already in place and the main job is simply to assign or sort objects out into their appropriate type; in unsupervised learning, we have to begin constructing the types from a mass of data – such as a pottery assemblage. Typically, this involves searching for similarities between objects based on a set of features and usually requires several runs or iterations to identify relevant, yet provisional ‘types’ which are then honed in a second stage of semi-supervised learning. Finally, once the types have been established they are tested in a third and final stage of supervised learning.

One should make clear at this stage that the primary step of unsupervised learning is not free of human involvement, despite how it sounds. The ‘mass of data’ subject to data mining has still been pre-processed to some extent, both in terms of defining the population parameters (i.e. assemblage of ceramics) and the coding of the objects (i.e. how the ceramics have been recorded and entered into the database). Moreover, even during the initial data analysis, the search for similarity between objects relies heavily on human input to make decisions on relevance; this is made very explicit by Hörr et al. in their analysis of late Bronze Age pottery where they reduce the number of relevant features to construct their typology down to two – size and morphology (Hörr et al. 2014). In other words, the ‘unsupervised’ creation of types is based on an already highly structured set of data. This is not to suggest there is no difference between contemporary computational approaches to classification and more traditional ones. Clearly the role of computers and machine learning does play an important part in this process, but my point is rather to underline the fact that it is not as automated or unsupervised as the terminology would imply. Indeed, it is better to see this perhaps as more of a continuum than an either or – that is, the degree to which humans are involved.

If one imagines a traditional, intuitive approach to classification, it may involve nothing more than the lone scholar sitting in front of the collection of pottery and manually creating types based on visual inspection. In some of the older typologies, types may not even be explicitly defined – simply named with an image to show what the type looks like (e.g. amphora); in the better ones,

specific criteria might be given (e.g. long neck, biconical body, lug handles). With the advent of numerical and paradigmatic approaches to classification in the 1950s and 1960s, these criteria or features became subject to more explicit, statistical operations such as attribute and matrix analysis (Spaulding 1953; Clarke 1962), and later cluster analysis performed with the aid of computers (Doran & Hodson 1975; Whallon & Brown 1982). But none of this work was really automated with the exception of computer-aided cluster analysis and then, only minimally so. What the contemporary computational approaches do is increase the degree of automation; the most common way has been to use digital profiling, i.e. use of algorithms to determine the shape or curvature of ceramic vessels from their profiles or to apply convolutional neural networks (CNN) for identifying decorative motifs or patterns on the surface of pottery (e.g. Meghini et al. 2017; Gualandi et al. 2021; Pawlowicz & Downum 2021). For example, Gilboa et al. used digital profiling to differentiate between two types of Iron Age torpedo-shaped storage jars found at sites in the Levant (Gilboa et al. 2004), while Pawlowicz & Downum (2021) used CNN to discriminate between different decorative motifs on sherds from sites in North America which have important chronological implications. Moreover, the promise of a much broader use of automation in building typologies is outlined in the study by Hörr et al., already cited. Yet as we have already made clear, even the most contemporary computational processes are not fully automated, i.e. involve no human input. The difference is always a matter of degree, rather than of kind.

Given that, how do we characterize the novelty that data mining brings to typological work and especially, how does it differ from more traditional means of building typologies? Here, we really need to be very attentive to the practices involved, in particular what, in practical terms, is involved in typological work. Consider a traditional approach of intuitive classification or even manual sorting of objects into pre-given types; in this situation, very few instruments or tools intervene between the archaeologist and the resultant work. When sorting an assemblage of pottery into types, the sheer act of moving sherds around the table into separate piles acts to concretize the types. The only mediation that might be present is if a reference collection or publication is used as an aid in the sorting process. In terms of creating a typology however, the situation starts to change; while the mediation might take the form of measuring instruments (e.g. calipers, diameter charts), the really crucial stage happens when the physical sherds are converted into drawings, numerical or text string data because thereafter, it is upon these that the typology is constructed, not the sherds themselves. Similarly with

digital profiling or CNN which works off scans of drawn profiles or digital photographs of sherds, it is the digital image that now substitutes for the physical object. Once the pottery has been converted into a digital medium, then a very different set of potentials are opened up, whether this is explored manually or through AI.

