

Penultimate draft. This article will appear in the forthcoming volume Sterpetti, F., Bertolaso, M. (edited by). *Will Science Remain Human?*. Springer

PHRONESIS AND AUTOMATED SCIENCE: THE CASE OF MACHINE LEARNING AND BIOLOGY

Emanuele Ratti, eratti@nd.edu

University of Notre Dame

Abstract. The applications of machine learning (ML) and deep learning to biology has fostered the idea that the automated nature of algorithmic analysis will gradually dispense human beings from scientific work. In this paper, I will show that this view is problematic, at least when ML is applied to biology. In particular, I will claim that *ML is not independent of human beings and cannot form the basis of automated science*. Computer scientists conceive their work as being a case of Aristotle's *poiesis* perfected by *techne*, which can be reduced to a number of straightforward rules and technical knowledge. I will show a number of concrete cases where at each level of computational analysis, more is required to ML than just *poiesis* and *techne*, and that the work of ML practitioners in biology needs also the cultivation of something analogous to *phronesis*, which cannot be automated. But even if we knew how to frame *phronesis* into rules (which is inconsistent with its own definition), still this virtue is deeply entrenched in our biological constitution, which computers lack. Whether computers can fully perform scientific practice (which is the result of the way we are cognitively and biologically) independently of humans (and their cognitive and biological specificities) is an ill-posed question.

Keywords: automated science; machine learning; experimental biology; phronesis

1. INTRODUCTION

Recently, specialized and popular journals have emphasized the revolutionary role of big data methodologies in the natural and social sciences. In particular, there are two claims that fluctuate, in more or less explicit ways, in both expert and popularized pictures of science. First, big data generates new knowledge that traditional forms of the scientific method (whatever we mean by that) could not possibly do. Machine learning, it is said, is advancing and will advance science in unprecedented ways (Maxmen 2018; Zhou et al 2018; Zhang et al 2017; Schrider and Kern 2017; Angermueller et al 2016; Sommer and Gerlich 2013). Second – and this is strictly related to the first claim - methodologies and tools associated to big data can automatized science

in a way that humans will become dispensable. This is because humans' cognitive abilities will be rapidly surpassed by AI which will take care of every aspect of scientific discovery (Yarkoni et al 2011; Schmidt and Lipson 2009; Alkhateeb 2017). However, it is not clear what automated science in this context means. In this chapter, I will show that the idea behind automated science is not substantial, *at least in biology* (and in particular molecular genetics and genomics, which are my main case studies). I will make one specific claim, namely that *ML is not independent of human beings and cannot form the basis of automated science, if by science we have in mind a specific human activity*.

The structure of the chapter is as follows. First, I will define the scope of ML (Section 1.1). Next, I will show a number of concrete cases where at each level of computational analysis human beings have to make a decision that computers cannot possibly make (Section 2.1). These decisions have to be made by appealing to contextual factors, background, tacit knowledge, and researcher's experience. I will frame this observation in the idea that computer scientists conceive their work as being a case of Aristotle's *poiesis* perfected by *techne*, which can be reduced to a number of straightforward rules and technical knowledge. However, my observations reveal that there is more to ML (and to science) than just *poiesis* and *techne*, and that the work of ML practitioners in biology needs also the cultivation of something analogous to *phronesis*, which cannot be automated (Section 2.2 and 2.3) for the very simple of reason that its processes are completely opaque to us. Finally, I will frame this claim in psychological words by arguing (Section 2.4) that *phronesis* is exercised in terms of intuitions, and that how we develop intuition is not transparent. But even if knew how to frame the production of intuitions in terms of rules, still computers lack the biological constitution that is probably relevant to the way our intuitions are produced.

1.1 Machine Learning and its scope

Before delving into my analysis, let me briefly introduce ML and its scope. Let us start with a traditional definition of data science and then tailor it to ML. Dhar defines data science as “the study of the generalizable extraction of knowledge from data” (2013, p 64). This definition implies that there is something that we call ‘knowledge’ which is present in data sets. Knowledge here is not meant in any philosophical sense. Knowledge is understood as ‘pattern’, namely a discernible regularity in data. When Dhar says ‘generalizable’, he means that patterns detected

will somehow occur in data sets which are similar to the one from which patterns have been extracted. The regularity is numerical and devoid of content. Similarly, ML is the study of generalizable extraction of patterns from data sets. However, this does not tell much about ML works, and it does not tell exactly what ‘regularity’ really means. Let me address these issues.

