UNIVERSITY OF BIRMINGHAM | THE JUBILEE CENTRE FOR CHARACTER & VIRTUES

Citizenship in a Networked Age

Dominic Burbidge

This is an unpublished conference paper for the 8th Annual Jubilee Centre for Character and Virtues conference at Oriel College, Oxford University, Friday 3rd – Sunday 5th January 2020.

Jubilee Centre for Character and Virtues

University of Birmingham, Edgbaston, Birmingham, B15 2TT United Kingdom

T: +44 (0) 121 414 3602 F: +44 (0) 121 414 4875

E: jubileecentre@contacts.bham.ac.uk W: www.jubileecentre.ac.uk

a) The networked age

The networked age can be defined as that period of history where interpersonal distance was eroded and decision-making became human-machine cooperative. This paper asks the following question: in the networked age, what is the space for civic engagement? Both components of the networked age have profound effects for the political, civic and social dimensions of what counts as a flourishing world. They challenge the fundamentals of our civic engagement. As interpersonal distance is eroded and decision-making becomes human-machine cooperative we lack a robust account of *good citizenship aimed at human flourishing*, and this is the gap the present paper seeks to fill.

An aim that lies behind both the erosion of interpersonal distance and human-machine cooperation is *efficiency*. Efficiency is a coordinating good for the networked age because it accelerates the satisfaction of individual needs and desires without passing judgment on those needs and desires. The goal of efficiency is politically liberal in that it achieves minimal consensus, tolerant of individual-level diversities, differences and valuations. While a commitment to efficiency, in this sense, rejects principled engagement with the ideological debates of the 20th century, it retains much of that century's fascination with progress through science and rationalisation, directing that fascination to non-totalitarian ends.

Progress is fastest where there is social agreement—or at least belief in future likely agreement—and that is a process inclusive of civil society to some extent. The age of totalitarianism provides a constant reminder of the need to maintain communicative and educational links with the public when advancing science and reason. But it often seems to the experts that the advancements are not much helped by popular review of their ethical implications. Scientific and technological developments are more specialised than ever before, making it harder and harder for the non-specialist to know what is going on or to make predictions about where the developments will likely take us. Even more difficult, it seems the advent of human-machine cooperation means specialists are themselves partially cut-out from giving full explanations of what progress is and where it will lead.

While the age of totalitarianism's requirement of obedience without thinking from vast swathes of the human race is easy to regret, the networked age we are entering into does not have a clear account of the role of citizens and so presents a similar problem. Distrust between the "haves" and the "have nots" is now not simply an economic question of appropriate distribution but a growing existential question about who the decision-makers really are in the new world that is shaping up.

A defence against the threat that decision-making is being taken out of the hands of ordinary people is that scientific and technological progress simply helps realise the self-determination of individual people, meaning they remain ultimately in charge and will not have to obey arbitrary authority in the use of these technologies. That holds true as a defence not ethically, existentially nor technologically. *Ethically* speaking, it requires consensus on a libertarian position of everyone's self-

defined morality being tolerable, and such consensus is not apparent. Libertarianism enjoys little popularity and morally incompatible options cannot actually be tolerated on grounds that truth itself is self-determined. Then, *existentially* speaking, the defence that scientific and technological progress serves individual self-determination also does not hold. Humans are social by nature and realise their aims as part of groups (families, schools, social classes). This means that pursuing one's self-determination as an individual does not existentially represent what society is or the fact that we often seek group-determination, not self-determination. Finally, *technologically* speaking the self-determination justification also fails because machine-human cooperation is leading to detachment from human agency in the execution of tasks and the optimisation of solutions.

It follows, therefore, that accelerating efficiency in the pursuit of self-determination cannot be a robust coordinating good for the networked age. And, without overall purpose to our technological progress, the role of civic engagement and democracy itself becomes less and less clear. Despite, in principle, retaining commitment to democratic rule, the space for civic engagement going forward is hard to define. Indeed, advances in science and technology point to an upcoming tough choice between efficiency and democracy, framing civic engagement as sub-optimal and rationally backward. What is this goal of efficiency that at times seems juxtaposed to democracy?

Henry Kissinger, Eric Schmidt and Daniel Huttenlocher write that the enlightenment replaced Christianity's emphasis on the divine with an emphasis on individual reason. They further that the age of the internet and artificial intelligence threatens to displace that emphasis on individual reason as artificial intelligence automates the drawing of conclusions from data and the action that follows from those conclusions.[1] Over the different ages there have been changes in the coordinating goods that society can be said to be seeking, though such paradigms are near impossible to fully demarcate—from seeking the divine (oneness with the creator), to sound individual reason (oneness with oneself), to efficient data-gathering and automated execution (oneness with doing). Alongside these shifts are concurrent shifts in our conceptualisations of the opposites to the goods that we seek—the evils we try to avoid. Vaguely, one can outline the following transitions through history on what counts as bad or evil, with our current networked age driven by the avoidance of inefficiency, a commitment begun in the industrial revolution:

[1] Kissinger, H. A., Schmidt, E. & Huttenlocher, D., 'The Metamorphosis'. *The Atlantic* (Aug 2019). https://www.theatlantic.com/magazine/archive/2019/08/henry-kissinger-the-metamorphosis-ai/592771/ (emphasis in original).