In a sense, the kind of novelty or new knowledge that emerges all depends on these processes of translation and the technologies and instruments used to facilitate this. In other words, what matters is the array of potential or virtual possibilities opened up by working on different translations of the object. Data mining scans of pottery profiles offers up different possibilities to digital photographs which are different again to 3D laser scans. Although I do not want to downplay the capacity of data mining to capture more complex variability through its access to larger datasets, I would like to stress that much of the potential for new knowledge is already prefigured at the moment of data capture – especially what technologies, instruments and codes are employed in this phase. In summary, I would argue that the *parameters* of epistemic novelty are set by the way in which data are initially constructed as much, if not more than how they are subsequently analysed – whether by algorithms or humans.

Implications

What broader implications can one draw from these observations on automation and novelty in the context of computation? At this point, I would like to situate these developments within discussions on the rise of big data and machine learning insofar as they connect with current theory, especially what is broadly defined as new materialism (e.g. Witmore 2014). There are potentially two ways to explore this question: one from a purely epistemological angle, the other more ontological. However, separating these two aspects is somewhat problematic as we shall see but to begin, I will assume this distinction and start with the epistemological angle.

There have been two main views on the epistemological implications of the machine learning and big data revolution (Kitchin 2014). The first is a, presumably, deliberately provocative position which proclaims the end of theory, as outlined in Anderson's oft-cited online paper. Anderson argues that

“with enough data, the numbers speak for themselves” (Anderson 2008). For Anderson, the data deluge means that no theory or model can possibly account for the complexity of available data; rather we need to draw on statistical algorithms to look for patterns and correlations. Connected to this is the idea of a machine epistemology – the idea that knowledge will emerge through a kind of super-induction that only modern computing makes possible (Wheeler 2017). This kind of ‘empiricism reborn’ has been heavily criticized for making all the same basic errors of any naïve inductivism; the theoretically laden nature of any data, the ontologies implicit in coding and classification and the myth of raw data all underline how even big data is still theoretical (Leonelli 2016; Bowker & Starr 1999; Gitelman 2013).

The other perspective on big data acknowledges these well-known issues and rather than advocating any end of theory, rather suggests that the asymmetry of the usual relationship between theory and data be reversed. Instead of questions and problems being framed by theory, and where data is used to test theories, data-driven science argues that the questions and problems emerge from the data themselves, with theory being brought in afterward. Such an approach has been aligned more generally with the Peircean notion of abduction (Kitchin 2014). On the whole, discussion of big data in archaeology has tended to follow the second view, though not always explicitly so (e.g. see Cooper & Green 2016; Wesson & Cottier 2014). More explicit calls for a data-driven archaeology have been made by Gattiglia who suggests that “...from a theoretical point of view, archaeological theory should shift towards data-driven research and a Big Data approach.” (Gattiglia 2015: 118; also see Kristiansen 2014.).

But the views of how science works in both of these positions seems to be stuck in some time warp of the mid twentieth century, opposing a positivist model of science that expired long ago. Since the practice turn of science studies (Solar et al. 2014), it seems very naïve to separate this kind of purely epistemological view of science from ontology – specially, from the ontological conditions through which scientific knowledge is produced. This is not about acknowledging the social context of science as was common early on (e.g. Bloor 1976; Fuller 1988) but rather more broadly about the *materiality* of science in terms of the bodies, instruments and spaces implicated in its practice (Latour 1987; Latour & Woolgar 1979; Barad 2007). One cannot view science as if it was some disembodied activity and nor can one fully understand issues of knowledge production, of epistemology, without including the material conditions in which such knowledge subsists. As most clearly

articulated by Latour (Latour 1987), this was also about recognizing the fact the knowledge production is not simply the preserve of humans but was distributed between a network of agents which in archaeology, might include artefacts, calipers, computers and reference books as well as the human archaeologist (Lucas 2012). More generally, it unsettles the ontological boundaries between the human scientist and the world of instruments and other things. As such, it should also make us re-consider how we view the two themes discussed in this paper: automation and novelty. What does automation or epistemic novelty mean if the boundary between the scientist and their instruments is blurred? In the rest of this section, I want to focus primarily on the former issue of automation; for the latter, see Lucas (forthcoming). In particular, I want to explore the connection between automation and our conception of what a machine is.