It is not that ML is presented with a data set, and then it extracts patterns. ML can extract generalizable patterns present in a data set starting from a *problem*. This is defined as a given set of inputs variables, a set of outputs which have to be calculated, a sample (previously input-output pairs already observed), and a set of real-world situation assumed to be similar to the one described in the sample – this is like the scope of the problem. What ML aims to calculate are the outputs, which stand in a quantitative relation of *predictive nature* with inputs – anytime that there is x (the input), then the probability of having y is such and such (the output). The sample will instruct the algorithm about this relation by providing previous observed pairs of “having x and y ”. Therefore, a first goal of ML is to generate *predictions*. Put it more precisely, ML generates *predictive models* where a prediction is defined as *the computation of the values of outputs for an input variable whose outputs are unknown*. The easiest example is e-commerce. ML is applied to find out quantitative relations between inputs (consisting of a set of customers’ characteristics) and outputs (defined as buying specific products). The problem will be then composed of an input (i.e. the customers’ characteristics), a set of input-output pairs (customers’ characteristics and the products they bought), while the scope of the problem will be probably something like the class of products we are interested in. The solution to the problem will be to find out exactly the quantitative relation between inputs and outputs. Let’s make another example, this time from biology. A common example of ML in biology is the class of algorithms aimed at identifying transcription start sites¹ (TSS). Let’s say that we have an algorithm that we think is suitable to identify TSSs. The algorithm is provided with a large data set of TSS sequences. The algorithm processes such sequences and then it generates a model. At this point, the inputs for which we want to calculate the output (let’s say, a whole-genome sequence) is provided and the model identify the sequences that are likely to be TSS.

Another thing worth to emphasize is that predictions are about a specific context, where the context is defined by the scope of the problem. In the case of TSSs, the scope of the problem

¹ To simplify a bit, a TSS is where transcription of a gene starts, precisely at the 5’-end of a gene sequence

is DNA sequences and not, for instance, amino acids sequences, even though these are somehow related. Moreover, an algorithm that is trained to recognize TSSs in human whole-genome may not be appropriate for other organisms, or it will not be functioning well for recognizing, let's say, promoters. A related goal of ML is *classification*, namely the use of classes for categorizing data points, where classes are sets of objects which share some features such that they can be relevant for some purpose. Usually, data points are assigned to a label, which indicates what the particular data point is about (e.g. a TSS). Related but different, *clustering* is the determination of which data points are similar to each other without having a labelled data set. There are other similar goals as well.

Given these aims, there are different modes of 'learning' for algorithms in ML. First, there is *supervised learning*, which is when ML is provided with labelled data sets, namely when data points are already classified in groups. Labelled data points may function as training data for ML algorithms, and from this structured data sets ML learns how to predict the way data points labelled are correlated in new (but similar from the standpoint of a problem) data sets. The example of predicting TSSs is supervised learning, since the data set of sequences provided to the algorithm contain labelled data points (i.e. sequences that are known to be TSSs). Classification is an example supervised learning. Differently, *unsupervised learning* analyzes unlabeled data sets. This means that ML in this case has to *cluster* data points in various groups or categories. Clustering is an example of unsupervised learning..

2. AUTOMATED SCIENCE

What do we mean by 'automated'? Which are the aspects of scientific practice that we are talking about? By drawing from Humphreys (2011), let's distinguish two scenarios. First, there is the present situation (called the *hybrid scenario*), where science is carried out both by humans and machines. The idea is that AI optimizes certain aspects of scientific practice – mainly the quantitative aspects. This is uncontroversial. Much more controversial is the *automated scenario*, where computers replace all humans' activities and humans' epistemic authority in scientific practice. The idea is that in *science as we know it*, AI will replace humans and their expertise and their ways of gathering and producing knowledge.

In this section I will argue against the *automated scenario*. My claim is that *machine learning is not independent of human beings*. My argument on this issue is to argue that, as it

happens in experimental biology, also in computer science the contextual situation in which a scientist acts underdetermines the technical rules she is equipped with, exactly those technical rules that a computer will follow to solve a scientific problem. Therefore, ML needs human judgment at every significant step of its discovery path. I will spell out the idea of ‘judgment’ by appealing to a virtue ethics/epistemology terminology and the difference between *techne* and *phronesis* in Aristotle.

However, this argument is not just another version of the strategy “computers cannot do *x*”. My idea is to challenge the idea of the *automated scenario* itself. First, I will frame the idea of acting in the *phronesis* sense by appealing to the psychology of intuitions. Acting (well) by using intuitions seems to provide a contemporary psychological framework to the idea of *phronesis*. What emerged from the psychology of intuitions is that the way we educate our intuitions and the way we appeal to them to choose the right course of action is *opaque*. Therefore, because we do not know how our intuitions work, and scientists appeal to intuitions more than they are ready to admit, then we cannot automatize science. But – and here comes a possible objection - one may say that it may happen that one day we will make such a process transparent and hence elaborate a set of rules about it. However, the way we educate, accumulate and use intuitions is strictly connected to our biological constitution. Science is a human activity, and the way it is practiced and the shape of the results we obtain are dependent on the way humans are, not only cognitively but also biologically. But computers are not humans even if they could simulate human activities; computers are just different. This is of course applied to science only if science is seen as an intrinsically human activity; however, we do not know any other science that is not human.

2.1 Rules are not enough in machine learning

Can scientific discovery be automatized so as to make human beings dispensable? The answer to this question is rather complex. I will show first the problems of automated science in the practice of ML and then I will conceptualize such problems in a bigger philosophical picture.