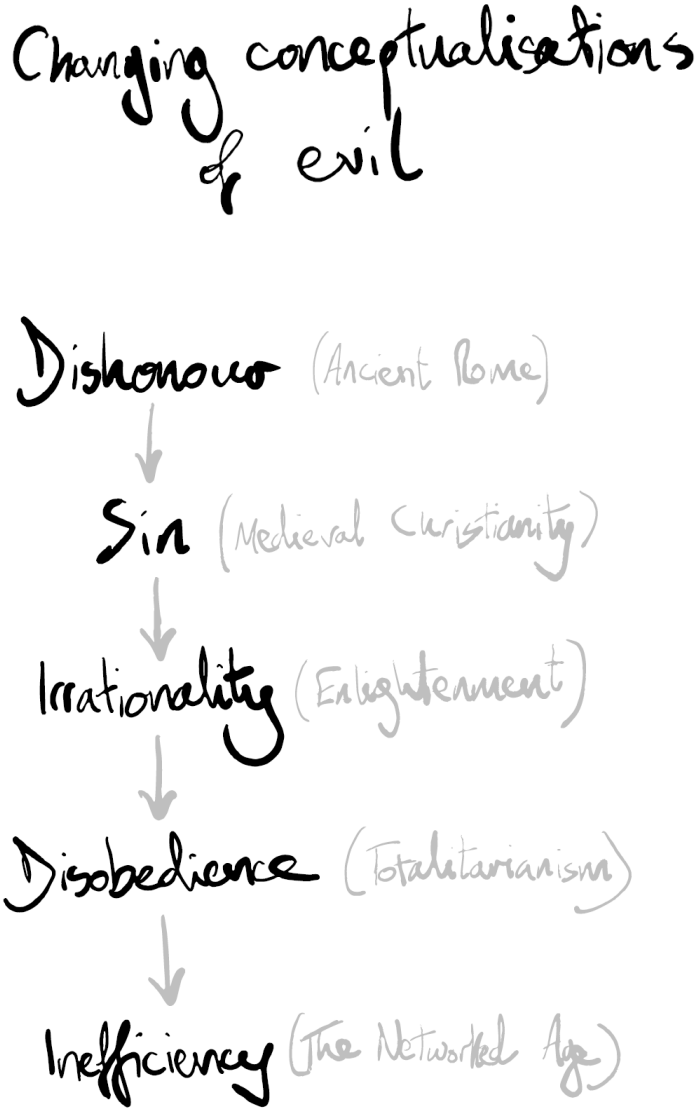

Figure 1.1 Key differences in conceptualisations of evil through the ages

Kissinger writes:

The Enlightenment sought to submit traditional verities to a liberated, analytical human reason. The internet's purpose is to ratify knowledge through the accumulation and manipulation of ever expanding data. Human cognition loses its personal character. Individuals turn into data, and data become regnant.[2]

[2] Kissinger, H. A., 'How the Enlightenment Ends'. *The Atlantic* (Jun 2018). https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/.

In the face of this chilling assessment, and the apparent struggle in process between efficiency and democracy, the present paper seeks to make a starkly positive contribution: we can and will find space for citizenship in the networked age and it is a space that completes, rather than competes with, our journey towards human flourishing.

b) Big data & big computation

A game changer in human-machine cooperation has come through big data and big computation. This refers to the way in which we are now able to amass data on an unprecedented scale, and then apply fast computational processes in sorting and analysing that data. Importantly, the process can be made to engage in a loop of further applying computation to the sifted data and further refining the predictive model, a process also known as machine learning.

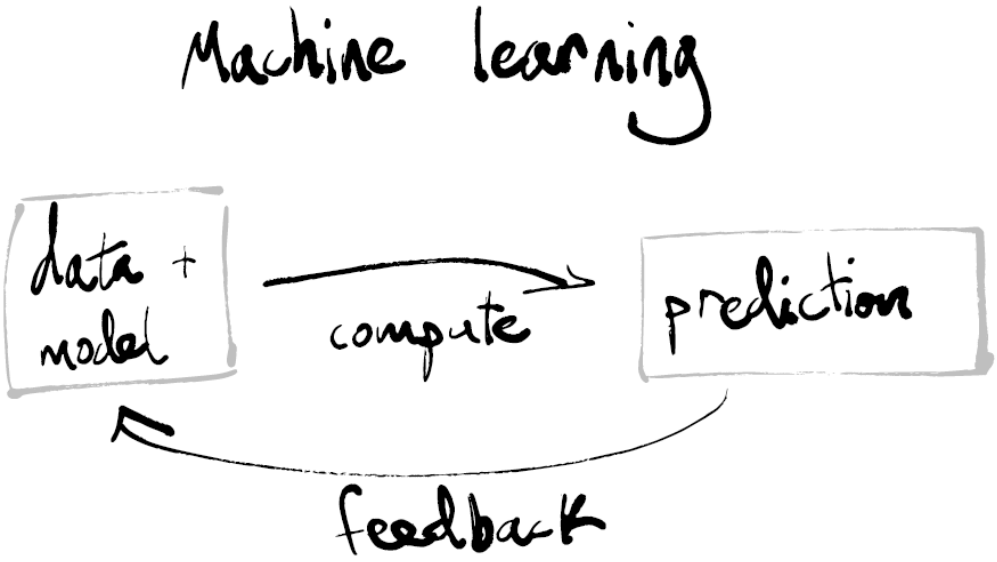

Figure 1.2 A simple diagram of machine learning

The speed with which data can be analysed in this way outstrips what humans can achieve on their own, and the in-built feedback mechanism means data can be refined independently of human oversight. That gives the process a life of its own, especially when the feedback mechanism finds patterns that the human initiators had not, and might not have, thought of.

These advances in capacity are mainly in terms of big data and big computation. Alongside this, however, there is additional advancement in method because machine learning can refine the explanatory model over the course of the investigation. This automates the traditional scientific approach of writing down a hypothesis and then looking at the data because in a sense the hypothesis can be constantly re-written while computation is evaluating the data and providing feedback on what model would best fit.[3]

In a controversial article entitled 'The End of Theory', Chris Anderson pushes the point still further by arguing that big data means we no longer need scientific models.[4] His view is that the amount of data now available qualitatively shifts what sorts of methodologies we should use to understand the world, and as such debunks the idea that we need to rely on the causal models scientists traditionally have in their minds when they carry out experiments. Instead, data itself can frame the research design and lead the way in identifying consistencies and inconsistencies to tell us directly about the way the world works. He explains:

Scientists are trained to recognize that correlation is not causation, that no conclusions should be drawn simply on the basis of correlation between X and Y (it could just be a coincidence). Instead, you must understand the underlying mechanisms that connect the two. Once you have a model, you can connect the data sets with confidence. Data without a model is just noise.

But faced with massive data, this approach to science—hypothesize, model, test—is becoming obsolete. […] Petabytes allow us to say: "Correlation is not enough." We can stop looking for models. We can analyze the data without hypotheses about what it might show. We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot.[5]

Of course, in terms of logic, Anderson's view that theory has therefore ended is, itself, a theory, which makes him part of the old guard that need to be transcended. But even when stepping aside this minor difficulty, we run into the problem that there are, clouded within his argument, two distinct ways in which big data can of its own accord be thought able to overcome the shortcomings of previous methods of analysis, and these two ways need to be unpicked.