Our common understanding of a machine usually implies two things: one, that it is made of moving parts driven by an external energy supply and which function together to perform a certain task; second, that such assemblages are artificial. The first aspect is a way of distinguishing machines from tools – thus a chisel is a tool, but a lathe is a machine. Note that a computer here, although a machine, does not seemingly have obvious moving parts (except perhaps the fan) although clearly its operation relies on the controlled movement of electrical currents. In a way, a computer is a machine that collapses the distinction between the external energy supply and its moving parts. But this first aspect is not what I want to focus on, but rather the second: the artificiality of the machine. Machines are generally taken to be products of human manufacture, not natural growth.

This second aspect is in many ways, more interesting because it marks an ontological rift between nature and culture; a machine that needed no human involvement, would not be a machine but a natural entity, i.e an organism. At the same time, the image of a machine that becomes fully independent of humans – fully automated and autonomous – questions this ontological division. Such a machine, although a central motif in much science fiction (e.g. androids) has long been a source of philosophical reflection. Even simple or quasi- automata like clocks captured the minds of European thinkers in the 17th century when they used the machine analogy to talk about living organisms (e.g. the heart as a pump). Yet it was only in the twentieth century that the concept of the machine started to be radically re-thought.

In this context, it is both instructive and ironic to consider the etymology of the

term 'robot'. It derives from the play *R.U.R. (Rossum's Universal Robots)* by the Czech writer Karel Čapek, which was translated into English in 1923 and in which robots were organically engineered humanoids, essentially humans manufactured for one purpose: work (Čapek 2011; in Czech, the word 'robota' means drudgery or servitude and a Robotnik is a menial labourer). This brings us right back to the discussion of automation as labour-saving and the Marxist critique, where robots or machines are essentially devices for the delegation of work or labour, and although today we tend to think of them as non-human, in Čapek's play, a robot is rather a being which renders this boundary much more fluid.

Challenging the boundary between human and machine also received more scholarly attention though. Lewis Mumford, coined the term megamachine to refer to larger, self-organizing systems, specifically states and other authoritarian regimes where humans became cogs in a social machine (Mumford 1967). The machine as in inherent form of human collectives long before its rendition in artificial, mechanical devices. In a different way, we can also invoke Georges Canguilhem's inversion of the 'body as machine' metaphor where the machinic becomes an extension of the organic (Canguilhem 1992). However, where Mumford and Canguilhem both turned the enlightenment metaphor of organism as machine on its head by seeing the machine as inherent in or extended from the organic, with Gilbert Simondon – who was a student of Canguilhem - the very boundary between the organic and mechanical, natural and artificial was challenged (Simondon 2016) in a way that anticipated later writings, especially those of Donna Haraway on the cyborg (1991). Simondon's work on technical objects is important because it homes in on the key issue of autonomy and its relation to individuation as broader ontological process (also see Simondon 1992).

For Simondon, distinctions between nature and culture, organic and mechanical are overwritten by a concern for how dependent objects or beings are on their context or environment to work. He suggested a distinction between three levels of autonomy: elements, individuals and ensembles, where elements are the most dependent and ensembles, the least. For example, consider a computer. Its screen is an example of an element – on its own, it really does nothing, to work it needs to be connected to a power supply and the CPU etc. The computer itself might be considered an individual – it can function almost autonomously, but it is still dependent on an external power supply (at least long-term), a human operator and for some purposes, an internet connection. An ensemble would then be the computer, power

supply, human user and anything else involved in the activity of working at the computer. An ensemble is the collective which minimizes the degree of external dependence. Of course no ensemble is completely autonomous; even in our office, we are still dependent on the power grid working and remote servers functioning, at least for most long-term, practical purposes today.

So what has all this got to do with machine learning and archaeological typology? Defined in this way, the machine is no longer linked to an anthropocentric discourse which obsesses about the divide between human and artificial intelligence and the myth of automation. Rather it is about processes of individuation and autonomy which apply to the natural world as much as the cultural, to human society as much as technology. Such levelling of distinctions became a key part of Deleuze's work, especially with Guattari and their concept of machinic assemblages (Deleuze & Guattari 1987). Similarly, machine epistemology ought to be seen not as about how computers and software generates knowledge independent of human cognition but how they operate within a broader technical ensemble which includes both humans and computers. To acknowledge the fact that knowledge production has always been situated within, and distributed across a hybrid assemblage of actors, individuals and elements. The connections here to ideas of posthumanism should be obvious (e.g. see Parisi 2017) but more important is the fact that we should be really questioning, if not abandoning the whole terminology that infects our discussion of these issues in archaeology. Terminology like artificial intelligence, machine learning, and so on, which while on the surface seem to blur the divides between the natural and artificial, human and machine, only really serves to perpetuate these divisions.