Any ML task is faced with a number of issues that require humans to take a decision based on background knowledge, experience and other factors hardly reducible to computational abilities *tout court*. In a sense, I will claim that humans have intuitions formed on an experiential ground which we do not know how are formed (Hogarth 2001). This means that ML cannot be

completely automated because humans' judgment and intuitions² are required anytime uncertainty emerges, and an acceptable level of uncertainty will change according to the context and who will take the decision in the first place. Such issues are summarized in Table 1 and discussed below.

The first case where human's judgment is required is in the choice of the type of methodology. As Libbrecht and Noble reports (2015), sometimes you do not have much of a choice. For instance, if no labeled data set is available, then unsupervised learning is the only option. But, they note, having a labeled data set does not necessarily imply that supervised learning is the best option. The reason for this is that "every supervised learning method rests on the implicit assumption that the distribution responsible for generating the training data set is the same as the distribution responsible for generating the test data set" (Libbrecht and Noble 2015, p 323). This means that sometimes data sets are generated differently and they underlie different characteristics even if they are labeled with the same tags. As an example, Libbrecht and Noble report the case of an algorithm trained to identify genes in human genomes; this probably will not work equally well for mouse genomes. However, it may work *well enough* for the purpose at hand. In such situations the researcher has to identify, first of all, the fact that there might be a divergence between training and test sets. This divergence in ML applied to biology may also have a biological meaning, and it has to be scrutinized biologically and not necessarily numerically. Next, one has to judge whether such divergence is such that a supervised learning might generate unreliable results. But unreliable results are a function also of the type of expectation and ability of interpretation that scientists are employing. Hence, there is not a clear-cut threshold between being reliable and unreliable. In other words, the researcher has to judge whether the level of uncertainty is acceptable or not.

Another issue worth to mention is the problem of *encoding prior knowledge*. Especially in the case of supervised learning, this might be understood as the choice between framing the problem in terms of either regression or classification. Regression is when the output of the problem is a real value, while classification is when the output is a category. Let's consider the the analysis of nucleosome positioning from primary DNA sequences. Example of handling such problem in terms of regression is determining the coverage base of nucleosomes. However, we

² I will equate 'intuition' and judgement' because I am thinking about 'judgements' as 'intuitive judgements'.

may want to frame the problem in terms of *classification*, namely predicting nucleosome-free region. It is important to note that, in light of researcher's judgment and interests, we may switch from classification to regression and *vice versa*, and this will also change the way prior knowledge is encoded. The issue is, again, how one will understand the problem to be solved and how framing it as either regression or classification will be the best option to obtain a valid result. Related to encoding of prior knowledge there are other close issues. First, there is *feature selection*, namely the set of decisions related to the selection of the features of a data set that are relevant to the problem one is facing. It is a case of implicit priori knowledge, as in any application of ML, "the researcher must decide what data to provide as input to the algorithm" (Libbrecht and Noble 2015, p 328). But the problem is not just how to deal with data you have; rather, it can be problematic also to judge data that you *do not* have, or data sets with poor quality. This is the problem of *missing data*, which can be of two types. First, values that are missing randomly or for reasons that have nothing to do with the problem one is trying to solve (e.g. poor data quality). Next, there are values that, when missing, are important for the task one is executing. Different strategies derive from the different types of missing data.

How to select data and how to deal with tags is not a problem that only users of databases have. There is compelling empirical evidence that computer scientists creating databases (i.e. database curators) have similar problems (Leonelli 2016, Chapter 4). Rules for standardizing heterogeneous data sets, label them, and make them available are underdetermined by the contexts and the specificities in which such procedures are actually used, because databases are curated on the basis of the foreseeable uses that other scientists will do of the database, the type of user that one imagines will use the database, etc and all these factors will vary substantially, even within well cohesive scientific communities. Also, the very procedure of curating data and thinking about effective metadata has been described by Leonelli as a procedure to codify *embodied knowledge* of experiments and their environments and this "involves making a decision about which aspects of the embodied knowledge involved in generating data are the most relevant to situating them in a new research context" (Leonelli 2016, p 97). In order to do this, being a computer scientist is not enough; you need also an expertise of the particular research context the database will be relevant to. Because of the importance of contextual judgments on the side of both curators and users, "human agents remain a key component of the

sociotechnical system, or regimes of practices, that is data-centric biology” (Leonelli 2016, p 112).

Furthermore, there is an additional problem if one is provided with a predictive model generated by unsupervised learning. Usually a solution from unsupervised learning provides a set of labels which are not interpreted – they are empty. A ML practitioner has to interpret such labels and connect them to the content of the science to which she is applying the algorithm – in our case, molecular biology. This may require collaborative research and integration of different epistemological cultures, which can be hardly automated. But even if one possesses both expertise, still there can be many ways of interpreting biologically results, especially in light of different epistemic aims. This is a case where pluralism should be cultivated or, as Chang says, “the great achievements of science come from *cultivating* underdetermination, not by getting rid of it” (Chang 2012, p 151).

Another issue worth to be mentioned has to do with the evaluation of the performances of algorithms. According to ML practitioners, there is a tradeoff between *sensitivity* and *precision*. Let’s say we want to predict the location of enhancers in a genome. We can either *value sensitivity* – the fraction of positive examples (enhancers) identified – or *value precision* – the percentage of predicted enhancers which are truly enhancers. These values stand in a tradeoff relation, and hence according to the task at hand we may end up valuing either one or the other.