The first is that big data enhances and accelerates capacities for identifying correlations. Via this route, through more data an analyst can be surprised at a correlation that was unexpected, like if he or she were to run a multilinear regression on the demographics that tend to watch YouTube videos of cats and find—surprisingly—that they are mostly watched by those in the financial sector. In a

[3] Udrescu, S. & Tegmark, M., 'AI Feynman: a Physics-Inspired Method for Symbolic Regression' (27 May 2019). arXiv:1905.11481.
[4] Anderson, C., 'The End of Theory: The Data Deluge Makes the Scientific Method Obsolete'. *Wired* (23 Jun 2008). https://www.wired.com/2008/06/pb-theory/.
[5] Ibid.

sense, the theory that the finance sector loves cat videos did not exist in the mind of the researcher prior to the observed correlation, but this does not amount to an elimination of the need for a model because, notwithstanding the surprise, such a model is latent in the choice of demographic factors included in the study, as well as in the more general assumption that correlations identified through the method likely indicate meaningful relationships. It may be that machine learning increases the number of parameters for consideration during the process of feedback and computation,[6] but this does not mean no model is needed at all, just that the model is being constantly refined. Continuous machine choice of what parameters or factors to include in the model is still based on and linked back to the original criteria of the research design—seeking explanation or optimisation of a particular dependent variable. That enduring link with the *explicandum* (what is to be explained), means there will always be a model, however buried.[7]

The second distinct way in which big data is considered a resource so comprehensive that it amounts to a new scientific method is in the way it allows a personalisation or individuation of prediction hitherto unprecedented.[8] In other words, big data means small modelling—so small, in fact, that to some it does not look like modelling at all. It is a bit like saying that because I know so much about you, I can predict when you will go for your morning coffee without the aid of any model suggesting what times people like you tend to go for their morning coffee. What is going on here? The idea is that, through big data, I know you inside out and no longer need to collate studies of other people like you to get at your daily routine.

Again, however, there is a simplification at play that does not do justice to the underlying scientific methods. The simplification is that because my prediction works, I no longer need reference the model, even though there was a model involved in the first place and there continues to be one as I go about making correct predictions. Dropping explicit reference to hypotheses when certain of a property's causal dynamics is nothing strange in science. If we know how the cell works, for example, we know how the cell works; by this we mean we know cells in reality, not just in our imagined models. We can drop reference to what is hypothesised about the cell in front of us and speak with confidence about what we know. Likewise, if it is the case that I did actually manage to know enough about you to be certain of when you go for your morning coffee (unlikely though that is, given the way big data is not yet that big at all[9]), it would be because my model of you corresponds to reality and is true, not because a hefty amount of data means I can dispense with

---

[6] Ghahramani, Z., 'Probabilistic machine learning and artificial intelligence'. *Nature*, Vol. 521 (2015), pp. 452-459, pp. 454-5.

[7] This point is applicable to the formation of research designs throughout the sciences, and is made here to demonstrate that machine learning likewise does not escape the need for human-instigated modelling. For further debate on this in biology over the application of systems theory, see Leyser, O. & Wiseman, H., 'Integrative Biology: Parts, Wholes, Levels and Systems'. Ch 2 of Reiss, M. J., Watts, F. & Wiseman, H. (eds.), *Rethinking Biology: Public Understandings* (New Jersey: World Scientific, 2019); Gatherer, D., 'Modelling versus Realisation: Rival Philosophies of Computational Theory in Systems Biology'. Ch 3 of Reiss, M. J., Watts, F. & Wiseman, H. (eds.), *Rethinking Biology: Public Understandings* (New Jersey: World Scientific, 2019).

[8] Ibid, p. 458.

[9] Graham, M., 'Big data and the end of theory?' *The Guardian* (9 Mar 2012). https://www.theguardian.com/news/datablog/2012/mar/09/big-data-theory.

modelling. Though I may cease to speak in hypothetical terms, I am still committed to a behavioural model—it is just that in this case the model happens to be right. When it is not right I talk about what I thought would be the case, which reveals once again the predictive model I had in my mind all along.

At the bottom of this debate is the unyielding fact that we get to know things with respect to what they do, and that understanding of what they do involves a model of what things tend to do. There is always an explicandum or dependent variable that needs to be explained via a model and data. The model can be changed radically and frequently, but dispensing with it altogether would amount to dispensing with knowledge itself.

Anderson's assessment of the way in which big data lets go of the need for human evaluation of causality through modelling is, therefore, extremely difficult to maintain. His conclusion begs more questions than it answers:

The new availability of huge amounts of data, along with the statistical tools to crunch these numbers, offers a whole new way of understanding the world. Correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all.

There's no reason to cling to our old ways. It's time to ask: What can science learn from Google?[10]

Despite the above reservations with this conclusion, let us suppose it is accurate and ask: what would a world where correlation supersedes causation look like? It would be one where all impetus is placed on collecting and codifying data rather than thinking through the meaning behind how data points relate to one another. Data would accumulate and the speed of computation advance, but the quest to make sense of the process—either in the form of predictions or models—would recede in importance.[11] Interest and respect would switch to those studies able to amass large datasets over those that develop single lines of causal inquiry.

That supposition has increasing relevance for society as a whole because the extent to which the search for truth needs human direction dictates our level of civic participation over the long-term. In rejecting Anderson's proposed approach of letting data correlations obviate theories of causation, one instead commits to the position that no matter the advances in big data and big computation,

---

[10] Anderson, 2008.

[11] Jonathan Zittrain makes the added point: 'Intellectual debt accrued through machine learning features risks beyond the ones created through old-style trial and error. Because most machine-leaning models cannot offer reasons for their ongoing judgments, there is no way to tell when they've misfired if one doesn't already have an independent judgment about the answers they provide.' Zittrain, J., 'The Hidden Costs of Automated Thinking'. *The New Yorker* (23 Jul 2019). https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking.

there will always be a space for goal-setting and identification of dependent variables and explicanda, a space that is uniquely human.