Concluding Remarks

In ending, I would like to bring my discussion back to the other concept at the heart of this volume: the type concept and its relation to new materialism. Once again, Simondon offers a useful way into re-thinking this concept. His distinction between elements, individuals and ensembles can in many ways be mapped onto the archaeological distinction between attributes, artefacts and assemblages. In relation to typology and classification, these distinctions take on relevance especially in the context of typological debates which revolved around whether types were best viewed as clusters of attributes or as prototypical objects (e.g. see Whallon & Brown 1982). Of course in the end,

we do both; often attribute clustering becomes the basis of creating prototypes (e.g. see Hörr et al. 2014) and in psychological terms, most of us operate under a prototypical conception of types (Bowker & Starr 1999). In addition, added to this is that attribute clustering and identification of prototypical objects will also be circumscribed by the nature of the assemblage: one of the major criticisms against culture historical archaeology and its concept of the type is that it regarded the assemblage in very broad terms – a whole region and period. One of Spaulding's key points in his critique of the taxonomic approach in US archaeology was that the assemblage always ought to be minimally defined (i.e. by the site) and any extensions of a typology beyond the site level have to be worked out on a site by site basis.

I think we can draw two important conclusions from these earlier debates. One is the relation between attribute and object in the construction of types needs to be seen as a fluid one; that for types to be distinguishable at all, is not something we can assume, but has to be explained. Here, I want to come back to Marie-Louise Sorensen's point: the very fact that objects display a repetition of form or appearance that enables us to construct a typology in the first place is, in itself, of significance regarding past practices of production and standardization (Sørensen 1997). The iteration of attribute clustering and the emergence of prototypical objects is something we should always acknowledge is the result of specific processes and as we all know, artefacts can vary quite substantially in the degree of standardization and thus amenability to typological analysis. What we need to be more attentive to is the degree of standardization and the elasticity of 'types'. The second conclusion is that the relation between objects and ensembles also needs to be seen as a fluid one; that even where extensive iteration and standardization is present, how widely distributed is it? Is it specific to a site, a region, a time period?

There is perhaps nothing that new about these observations, yet at the same time they are often forgotten in the quest to construct or use typologies. Yet foregrounding these points is one of the more obvious ways we can connect typology to current theoretical concerns with relational ontologies and fluid entities. Although concepts like the 'type' would seem to be antithetical to these theoretical approaches, I would argue that this is only the case if we operate under a very simple and essentialist conception of what a type is. If we recognize that types themselves can be sharp or fuzzy, can be local or widespread, and that what is important is understanding the conditions which determine these characteristics, then the type-concept still has great utility for

archaeology. Indeed, the perception and use of archaeological types is much bound up with Simondon's concept of individuation as is the notion of digital technology

Acknowledgements

I would like to thank first of all, the editors Shumon Hussain, Sébastien Plutniak and Felix Reide for inviting me to contribute to this volume and in so doing, get me to think deeper into these issues. More particularly however, I owe a great debt to Rachel Crellin and another, anonymous, reviewer who both offered extremely important feedback and comments on the paper, improving it immeasurably. To my anonymous colleague, I am particularly grateful for the reference to Karel Čapek's concept of robots of which I was previously unaware. Needless to say, all flaws remain my own responsibility.

References

Åberg, N. 1929. Typologie. In *Reallexikon der Vorgeschichte. Band 13*, edited by M. Ebert, 508–516. Berlin: Verlag Walter de Gruyter & Co.

Anderson, C. 2008. "The end of theory: The data deluge makes the scientific method obsolete". *Wired*, 23 June 2008. http://www.wired.com/science/discoveries/magazine/16-07/pb_theory

Barad, K. 2007. *Meeting the Universe Halfway. Quantum Physics and the Entanglement of Matter and Meaning*. Durham, NC: Duke University Press.

Bloor, D., 1976. *Knowledge and Social Imagery*. Chicago: University of Chicago Press.

Bowker, G. and S. Leigh Star. 1999. *Sorting Things Out: Classification and its Consequences*. Cambridge: MIT Press.

Canguilhem, G. 1992. Machine and Organism. In *Zone 6: Incorporations*, edited by J. Crary & S. Kwinter. 45-69. New York: Zone Books.