Technical Issue in ML	Description of the Issue	What Cannot be Automated
Choice of the algorithm	Is there a strict set of rule for choosing between supervised and unsupervised learning?	Whether one judges the divergence of training and test sets as impacting significantly the reliability of results
Encoding Prior Knowledge	Shall we frame the problem in terms of regression or classification?	Choice between regression and classification is a function of the task, background knowledge, and experience of the researcher
Feature Selection	Selection of the features of a data set that are relevant to the problem one is facing	Choice of features is a function the task, background knowledge, and experience of the researcher
Data Curation	How do we choose significant tags to label a data set?	Choice of labels depends on the foreseeable uses of the database, which depend in

		turn on the task, background knowledge, and experience of the curator
Labels interpretation	Semantics must be manually assigned to labels	Interpretation of labels in a biological sense requires collaborative research
Evaluation of the performance of an algorithm	Should we value precision or sensitivity?	Choice is a function the task, background knowledge, and experience of the researcher

Table 1. Aspects of Machine Learning practice that cannot be automated

The upshot of all these issues is that working in ML requires making important decisions all the time. Algorithms are not independent, and sometimes in addition to ML knowledge of theoretical and practical constraints one needs to be proficient also in biology. Being an expert in both fields is very difficult. This is why bioinformatics is a collaborative enterprise; collaboration requires the ability to mediate and negotiate between very different areas of expertise; these are skills which are cultivated through experience, and they cannot be standardized in rules.

How do we frame philosophically all these problems? Which are the consequences that we should draw from such technical issues? What I will do is to interpret the view (and its denial) of automated science in light of some virtue ethics/epistemology basic distinctions.

2.2 Experimental science and rules

In order to frame the analysis of the problems mentioned above in light of some basic virtue-ethics/epistemology concepts, we need to take a step back to experimental biology.

Experimentation in molecular biology is characterized by an intense material manipulation of biological systems. The development of standardized techniques and the ingenious tinkering of biological entities are hallmarks of molecular biology (Rheinberger 1997; Ratti 2018). Experiments in molecular biology are rarely performed to test theories in the strict sense of *crucial experiments*. Usually the work of a molecular biologist is to develop a working model or a hypothesis about a biological system. Molecular biologists start from a general hypothesis about how a biological entity influences a set of processes. On the basis of a general hypothesis, they derive predictions and instantiate experiments to make sense of these predictions. Results provide biologists with other predictions that call for other experiments. Ensuing experiments refine the initial model by providing specific information about the entity

under scrutiny. The experimental system is materially stimulated to reveal increasingly precise information:

“This process of shaping and refining hypotheses through experiments continues, virtually, *ad libitum*. This is a sort of progressive and ramified (but not-linear) deductive process, developed by poking and prodding experimental systems” (Boem and Ratti 2016, p 150)

The aspect to emphasize is that, at each step, biologists have to stimulate the experimental system in unexpected ways, depending on the task at hand, the availability of resources and the nature of the experimental system itself. Even though there are plenty of standardized experimental techniques in molecular biology, still biologists have to modify their protocols to adapt them to the characteristics of the experimental system at hand. It seems that biologists have to ‘feel’ the experimental system or have to (metaphorically) ‘dialogue’ with it and interpret its responses accordingly (Keller 1983). Shannon Vallor makes sense of this complex dynamics by appealing to the cultivation of a specific scientific virtue called *perceptual responsiveness*, which is defined as

“a tendency to direct one’s scientific praxis in a manner that is motivated by the emergent contours of particular phenomena and the specific form(s) of practical and theoretical engagement they invite” (Vallor 2014, p 271)

In other words, perceptual responsiveness is a disposition to direct scientific praxis towards a proper appreciation of the affordances that a phenomenon under investigation yields. The virtuous experimental scientist is one who “properly reads or ‘decodes’ all of the salient invitations to measurement implied by the phenomenon (...) and creatively finds a way to take up just those invitations whose answer may shed the most light” (Vallor 2014, p 276). However, an aspect that Vallor does not emphasize enough is the dependence of the scientific praxis on the experience of the researcher. I understand the phenomenological (Husserlian) concept of ‘invitation’ in terms of ‘affordances’. The idea of ‘affordance’ comes from psychology, and it can be defined as a property of an object that suggests how to interact with or use that object (Ratti 2018). While ‘affordances’ exist independently of the observer, the fact that they can be

‘exploited’ depends on the actor interacting with the environment. The scientist must be sensitive to the ways affordances emerge from phenomena, and for this reason she must cultivate other important virtues such as perseverance, diligence, insight, adaptability, etc; *perceptual responsiveness is an umbrella virtue because it implies many other virtues.*

The way a phenomenon is modelled is the result of the interaction of scientist’s background, experiences, aims, the ability to properly see the affordances and know how to deal with them, and the phenomenon’s characteristics. Scientific inquiry on this matter looks like value judgment rather than a mere clear-cut procedure by applying rules. Perceptual responsiveness is the disposition to see and respond appropriately to the affordances of biological phenomena. This is exactly the excellence that the experimental biologist strives to cultivate; by starting from a general hypothesis or an incomplete model, the biologist poses certain questions in forms of experiments to an experimental system, and by interpreting its answers in light of her experience and the components of the experimental system itself, she poses more accurate questions, until she elaborates a model of the phenomenon which is adequate – or, to use Rheinberger’s terminology (1997), until she brings into light the opaque aspects of the *epistemic thing* (i.e. the phenomenon under investigation) and she turns it into a *technical object*.