In line with this, it is worth at the outset affirming these following three home truths:

1. Models will remain directional for knowledge acquisition no matter the size and extent of big data and big computation.
2. The model we have of the human person is our human nature.
3. Debate over our human nature and what society *is* and *is for* will therefore be fundamental to our citizenship in a networked age.

c) Automation, algorithms & artificial intelligence

Algorithmic decision-making can be defined as optimised responses following pattern identification. It can be automated computationally and therefore executed at great speeds, but it essentially rests on these two core elements. All algorithmic decision-making involves some initial codification of data, and then an execution command based on the results of that data codification. It is an ongoing question whether this amounts to *intelligence*. Usually, commentators will begin by describing surprisingly effective examples of algorithmic decision-making or machine learning (such as AlphaZero playing chess), but that risks putting the cart before the horse. To answer definitively whether such processes count as intelligence requires starting with a definition of intelligence before specific examples are considered. Otherwise what is most interesting or humanlike in the examples one sees biases one's sense of what intelligence looks like. If intelligence is first defined in terms of being able to give optimised responses following pattern identification, then yes there is a thing called artificial intelligence that is becoming increasingly dominant. If, instead, there is something else to intelligence, one has to see whether machines are capable of it. If not, and humans are capable, one has to conclude that there is no such thing as artificial intelligence properly understood.

It is a regular concern that in the enthusiasm for technological progress, commentators anthropomorphise artificial intelligence, believing computers to be smart because they mimic human thought. However, an opposite problem can also come in *anthropomorphising intelligence*. Intelligence is not an exclusively human faculty, and there is no reason why many other things cannot be understood as intelligent.[12] There is a danger in confusing the question, "Is this intelligent?", with the closely related but distinct question, "Is this intelligent in the way humans are intelligent?" In this sense, the Turing Test does more to confuse than enlighten. The Turing Test is an

[12] See Templeton World Charity Foundation, 'Diverse Intelligences' (2018). https://www.templetonworldcharity.org/our-work/diverse-intelligences.

'Imitation Game'[13] of asking 'whether or not a computer is capable of thinking like a human being'.[14] This is a confusing anthropomorphising of intelligence because it is not clear whether one is testing for intelligence or testing for being humanlike. An especially intelligent human may not look and act like the average human and so might fail the Turing Test by failing to properly imitate what is thought average. In this sense, the Turing Test is more like a test of social congruence.

For the sake of understanding the strengths and weaknesses of automation, algorithms and artificial intelligence, we therefore have to temporarily put some distance between our understandings of human nature and intelligence and answer whether machines can be humanlike, distinct from answering whether machines can be intelligent. Of course, the two questions come back together when asking whether machines can be intelligent like humans are, but that assumes there is a uniquely human type of intelligence which first needs to be established.

Let us consider first a definition of intelligence. Intelligence is 'the ability to learn, understand, and make judgments or have opinions that are based on reason'.[15] It seems that under the general descriptions already outlined, machines are capable of this if, for the sake of argument, we further specify *understand and make judgments based on reason* in terms of *following what is logical*. The definition above of algorithmic decision-making as *optimised responses following pattern identification* does, in this sense, count as making judgments based on reason in that the outputs logically follow from the inputs.

Giving the benefit of the doubt, let us now turn to the more specific question of whether machines are intelligent like humans are, bearing in mind that humans are not necessarily the most intelligent beings around and so the fact that something is very intelligent does not make it humanlike necessarily.

Michael Jordan, of University of California, Berkeley, explains that those narrating the rise of artificial intelligence often compound two different things, and this has direct relevance for whether machines are becoming intelligent like humans are. One area of progress is the development of software and hardware that seeks to approach or copy human-level intelligence, what he calls 'human-imitative AI'. That was the original use of the captivating acronym A.I. as coined in the 1950s. The second is what he terms 'intelligence augmentation', something that has enjoyed great progress over the past two decades and is where 'computation and data are used to create services that augment human intelligence and creativity.'[16] Intelligence Augmentation flips the acronym to

---

[13] Alan Turing's original description of the Turing Test. Christian, B., *The Most Human Human: What Artificial Intelligence Teaches Us About Being Alive* (London: Penguin Books, 2011), p. 10.
[14] Rouse, M., 'Definition: Turing Test'. *TechTarget* (2019). https://searchenterpriseai.techtarget.com/definition/Turing-test.
[15] Cambridge Dictionary (Cambridge University Press, 2019). https://dictionary.cambridge.org/dictionary/english/intelligence.
[16] Jordan, M., 'Artificial Intelligence – The Revolution Hasn't Happened Yet'. *Medium* (19 Apr 2018). https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7.

IA, and is in some ways an opposite to AI. It puts human intelligence first, with computation used merely to accelerate the execution of human ideas and goals. Jordan points out that IA is very unhelpfully subsumed under the general heading of AI because an increased ability to execute human ideas often looks like growth in non-human intelligence, but it is not. Jordan explains the confusion:

While related academic fields such as operations research, statistics, pattern recognition, information theory and control theory already existed, and were often inspired by human intelligence (and animal intelligence), these fields were arguably focused on "low-level" signals and decisions. […] "AI" was meant to focus on something different—the "high-level" or "cognitive" capability of humans to "reason" and to "think." Sixty years later, however, high-level reasoning and thought remain elusive. The developments which are now being called "AI" arose mostly in the engineering fields associated with low-level pattern recognition and movement control, and in the field of statistics—the discipline focused on finding patterns in data and on making well-founded predictions, tests of hypotheses and decisions.[17]

In this way, optimisation or statistics researchers 'wake up to find themselves suddenly referred to as "AI researchers"'[18] even though their work is more IA than AI. Statistical methods have very little to do with the human-imitative AI that is doing so much to heighten expectation in future capabilities for general AI and machine-led decision-making,[19] but the general assumption is that they are the same thing. The distinction between IA and AI holds even though machine learning systems are able to make 'predictions which sometimes far exceed the capabilities of the top humans and competing computer programs in the world.'[20] It is not about processing speeds so much as the original reason why the calculations are being done in the first place that dictates whether the process is properly described AI or IA. If the computation is being done to automate goals that human creators originally set, it is IA. If the computation is to artificially recreate human behaviour or thought, it is AI. Success in human-imitative AI is neither sufficient nor necessary to accelerate progress in IA, and vice versa. After all, we already are humans, so why would creation of human-imitative AI augment our intelligence?