- Čapek, K. 2011. *RUR and War with the Newts*. London: Orion Books.
- Caraher, W. 2019. "Slow Archaeology, Punk Archaeology, and the 'Archaeology of Care'". *European Journal of Archaeology* 22(3): 372-385.
- Clarke, D.L., 1962. "Matrix analysis and archaeology with particular reference to British Beaker Pottery". *Proceedings of the Prehistorical Society* 28: 371-382.
- Cooper, A. & C. Green. 2016. "Embracing the Complexities of 'Big Data' in Archaeology: the Case of the English Landscape and Identities Project". *Journal of Archaeological Method and Theory* 23: 271-304.
- Crellin, R. 2020. *Change and Archaeology*. London: Routledge.
- Deleuze, G. & F. Guattari 1987. *A Thousand Plateaus*. Minneapolis: University of Minnesota Press.
- Doran, J.E. and Hodson, F.R., eds. 1975. *Mathematics and Computers in Archaeology*. Edinburgh: Edinburgh University Press.
- Ford, J. A. (1954) "The Type Concept Revisited", *American Anthropologist* 56: 42-53.
- Fuller, S. 1988. *Social epistemology*. Bloomington: Indiana University Press.
- Gattiglia, G. 2015. "Think big about data: Archaeology and the Big Data challenge". *Archäologische Informationen* 38: 113-124.
- Gilboa, A., A. Karasik, I. Sharon, and U. Smilansky. 2004. "Towards Computerized Typology and Classification of Ceramics". *Journal of Archaeological Science* 31, no. 6: 681-694.
- Gitelman, L. ed. 2013. *"Raw Data" Is an Oxymoron*. Cambridge, MA: MIT Press.
- Gorodov, V. A. 1933. "The Typological Method in Archaeology", *American Anthropologist* 35: 95-102.
- Gosden, C. & L. Malafouris 2015. "Process archaeology (P-Arch)". *World Archaeology*, 47:5, 701-717.

Griffiths, S. 2017. "We're All Cultural Historians Now: Revolutions In Understanding Archaeological Theory And Scientific Dating". *Radiocarbon* 59: 1347-1357.

Gualandi, M.L., G. Gattiglia and F. Anichini, 2021. "An Open System for Collection and Automatic Recognition of Pottery through Neural Network Algorithms". *Heritage* 4: 140-59.

Haraway, D. 1991. *Simians, Cyborgs and Women: The Reinvention of Nature*.

London: Free Association Press

Hargrave, L. L. 1932. *Guide to Forty pottery types from Hopi country and the San Francisco mountains, Arizona*. Museum of Northern Arizona Bulletin 1.

Hörr, C. Lindinger, E. & Brunnett, G. 2014. "Machine Learning Based Typology Development in Archaeology". *ACM Journal on Computing and Cultural Heritage* 7, no. 1, Article 2 (March), 23 pages.

Kitchin, R. 2014. "Big Data, new epistemologies and paradigm shifts". *Big Data and Society* (April–June), 12 pages.

Kristiansen, K. 2014. "Towards a new paradigm. The third science revolution and its possible consequences in archaeology". *Current Swedish Archaeology* 22, no. (1): 11-34

Latour, B. & S. Woolgar 1979. *Laboratory Life: The Construction of Scientific Facts*. Beverly Hills: Sage Publications.

Latour, B. 1987. *Science in Action*. Cambridge MA: Harvard University Press.

Leonelli, S. 2016. *Data-Centric Biology: A Philosophical Study*. Chicago: University of Chicago Press.

Lucas, G. Forthcoming. "Tales of the Unexpected. Epistemic novelty in archaeology". In *The End of Epistemology in Archaeology? Indigenous Knowledges, Ontologies and the Crisis of Traditional Philosophies of Knowledge in Archaeology* edited by O.M. Abadía, K. Pollard, O. Marek-Martinez, C. Cipolla and S. L. Gonzalez. Tuscon: University of Arizona Press