An aspect that Vallor emphasizes is that perceptual responsiveness is a virtue because it is not rule-driven, but it is cultivated as a habit of seeing and acting in specific (both theoretical and experimental) contexts with the aim of achieving epistemic goals. Moreover, such habits are what make “a scientist exceptionally praiseworthy and a model of excellence” (Vallor 2014, p 277).

My claim is that a similar virtue has to be cultivated by ML practitioners as well. This claim is motivated by the fact that even a quantitative enterprise such as ML requires a deliberation which is not just technical, but it depends on many other factors that will change from context to context, depending both on the nature of the data sets analyzed and the characteristics of the ML practitioner herself, and hence cannot be expressed fully in an algorithmic form. The issues emphasized in Table 1 are just a few of the contextual decisions that ML practitioners have to take, and that each of them will take in a different way, depending on their background and other contextual factors. In order to spell out this view properly, let me characterize the position of automated science in light of some basic distinctions.

2.3 Techne, phronesis and automated science

In Aristotle, there is a distinction between theoretical knowledge and practical knowledge³. Theoretical/scientific knowledge is *episteme*, which can be understood as demonstrative knowledge of the necessary and the eternal (EN VI.3). Theoretical knowledge has no practical import (Dunne 1997, p 238). On the other hand, practical knowledge includes *praxis* (usually translated as ‘acting’ or ‘doing’) and *poiesis* (usually translated as ‘making’). The distinction is often put in the following terms; *poiesis* is related to the production/fabrication of things, and the success of an action is measured by the goodness of its product, while in the case of *praxis* the aim of an action is the good of the action itself⁴ (EN VI.4-5). Examples of *poiesis* would be making a chair or building a bridge. Examples of *praxis* would be actions that are constitutive to the good life, to human flourishing etc, which are usually understood in the context of morality or ethics. *Poiesis* is perfected by *techne*, which is understood as the knowledge of a set of rules and standards that are applied in order to make a well-constructed and well-formed external product. *Praxis* is perfected by *phronesis*, which is a very complex excellence, being an umbrella virtue as Vallor’s perceptual responsiveness. *Phronesis* is practical knowledge, namely being able to recognize the salient features of a situation and act according to the choice of the right mean between extremes. Dunne (1997) puts the distinction in terms which are useful for the argument of this section:

“Techne provides the kind of knowledge possessed by an expert (...), a person who understands the principles underlying the production of an object or state of affairs (...) Phronesis, on the other hand, (...) is acquired and deployed not in the making of any product separate from oneself, but rather in one’s action with one’s fellow. It is personal knowledge in that, in living of one’s life, *it characterizes and expresses the kind of person that one is*” (p 244, emphasis added)

Whether Aristotle implies that we can sharply isolate *poiesis* from the constraints of *phronesis* - in the sense that ‘making’ will be independent from ‘ethics’ and from all the tacit and implicit constraints that *praxis* is subjected to – is an open question, which I do not seek to answer.

³ My reading of these aspects of Aristotle is mostly drawn from (Dunne 1997)

⁴ “[T]he end of production is something other than production, while that of action is not something other than action, since doing well in action is itself action’s end” (EN VI.5, 11140b)

Rather, I want to focus on the consequences that insulating *poiesis* from *praxis* have for contemporary science.

When I talk about the constraints of *praxis* I have in mind a specific interpretation of this claim. Dunne (1997) tries to make sense of *praxis* with the idea of *incommensurability*. This emphasizes the plurality and heterogeneity of measures and factors to which we appeal to make sense of one's action. Incommensurability insists "on the disparateness of the factors that will weigh with *different people* faced with the same situation" (Dunne 1997, p 47). Dunne connects this idea both with Aristotle's notion of *praxis* and deliberation and, interestingly enough, with Kuhn's remark that there cannot be algorithms for theory choice in science because every scientist will value different epistemic desiderata in different ways (Kuhn 1977). This happens because background and tacit knowledge, experience, goals and other contextual factors will make each scientist take a specific and unique turn towards theory⁵. Understood in this way, *praxis* has a parallel in the numerous factors through which perceptual responsiveness is exercised in the experimental sciences.