What about those who say machine learning is so advanced, or will be so advanced, that it produces outcomes that were not envisaged by the original human designers?[21] As should be clear from the preceding discussion, that framing already makes the mistake of *anthropomorphising intelligence* by taking human intelligence as the definition of what intelligence is such that going past that limit in any respect makes the machine potentially at least as competent in all respects. As human capacity is not a sovereign indicator of intelligence, going beyond what humans can do in one respect is

[17] Ibid.
[18] Ibid.
[19] On the creation of human-imitative AI that heightens expectations of general AI, see the work of anthropologist Beth Singler.
[20] Hawley, S. H., 'Theopolis Monk: Envisioning a Future of A.I. Public Service'. Ch 14 of Lee, N. (ed.), *The Transhumanism Handbook* (Cham: Springer, 2019), part 2.
[21] Bostrom, N., *Superintelligence: paths, dangers, strategies* (Oxford: Oxford University Press, 2014); Tegmark, M., *Life 3.0: being human in the age of artificial intelligence* (London: Allen Lane, 2017).

limited to describing advancement in that respect only. So the fact that human designers did not envisage the action is insufficient for determining whether the being is intelligent. Indeed, accidents are often unexpected and sometimes they even give the impression that the thing that caused the accident has a life of its own, but that says nothing about whether the thing is intelligent.

Joaquin Quiñonero Candela, Facebook's Director of Applied Machine Learning, gives the following rubric on how to enhance machine learning:

1. Get as much data as you can and make sure it is of highest quality.
2. Distill your data into signals that will be maximally predictive—a process called feature engineering.
3. Once you have the most awesome data and tools for feature engineering, keep raising the capacity of your algorithms.[22]

The key jump is in number two, where the researcher seeks to be 'maximally predictive'. It begs the question on what it is that the machine is, ultimately, trying to predict. Although that step is hidden here, it is essential for our question of whether technological advancements remove the need for human input. As Taina Bucher explains:

Feature engineering, or the process of extracting and selecting the most important features from the data, is arguably one of the most important aspects of machine learning. While feature extraction is usually performed manually, recent advances in deep learning now embed automatic feature engineering into the modeling process itself. If the algorithm operates on badly drawn features, the results will be poor, no matter how excellent the algorithm is.[23]

By omitting the human agency involved in selecting the thing that needs to be explained by step two, it is possible to give the impression that machines have the capacity to be self-learning and even self-determining for the whole process. However, the fact that human agency will always be present in step two—as a kind of limited first mover—means there will never be self-governing artificial intelligence in the sense of ultimate goals that are self-defined. While it is accurate to describe machines as often displaying a form of intelligence by making judgments based on reason (in terms of following what is logical according to the programming), there is a very different field of intelligence that humans have. This must be outlined and explored on its own terms if we are to fathom the true scope of the networked age's human-machine cooperation.

[22] Bucher, T., *If… Then: Algorithmic Power and Politics* (Oxford: Oxford University Press, 2018), p. 25.
[23] Ibid.

d) Human intention & civic ideals

Julie Cohen explains that '[i]nformation is never just information: even pattern identification is informed by values about what makes a pattern and why, and why the pattern in question is worth noting.'[24] The permanence of this feature in the human-machine learning process points to the long-term place of human values in our networked age. Rather than a force for regression, the inclusion of human thinking in the value part of our methodologies grants 'a real opportunity to conceive of something historically new—a human-centric engineering discipline.'[25]

Human decision-making is unique in the way it is able to order pursuit of the common good and find interpersonal consensus. Humans do so through *developing hierarchies of goods* while *retaining interest in the goods postponed*. Optimisation through algorithms and artificial intelligence instead involves best responses and best actions.

What does it mean to retain interest in goods postponed? It can be helpful to take a step back. Behind all of the successes of artificial intelligence, there is a clearly defined goal, in pursuit of which machine learning often does surprisingly well. The best examples are in games like chess or Go, though other examples abound. In all of these examples the defined end point—winning according to the rules, maximising explanatory power, or generating optimal best responses—is made clear and the programming is structured towards it. In contrast to this, humans' end point is death, which is not a goal and not something we structure ourselves towards. So, the end point and the goal are different in humans, but in machines they are the same.

When computers execute their programming they are either in a loop or travelling towards the end. Either way, they are following the programme's path, and the programme will terminate if that path finishes. Humans, however, can terminate at any time and they are usually aware of that fact. Importantly, their termination is almost always unrelated to their goals. Humans are not able to follow loops infinitely; even when they seem to be in a temporary loop, they are all the time changing, growing, ageing. Humans' way of being—their behaviour and evolutionary journey—is structured around not a single goal but a wide plurality of goals in the midst of the unknown of death. Death is not just unknown in the sense of not being sure what it is like and what it means, but also in the more straightforward sense of not tending to know when it will happen.

The good life is not, therefore, about maximising a single variable and then terminating but about pursuing enduring goals in the midst of a difficult unknown. Leading a good life thus involves something of a mastering of one's relationship with death—being aware of death and its significance but not letting it paralyse pursuit of one's mission, duties and goals. From a machine's perspective, humans live very strangely indeed in seeking goals that are neither part of a loop nor steps towards one's termination. From a human's perspective, machines live strangely in never enjoying the

[24] Cohen, J. E., 'What Privacy Is For'. *Harvard Law Review*, Vol. 126 (2013), pp. 1904-1933, pp. 1924-5.
[25] Jordan, 2018.

moment and always moving on to the next thing. Human life is about finding and pursuing the meaning of life despite knowing one's forthcoming death can come at any time and for seemingly unrelated reasons. Our reasoning process is therefore hard-wired in being able to navigate the incommensurable push-and-pull of seeking atemporal goods despite temporal limitations.

Indeed, even those who believe we should choose when we die through euthanasia do not argue that it is because we have reached perfect human fulfilment and so are now ready for termination. They say that it is when we are unable to reach human fulfilment that we should be allowed to be terminated. In reverse, they likewise prove that humans' termination is not the natural step following achievement of goals.

As stated, the end point and the goal are different in humans but in machines they are the same. This stark and unyielding difference also reveals the problem with applying a utility framework to trying to live well as a human, despite the fact that a utility framework is often found to be useful and appropriate when programming algorithmic decision-making. The big *problem and advantage* of utilitarianism is that it is an ethical framework with no sense of time. All utilitarian justifications are defeated by the question, *when?* Present suffering for future gain can always be a good decision if no limit is given on the likely length of future time. Time, as we know, is potentially infinite, which means a utilitarian calculation of the greatest good for the greatest number has no in-built *number*. Actual application of utility frameworks requires an arbitrary demarcation of the timeline we must keep to—something never provided by the utility framework itself but through outside narration of the decision-making scenario. In the same way as a utility framework is always placed on an already structured problem—with humans describing the temporal limits to the problem—so too are all plans of how to live a good life developed in conjunction with an assessment of the likely time of one's end. Humans then engage in a praxis of working out their achievement of mission, duties and goals, in the face of the incommensurable fact of death.