- Mackenzie, D. 1984. "Marx and the Machine". *Technology and Culture* 25, no.3: 473-502.
- Marila, M. 2019. "Slow science for fast archaeology". *Current Swedish Archaeology* 27: 93-114.
- Marx, K. 1976. *Capital. Volume 1*. Harmondsworth: Penguin
- Meghini, C., R. Scopigno et al. 2017. "ARIADNE: A Research Infrastructure for Archaeology". *ACM Journal on Computing and Cultural Heritage* 99, Article 1.
- Mol, E. 2021. "'Trying to Hear with the Eyes': Slow Looking and Ontological Difference in Archaeological Object Analysis", *Norwegian Archaeological Review* 54, nos.1-2: 80-9.
- Morgan, C. & H. Wright, 2018. "Pencils and Pixels: Drawing and Digital Media in Archaeological Field Recording." *Journal of Field Archaeology* 43, no. 2: 136-151.
- Morgan, C., H. Petrie, H. Wright & J.S. Taylor 2021. "Drawing and Knowledge Construction in Archaeology: The Aide Mémoire Project". *Journal of Field Archaeology* 46, no. 8: 614-28.
- Mumford, L. 1967. *The Myth of the Machine. Technics and Human Development Vol. 1*. New York: Harcourt Brace/Jovanovich.
- Parisi, L. 2017. "After Nature. The Dynamic Automation of Technical Objects". In *Posthumous Life. Theorizing Beyond the Posthuman*, edited by J. Weinstein & C. Colebrook, 155-78. New York: Columbia University Press.
- Pawlowicz, L.M. and C.E. Downum 2021. "Applications of deep learning to decorated ceramic typology and classification: A case study using Tusayan White Ware from Northeast Arizona". *Journal of Archaeological Science* 130: 105375.
- Reide, F. 2006. "The Scandinavian Connection: The Roots of Darwinian Archaeology in 19th-Century Scandinavian Archaeology". *Bulletin for the History of Archaeology* 16, no.1: 4-19.
- Ribeiro, A. 2019. "Science, Data, and Case Studies under the Third Science Revolution". *Current Swedish Archaeology* 27:115–132.

Sharma, S. 2014. *In the Meantime. Temporality and Cultural Politics*. Durham, NC: Duke University Press.

Simondon, G. 1992. "The Genesis of the Individual". In *Zone 6: Incorporations*, edited by J. Crary & S. Kwinter, 297-319. New York: Zone Books.

Simondon, G. 2016. *On the Mode of Existence of Technical Objects*. Minneapolis: University of Minnesota Press

Soler, L., S. Zwart, M. Lynch & V. Israel-Jost, eds. 2014. *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science*. London: Routledge.

Sørensen, M.L.S. 1997. "Material Culture and Typology". *Current Swedish Archaeology* 5: 179-192.

Sørensen, M.L.S. 2014. "'Paradigm Lost' – on the state of typology within archaeological theory", In *Paradigm Found*, edited by K. Kristiansen, L. Smejda and J. Turek, 84-94. Oxford: Oxbow.

Spaulding, A. C. 1953. "Statistical Techniques for the Discovery of Artifact Types". *American Antiquity* 18: 305-313.

Stengers, I. 2017. *Another Science is Possible: A Manifesto for Slow Science*. Cambridge: Polity

Tyukin, I., Sofeikov, K., Levesley, J., Gorban, A.N., Allison, P. and Cooper, N.J. 2018. "'Exploring Automated Pottery Identification [Arch-I-Scan]'", *Internet Archaeology* 50

van Helden, D, Mirkes, E, Tyukin, I and Allison, P. 2022. "The Arch-I-Scan Project: Artificial Intelligence and 3D Simulation for Developing New Approaches to Roman Foodways". *Journal of Computer Applications in Archaeology*, 5, no.1: 78–95.

Wesson, C.B. and Cottier, J.W., 2014. "Big Sites, Big Questions, Big Data, Big Problems: Scales of Investigation and Changing Perceptions of Archaeological Practice in the Southeastern United States". *Bulletin of the History of Archaeology* 24, article 16.

Whallon, R. and Brown, J.A., eds. 1982. *Essays on Archaeological Typology*. Center for American Archaeology Press.

Wheeler, G. 2017. "Machine Epistemology and Big Data". In *The Routledge Companion to the Philosophy of Social Science* edited by L. McIntyre & A. Rosenberg, 321-329. London: Routledge.

Witmore, C. 2014. "Archaeology and the New Materialisms". *Journal of Contemporary Archaeology* 1, no. 2: 203-24.

Wylie, A., 2002. *Thinking from Things. Essays in the Philosophy of Archaeology*. Berkeley: University of California Press.