With the distinction between *techne* and *phronesis* framed in terms of the former being dictated by technical knowledge in the form of rules to produce something, and the latter as actions resulting from judgments which are influenced by a myriad of factors, let's turn to the natural sciences. How do we understand them? An interesting proposal is to see natural sciences as *poiesis*⁶. In fact, "[o]ur contemporary conception of scientific research (...) is deeply *poietic* (...) As with *poiesis* more generally, the goal of scientific research is a product (...) external to the scientist herself, and that product can be detached from the scientist and evaluated independently" (Stapleford 2018, p 4-5). It is not a coincidence that standards of scientific knowledge are usually expressed in form of adherence to protocols and technical skills which can be meaningfully understood in light of the typical *techne* language. Interestingly, Habermas

⁵ It is also important to point out that when it comes to the components that make human judgment, values of various sorts emerge, not only epistemic values. In philosophy of science this has been documented in several ways, both theoretically and empirically

⁶ A much more straightforward way to solve the issue would be to think about natural sciences as *theoria*, but if we pay attention to the way Aristotle meant this term and its cognate *episteme*, it is easy to see that natural sciences have not many things in common with *episteme* and *theoria*. Stapleford (2018) rightly emphasizes that *episteme* is described as knowledge of the eternal, it is achieved through demonstration, and it is about contemplation; "what admits to be known scientifically is by necessity" (EN VI.3 1139b). On the other hand, natural sciences are knowledge of the contingent, demonstration has not necessarily an epistemic priority over other discovery strategies (e.g. material manipulation, statistical analysis, etc) and theorizing – rather than contemplating – is an important aim.

explains the confusion that may emerge in blurring *techne* and *episteme* in terms of a deliberate operation instantiated by modern scientists, in the sense that “the modern scientific investigation of nature set about to pursue theory with the attitude of the technician” (Habermas 1974, p 61). In our case, ML practitioners think about science as *poiesis*, and *techne* as the recital of technical skills needed to increase its products. The fact that computer scientists think about a fully automated science is because they reckon that ‘making science’ can be fully isolated from all the implicit and tacit components that are required to deliberate in a *praxis* environment. Therefore, at the very core of the idea of contemporary science lies the idea of science as a form of *poiesis*; modern science can be conceived as *poiesis* that can be fully isolated from *praxis*. Possessing *techne* would imply only “a logical progression through means to a predetermined end” (Stapleford 2018, p 9). Automated science is an algorithm that possesses *techne*, and it will progress towards the achievement of certain products without the constraints typical of *praxis* that I mentioned above.

However, my claim – which is analogous to Stapleford’s analysis (2018) – is that contemporary science is not just *techne*. Actually, science aims at being *poiesis* dictated only by *techne*, but in fact it is more dependent on something analogous to *praxis* than scientists would admit. In my list of situations where contextual decisions are required for ML practitioners, the dominance of *techne* might be undermined. Stapleford draws an interesting parallel to *praxis* when it comes to scientific practice, and in the case of ML the parallel can be made even more precise. The ML practitioner wants to act justly, where ‘just’ is in relation to the adequacy of her model. But we have seen that judging what counts as adequate requires something analogous to *phronesis*. It is not just about following a set of predetermined rules to construct a model (namely, the technical knowledge required to handle data), but also to identify the best mean given the particular situation the practitioner is dealing with and the level of uncertainty. The issue of missing values is a good case in point; depending on the nature of missing values (e.g. poor data quality) the researcher must make a choice as how to cope with such missing values. Feature selection is another example; depending on what our interests are, and the way the data sets are structured, we may prefer a regression or a classification. But even more basic choices such as shifting from a supervised to an unsupervised algorithm require a sort of *phronesis*; judging whether the provenance of a data set and the way it has been labeled can provide a sufficient level of certainty to assure the reliability of the results we may possibly obtain will

depend on several constraints, many of which will be tacit and strictly connected to the researcher's background. Moreover, judging a numerical result such as a ML predictive model in a biological sense goes beyond a simple automatized task; it requires experience in the biological field that a researcher in computer science may not possess. If the researcher lacks such an experience, she may ask for collaboration with a biologist, and collaborative research is the result of the integration of disparate fields of knowledge that no computer can possibly achieve. Computers suggest, humans decide⁷. On this point, it is interesting to notice that both *phronesis* and Aquinas' *prudentia* require the cultivation of virtues (such as caution, foresight, docility, etc) that could apply pretty well in the case of ML. A practitioner may exercise *caution* when getting certain values, and she has to be open to learn (i.e. *docility*) in interpreting results by listening to what the biologist has to say about her result. A data curator must cultivate *foresight* in order to imagine the different ways in which her database could possibly be consulted. But this requires *docility* again (Bezuidenhout et al 2018), because a dialogue with actual users may change the way a data set is labeled. Docility is also fundamental when it comes to interpret the labels according to which a data set has been partitioned as a result of unsupervised learning; once the unsupervised algorithm has found the perfect way to partition a data set in terms of *n* labels, then a collaboration between biologists and computer scientists is in order if the aim is to *act justly* (i.e. to elaborate an adequate model).