The human peculiarity of *pursuing goals that are hard to relate to our limitations* means our evolution has specialised in a *building-up of the ability to retain interest in the goods postponed*. We do not just have culture to help us with habits of solving collective action problems, we have culture to help us remember goods postponed. The tragedy of death places confusing limits on our pursuit of the meaning of life, and so we need to be strong in retaining memory of our ideas on that meaning. All of what we do to stay alive—our food, our shelter, our health care—are important and yet not exhaustive of that meaning. They are confusingly conjoined, necessary yet insufficient. In the midst of the to and fro of navigating our mortality we write books about the meaning of life, and build places of worship, and follow advice on how to be happy, and try to know more and more about truth, when we can, *as these are the goods postponed*.

Our terrible mix-up of latent goods and disjointed urgent goods means human valuation can never be fully reconciled with a process of machine execution. A machine can be perfectly flexible in changing goals but is not specialised in retaining interest in goods postponed. A machine has its priority dictate the best response, the best action. But humans develop hierarchies of goods that specialise in keeping sight of goods we cannot currently pursue.

In community, and as a citizenry, we work through this unique method of reasoning. It is here to stay.

e. Our human space, going forward

We need to find ideals of citizenship that work for value-based decision-making in our networked age. As the earlier discussion made clear, the moral dimension to human decision-making will continue to be directional for the shaping of digital technologies, despite enthusiasm among some tech utopians to the contrary. Agreeing on and establishing ideals for civic engagement in our networked age will therefore be fundamental for healing democratic society. As Corrine Cath and colleagues write:

We are creating the digital world in which future generations will spend most of their time. [T]he design of a "good AI society" should be based on the holistic respect (i.e., a respect that considers the whole context of human flourishing) and nurturing of human dignity as the grounding foundation of a better world. The best future of a "good AI society" is one in which it helps the infosphere and the biosphere to prosper together.[26]

But before providing an account of *democratic society* we have to provide an account of *society*—the way we are bound together, our human space, going forward. Citizenship, as *a normative fact about who we are able to be when we come together politically*, connects the basic matter of society with our account of democratic society. It is a mixed normative and empirical endeavour, saying something both about who we are (empirical) and the kind of people we want to be (normative). Citizenship is hard to build up and hard to break down; something of a habit of the heart, resiliently playing out generation after generation.[27]

---

[26] Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M. & Floridi, L., 'Artificial Intelligence and the "Good Society": the US, EU, and UK approach'. *Science and Engineering Ethics*, Vol. 24, No. 2 (2018), pp. 505-528, p. 508.

[27] Julie Cohen remarks, 'Well-functioning state and market institutions cannot be built in the span of a grant-funded research project or a military campaign. Their rhythms and norms must be learned and then internalized, bringing into being the habits of mind and behavior that democratic citizenship requires.' Cohen, J. E., 'What Privacy Is For'. *Harvard Law Review*, Vol. 126 (2013), pp. 1904-1933, p. 1912. See also Putnam, R. D., *Making Democracy Work* (Princeton: Princeton University Press, 1993); Putnam, R. D., *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon & Schuster, 2000); Uslaner, E. M., *The Moral Foundations of Trust* (Cambridge: Cambridge University Press, 2002).

Our durable norms of togetherness are under attack. Who we are was broken up by an economics of individualism long before digital technologies came into existence,[28] but those technologies have capitalised on the atomisation of cultures and societies through rapid personalisation of technology's use and purpose. These technologies are not, therefore, optimising formation of common goals but, instead, optimising an individuation of goals. To some extent that was always the way with tools—they facilitate humanity's greater specialisation of activities over time. Nevertheless, there is a difference here in that the extent of personalisation rebukes the assumed togetherness of social living, bringing us away from questions of the economic effects of digital technologies and towards a relatively newer topic of its social and political effects. It may be that the personalisation of human experience through digital technologies is such that we no longer have the *common ground* that can work as the basis to a *common good*.

Through advances in augmented reality, for example, we will literally be seeing different worlds depending on how good one's phone or glasses are—the world will appear differently to each person in accordance with how much they can afford these technologies. And seeing is believing. The dynamic is already at play more broadly in inequalities of data gathering, whereby the people around us effectively look different depending on our level of access to their data profiles. In such a scenario, people are passive data generators, and organisations are the active users of that data for predicting and generating future trends.[29] Legal scholar Julie Cohen has coined the term "modulated society" to describe a world in which individuals are fed choices that suit their comfort level, rather like setting a thermostat to a preferred temperature. She describes a move from a liberal to a modulated society through increased surveillance and data-gathering techniques:

Citizens of the modulated society are not the same citizens that the liberal democratic political tradition assumes, nor do their modulated preferences even approximately resemble the independent decisions, formed through robust and open debate, that liberal democracy requires to sustain and perfect itself. The modulated society is the consummate social and intellectual rheostat, continually adjusting the information environment to each individual's comfort level. Liberal democratic citizenship requires a certain amount of *dis*comfort—enough to motivate citizens to pursue improvements in the realization of political and social ideals. The modulated citizenry lacks the wherewithal and perhaps even the desire to practice this sort of citizenship.[30]

It may be that one day humanity will conform to Cohen's image of citizens modulated beyond all ability to engage in self-direction, but we are not there yet. If it were true, there would be no value in writing and reading this paper; it is premised on the assumption that good ideas can help make a good citizenry.[31] The liberal democratic tradition has always worked with a view of citizenship

[28] Manent, P., *A World beyond Politics? A Defense of the Nation-State* (Princeton: Princeton University Press, 2006); Elshtain, J. B., *Sovereignty: God, State, and Self* (New York: Basic Books, 2008); Siedentop, L., *Inventing the Individual: The Origins of Western Liberalism* (London: Penguin Books, 2015); Deneen, P. J., *Why Liberalism Failed* (New Haven: Yale University Press, 2018).

[29] Zuboff, S., *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power* (London: Profile Books, 2019).

[30] Cohen, 2013, p. 1918 (emphasis in original).