Even though the predetermined end is the same (i.e. the construction of an adequate model), still the way this is achieved is the result of a complex network of constraints and decisions which are not 'end-directed automation'. What a highly sophisticated and technical field as ML needs is striking the right balance between *techne* and *phronesis*. A practitioner in ML needs both the technical knowledge in terms of protocols and rules provided by something like *techne*, and the experience, the capacity to 'sense' and 'feel' the context and a set of basic virtues that are provided by something like *phronesis*. It is no surprise that the situation is similar in an experimental field such as molecular biology. I am not claiming that what Vallor calls perceptual responsiveness is the *phronesis* required by computer scientists in bioinformatics. Actually, it does not matter how we call such a contextual *phronesis*. The important thing is that both biologists and computer scientists clearly need the technical knowledge to manipulate

⁷ This is analogous to Leonelli's remark that "[r]ather than replacing users' expertise and experience in the lab, metadata serve as *prompts* for users to use embodied knowledge to critically assess information about what others have done" (Leonelli 2016, p 106)

entities or data sets (e.g. protocols), but at the same time perceptual responsiveness and/or *phronesis* will dictate the way such technical knowledge is modified, adapted and used in specific contexts. *Techne* and *phronesis*/perceptual responsiveness here inform each other.

2.4 A possible objection and reply

My claim can be summarized by saying that the way a ML practitioner does research will depend not only on the technicalities of computer science, but also on the intuitions she has about her research object given her experience and research context – how she ‘feels’ and ‘interprets’ the specific context, and what her ‘guts’ say. Therefore, if the *automated scenario* requires that AI will perform science (as a human activity) by replacing humans in all their activities, then this is simply not happening and cannot possibly happen. The misunderstanding lies in conceiving science as *poiesis* driven exclusively by *techne*, while science is *informed* by both *phronesis* and *techne*.

However, one may buy this claim but not the consequences of this claim. One may say that yes; this is the way science is practiced today. But this does not mean that we cannot make the process of ‘feeling’ the experimental system less opaque, and somehow build a computer designed in a way that has ‘intuitions’ in the same way we do. And we should do this because our intuitions can be so good that we want to optimize them in a computer. In principle, this sounds right. But there are many things that in principle are legitimate but in practice are misleading. Let me be more precise of what is opaque and what is not, and then I will deal with the objection.

I have said that any scientist needs a sort of *phronesis*, namely a capacity to sense the situation and being able to recognize its salient features and act adequately to achieve a particular goal. Recognizing the salient features and act accordingly is having *an intuition* about the situation. Such intuitions are based on experience – we collect information about similar cases, somehow we internalize this set of information, and anytime we see an analogous case we almost automatically act in certain ways. However, the way we internalize such things and how these intuitions emerge is completely opaque – we have no clue about how we do it. Ironically, ML algorithms may be opaque, but human beings are opaque as well, and this is why we cannot turn into a set of rules the way a person possessing *phronesis* acts – or at least now.

What is exactly an intuition? Here I rely on some works in psychology, in particular (Hogarth 2001). Hogarth makes several examples of people taking decisions in their field of expertise without really deliberating about those things; “they have a feeling of knowing what decision to make (...) they depend on their ‘intuition’ or ‘sixth sense’, and this they cannot explain” (Hogarth 2001, p 3). What is important to highlight is that the process of intuition is characterized by a lack of awareness of how it happens – unlike deliberative processes, intuitive processes are characterized by the impossibility of being made transparent. We can rationally reconstruct them, but rational reconstructions may be made in several ways. Interestingly, intuitions seem to depend on personal experiences, and hence each of us will have his/her own intuitions. The idea is that

“over time, we all build up a stock of ways of interpreting the world (...) [we] are typically not aware of how we acquired the knowledge that we use daily (...) the outcomes of intuitive processes are typically approximate and often experienced in the form of feelings” (Hogarth 2001, p 9)

According to Hogarth, on pragmatic grounds we can think about two systems that control the processes by which we learn and we act accordingly to what we have learnt. First, there is the *deliberate* system, which deals with all processes that require attention and deliberation. Next, there is the *tacit* system, which controls all processes occurring with little or no use of conscious attention. In my understanding, intuitions are forms of expertise, in the sense that anytime an agent learns by associating things pertained to a certain domain or by acting in a certain way, then what has been learnt is somehow internalized in the tacit, and when similar situations emerge, intuitions about how to act or to understand the situation emerge as well. However, we do not know how these processes work. Interestingly enough, this is also what happens in ML; algorithms of ML – especially deep learning – evolve as they are trained with more and more data sets. Sometimes they are so complex that they are black-boxes – they generate results, but we do not know how they do it. In the context of human learning, we just know that, by having experience in a particular field, then most of the things we will do in that domain will be intuitive behavior. My idea is that *phronesis* or *perceptual responsiveness* is exactly the result of educating intuitions in specific fields.

However – and we return to the objection - one may say that we do not really know if these processes cannot be made transparent. Therefore, we can imagine a situation where

psychologists identify clearly such processes, and a computer scientist tailors them into an algorithm to foster automated science. I do not think this is a very powerful objection, and for the following reason. As Hogarth recognizes, humans are composed of many information-processing systems, and hence our intuitions are the results of all these systems working together. We have no idea of how these complement each other, but what we can say is that they are deeply embodied *in our biological constitution*. Therefore, we should expect to have intuitions in a certain way and to act in a certain way also because we are biologically in such and such a way. We are not just *res cogitantes*. There is nothing unreasonable in thinking that also science is the result of the way we are. But computers are not like us, and hence we should not expect to have them the same intuitions and ‘feelings’ that we have; computers are just different. Therefore, automating science by looking at what humans and computers can and cannot do is like trying to teach to a dog how to be a cat. Of course, this applied to ‘anthropocentric’ science, and it leaves open the possibility of a scientific practice AI-tailored, as some commentators suggested⁸ (Humphreys 2004), even though I do not think I fully understand this statement. But even if I lack a clear understanding of these claims, still referring to science and outdistancing it from the human is quite strange, because we do not know any other science that is not human⁹.