[31] Cohen describes the fear that '[s]timuli tailored to consumptive preferences crowd out other ways in which preferences and self-knowledge might be

slightly more morally competent than what we tend to find among real citizens; in that subtle idealism we give effect to our aspiration of a freer world.[32] Admittedly, this is citizenship as *a normative fact about who we are able to be when we come together politically*, breaking the boundaries of the fact/value distinction set by David Hume and Immanuel Kant in the 18th century.

What we value as citizens need not necessarily be in conflict with what we are empirically. It is not hard to see that as humans we have become utterly dependent on mutual social commitment: just imagine trying to make your own ball point pen, let alone your own smart phone. Much of this cultural dependence is longitudinal, depending on education to pass what is learned from one generation to the next. There is growing evidence that this kind of cultural evolution has occurred hand in hand with genetic evolution, to provide humans with the hardware in the form of brains and larynxes that facilitate cooperation.[33] Insofar as we believe that cooperation and education are good things for citizens to promote, there is a coincidence between what is and what should be. But it is not automatic.

Just as our notion of citizenship does not fully rely on the facts of who we are, but also on the normativity of who we want to be, so the changes brought through digital technologies tell only part of the story of the kind of humanity we will likely become. Our civic ideals are thus *empirically predictive* of many aspects of the society we will go on to inhabit, because while their realisation is frequently frustrated, they endure as an important causal factor for what gives us common direction for our human space.[34]

We turn now to two main accounts as to how ideals have been generated. These two accounts are not the only ones available for rebuilding our civic ideals for the networked age, but they provide a helpful guide on the richness of ethical and political thinking that so far forms the basis to current understandings of citizenship's normativity. The first is the citizen-slave distinction, and the second the ethics of navigating problems of scarcity and competing desires. These are two alternative routes for establishing norms and rules for what it means to be a person-in-community, a citizen. Both hold relevance for rebuilding our civic ideals for the networked age, and therefore the account that follows draws from both literatures. The concluding argument is that both accounts indicate

---

expressed, and also crowd out other kinds of motivators—altruism, empathy, and so on—that might spur innovation in different directions.' Cohen, 2013, p. 1926. While that is a relevant concern, if it really takes place there would be little point in writing about it. Academic writing is, ultimately, a project in social improvement through greater self-knowledge.

[32] A good example is Garton Ash, T., *Free World: Why a Crisis of the West Reveals the Opportunity of Our Time* (London: Penguin Books, 2005).

[33] Henrich, J., *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (Princeton: Princeton University Press, 2016); Christakis, N. A., *Blueprint: The Evolutionary Origins of a Good Society* (New York: Little, Brown Spark, 2019).

[34] See, for example, the role played by particular countries' tolerance for 'creative destruction' in their economic and political transformation over time, as accounted for in Acemoglu, D. & Robinson, J. A., *Why Nations Fail: The Origins of Power, Prosperity and Poverty* (London: Profile Books, 2012).

how the empirically-grounded problems of the networked age can, in fact, provide the material for an opposite rebuilding of our civic ideals.

f. The citizen vs slave distinction

Citizenship has both legal and normative dimensions; the focus here is on the normative: what is a good citizen, and what does a good citizen do? Defined as *a normative fact about who we are able to be when we come together politically*, citizenship is being looked at from the perspective of the person-in-community—social by nature, finding particular fulfilment in joint human efforts and partnerships.[35] In terms of a normative aspiration, there is an ideal citizen contribution to the life of the body politic that can be imagined and realised in degrees. In the liberal democratic order it tends to include a spirit of public service, justice, neighbourliness, democratic participation and moral reasoning. We often fall short in trying to achieve these elements in their fullness, and yet they continue to guide us, and even inform us on who is failing to contribute. In their achievement, we form an active and well-connected community in pursuit of a common good.

The normative dimension to citizenship was in part generated through 18th century reflection on the differences between citizens and slaves. The basic argument from Jean-Jacques Rousseau is that in order to participate fully in the life of the body politic, we must throw off the chains of our slave-like dependency on rulers and their institutions.[36] For Rousseau, this did not only apply as a criticism of the aristocratic and royal elites but to the church too—whose hierarchy, symbolism and narrative of the afterlife kept people subjugated. A coming of age for the citizen was about breaking free from these chains: 'What makes the work of legislation difficult is not so much what has to be established as what has to be destroyed'.[37] Rousseau held there to be something natural in achieving harmony between state and society, such that governance was about realising that citizenship through release of the general will of citizens.[38] The cunning of his thought came in establishing a connection between the freedom of citizens and the goodness of citizens. A flourishing society is less about producing good people as having a state that realises peoples' natural goodness through harmony with their self-direction.

---

[35] Aristotle, *The Politics* (London: Penguin, 1992); Aristotle, *The Nicomachean Ethics* (London: Penguin, 2004).

[36] Rousseau, J. J., *The Social Contract* (Cambridge: Cambridge University Press, 1997 [1762]).

[37] Ibid, p. 78.

[38] 'What makes the constitution of a State genuinely solid and lasting is when what is appropriate is so well attended to that natural relations and the laws always agree on the same points, and the latter as it were only secure, accompany and rectify the former. But if the Lawgiver mistakes his object, if he adopts a principle different from that which arises from the nature of things, if one principle tends toward servitude while the other tends toward freedom, one toward wealth, the other toward population, one toward peace, the other toward conquests, then the laws will be found imperceptibly to weaken, the constitution to deteriorate, and the State will not be free of turmoil until it is either destroyed or altered, and invincible nature has resumed its empire.' Ibid, pp. 79-80.

There are, of course, many themes here that carry over into liberal democratic appreciation of rights and freedoms as at the same time both a property of individuals and essential minimums for well-functioning states. On this reading, the liberal project is thoroughly naturalistic because it takes an a-historical account of human nature and asks that laws be harmonious with it. Such an approach was the source of political optimism for our networked age: the freedom afforded by the internet was first thought something of a state-of-nature experiment, for which our regulatory institutions would be wrong to impose out-of-touch rules. While that political optimism for the internet lasted for a time, two serious threats to the narrative arose. The first was that some online activities clash with human rights, which put the naturalistic argument in a difficult position of needing to choose between protecting what it holds to be, ultimately, "natural" rights or, instead, protecting the boundaries of a state-of-nature territory. The second was that the controllers of the internet have formed a new elite, setting its rules and providing corporate conditioning to all that goes on. This new elitism excites a naturalistic alternative in the form of the "hacker", who can "move fast and breaks things"—an adage that Mark Zuckerberg had to abandon once Facebook's size meant it had to take greater responsibility for all it was doing.[39] The "hacker" can be both an ordinary programmer (as in "hackathons") or an anti-system rebel (as in "Anonymous"). Either way, the sense of overall contribution comes in the belief that freedom naturally leads to good things, even if socially and politically disruptive—a logical step that can be traced back to Rousseau's connection between the freedom and goodness of citizens.