To sum up, because (1) any scientific activity is dependent upon *phronesis*, (2) *phronesis* is the way it is as a consequence of the way humans are, and (3) computers are not like us, then attempts to automatize completely scientific practice are hard to understand.

4. CONCLUSION

In this article, I have argued that science as a human activity cannot be automated on the basis of ML because a look at the practice of science in this context shows that rules are not enough to lead scientific discovery. The philosophical lesson is that something similar to Aristotle’s *phronesis* – the ability to sense and feel the specificities of the context and act (epistemically) justly – is required in dealing with complex quantitative models elaborated by ML algorithms.

⁸ “[A]n exclusively anthropocentric epistemology is no longer appropriate because there now exist superior, non-human, epistemic authorities” (Humphreys 2004, pp 4-5)

⁹ To use Humphreys’ terminology, I do think that the *hybrid scenario* in which we cannot completely abstract from human cognitive abilities in science is here to stay even with ML

REFERENCES

- Alkhateeb, A. (2017). Can scientific discovery be automated? *The Atlantic*.
- Angermueller, C., Pärnamaa, T., Parts, L., & Stegle, O. (2016). Deep learning for computational biology. *Molecular Systems Biology*, *12*(7), 878. <http://doi.org/10.15252/msb.20156651>
- Aristotle. (2014). *Nicomachean Ethics*. (C. D. . Reeve, Ed.). Indianapolis: Hackett Publishing Company.
- Boem, F., & Ratti, E. (2016). Towards a Notion of Intervention in Big-Data Biology and Molecular Medicine. In G. Boniolo & M. Nathan (Eds.), *Philosophy of Molecular Medicine: Foundational Issues in Research and Practice* (pp. 147–164). London: Routledge.
- Chang, H. (2012). *Is Water H2O? Evidence, Realism and Pluralism*. Springer.
- Dhar, V. (2013). Data Science and Prediction. *Communications of the ACM*, *56*(12), 64–73. <http://doi.org/10.2139/ssrn.2086734>
- Dunne, J. (1997). *Back to the Rough Ground - Practical Judgment and the Rule of Technique*. Notre Dame: University of Notre Dame Press.
- Hogarth, R. (2001). *Educating Intuition*. Chicago: The University of Chicago Press.
- Humphreys, P. (2011). Computational science and its effects. In M. Carrier & A. Nordmann (Eds.), *Science in the Context of Application* (Boston Stu). Dordrecht: Springer.
- Leonelli, S. (2016). *Data-centric Biology*. Chicago: University of Chicago Press.
- Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321–332. <http://doi.org/10.1038/nrg3920>
- Maxmen, A. (2018). Deep learning sharpens views of cells and genes. *Nature*, *553*, 9–10.
- Ratti, E. (2018). “Models of” and “models for”: On the relation between mechanistic models and experimental strategies in molecular biology. *British Journal for the Philosophy of Science*.
- Rheinberger, H.-J. (1997). *Toward a History of Epistemic Things: Synthetizing Proteins in the Test Tube*. Stanford University Press.
- Schmidt, M., & Lipson, H. (2009). Distilling Natural Laws. *Science*, *324*(April), 81–85. <http://doi.org/10.1126/science.1165893>

- Schrider, D. R., & Kern, A. D. (2017). Machine Learning for Population Genetics: A New Paradigm. *Trends in Genetics*. <http://doi.org/10.1101/206482>
- Sommer, C., & Gerlich, D. W. (2013). Machine learning in cell biology – teaching computers to recognize phenotypes. *Journal of Cell Science*, 126(24), 5529–5539. <http://doi.org/10.1242/jcs.123604>
- Stapleford, T. (2018). Making and the virtues: The ethics of scientific research. *Philosophy, Theology and the Sciences*.
- Vallor, S. (2014). Experimental Virtues: Perceptual Responsiveness and the Praxis of Scientific Observation. In A. Fairweather (Ed.), *Virtue Epistemology Naturalized: Bridges Between Virtue Epistemology and Philosophy of Science* (pp. 269–290). Springer.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8(8), 665–670. <http://doi.org/10.1038/nmeth.1635>
- Zhang, L., Tan, J., Han, D., & Zhu, H. (2017). From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug Discovery Today*, 22(11), 1680–1685. <http://doi.org/10.1016/j.drudis.2017.08.010>
- Zhou, Z., Tu, J., & Zhu, Z. J. (2018). Advancing the large-scale CCS database for metabolomics and lipidomics at the machine-learning era. *Current Opinion in Chemical Biology*, 42, 34–41. <http://doi.org/10.1016/j.cbpa.2017.10.033>