While there are positives at play in Rousseau's account in terms of equality (or at least equivalence) between citizens, it is important to keep in mind that his is ultimately a rejection of the perceived slave-like status quo, followed by appeals to less-defined naturalism. He provides little in terms of how one would purposefully restructure the relationship between citizens and authorities in the event the natural is not so good.

The Roman foundations of Western vocabulary and interpretations of citizenship also made general appeal to fairness, but it was in terms of a fairness in receiving contributions from the wealthy, for which non-citizens should not be entitled. The way in which Christians dedicated themselves to almsgiving towards the non-citizen poor put them at odds with Roman ideas of citizenship, contributing to the sense that they must be to blame for the fall of Rome.[40] A "citizen" is a term wedded to that of "city" and yet the Christian idea of almsgiving worked against the idea that the rich should give back to their home cities and citizens, arguing that God wanted giving to those least able to give back, and regardless of where they were. The provision of infrastructure and public entertainment from the wealthy to the citizens of one's native area gradually gave way, therefore, to monasteries offering hospitality, health care and spiritual direction, regardless of citizenship. While a causal connection between Christianity and the decline of the Roman Empire is difficult to maintain,[41] these changes did result in a rupture between normative and legal understandings of

[39] Taneja, H., 'The Era of "Move Fast and Break Things" Is Over'. *Harvard Business Review* (22 Jan 2019). https://hbr.org/2019/01/the-era-of-move-fast-and-break-things-is-over.

[40] Gianakon, S. E., 'Citizenship and the Holy in Late Antiquity'. Humboldt Universität, Fulbright essay (May 2018).

[41] Augustine, *City of God* (London: Oxford University Press, 1963 [426]).

Roman citizenship, with the legal retaining much of its sense of exclusivity and the normative instead giving up ground to a competing notion of universal human dignity.

Roman citizenship is therefore a privilege distinct from slavery, but Rousseau's citizenship is instead non-privileged and natural. They both look the same in that they can both be contrasted with slavery, but in fact Rousseau makes a clear break with Roman tradition by removing the category of members of the public who are neither slaves nor citizens. For Rousseau, everyone is either a slave or a citizen: the former by malicious design, the latter by natural right. Any sub-groupings of citizens through associations is, for Rousseau, damaging to the strength of the body politic.[42] To help explain his point in *The Social Contract*, Rousseau writes the following footnote to the word 'city':

The true sense of this word is almost entirely effaced among the moderns; most take a city for a City, and a bourgeois for a Citizen. They do not know that houses make the city but Citizens make the City. […] Only the French assume the name *Citizen* casually, because they have no genuine idea of it, as can be seen in their Dictionaries; otherwise they would be committing the crime of Lese-Majesty in usurping it: for them this name expresses a virtue and not a right.[43]

Such is the naturalism of Rousseau's position that he scorns the idea of citizenship as a virtue (to be attained) and instead asserts it as a right (arising from nature).

Citizenship as a normative status that compels itself from nature eventually translated into the civil rights movements of the 20th century.[44] Citizenship becomes inextricable from themes of social solidarity and the refusal to be subject to arbitrary authority.[45] In its democratic form, it means dynamic engagement with the decision-making that directs the people as a whole. All this is predicated on the naturalism of Rousseau: citizenship in its normative sense is inclusive of universalistic conceptions of human dignity, with citizens demanding equal say in the way society is directed by natural right.

Those who defended slavery sometimes tried to present it as natural.[46] Citizenship is, of course, easy to imagine as legally constructed and therefore not so natural after all, by way of contrast. The contention—largely dismissed now—argued that slavery is natural, by virtue of some people always being able to dominate others, and citizenship is unnatural, in that it is an attempted legal affirmation of equality which is not there in nature. The naturalism for citizenship as *a normative*

---

[42] Rousseau, 1997, p. 60.
[43] Ibid, p. 51 (emphasis in original).
[44] See, for example, Allen, D. S., *Talking to Strangers: Anxieties of Citizenship since Brown v. Board of Education* (Chicago: The University of Chicago Press, 2004); Nash, K., *Contemporary Political Sociology: Globalization, Politics, and Power* (Chichester: Wiley-Blackwell, 2010), 2nd Ed., pp. 131-2.
[45] Pettit, P., *Republicanism: A Theory of Freedom and Government* (Oxford: Oxford University Press, 1997).
[46] BBC, 'Ethics guide: Philosophers justifying slavery' (2014). http://www.bbc.co.uk/ethics/slavery/ethics/philosophers_1.shtml.

fact about who we are able to be when we come together politically won the day, and the counter-argument that slavery is evident in history and therefore natural is largely confined to the classroom as a devil's advocate position.

While it may seem that the debate on the citizen-slave distinction is dead, this paper brings it up afresh because citizenship is now itself pitted as unnatural, as compared to the natural progress of technology and artificial intelligence. Put simply, technology seems to have a *natural growth*, which outpaces in functionality human crafts, now including the most human of all crafts: democracy.

The rapid transformation of our world through technology, artificial intelligence and algorithmic decision-making offers an altogether different proposition: decision-making is best conducted by artificial intelligence, which means any natural right for civic engagement should be wilfully foregone. Big data and algorithmic decision-making are the "new natural". They lift off from strong economic growth and solve many of the reasons for human discoordination, in turn removing the need for democratic deliberation. On this reading, "artificial" intelligence is "artificial" from the perspective of it being non-human, but it is natural intelligence from the perspective of the natural growth of the economy and civilization. The civic contribution comes, instead, in ethical monitoring and evaluation of the problem-solving progress of artificial intelligence.

From the point of view of the citizen-slave distinction, we are entering unchartered territory. It looks to some that increased reliance on, and guidance from, artificial intelligence amounts to a return to slave-like dependency. It looks to others that the optimisations achieved through computer-based techniques are finally able to help us realise the *general will* (*volonté générale*) that Rousseau deemed most true to our nature when all arbitrary restrictions are cast off.

Here the citizen vs slave distinction falls short. It has been useful in helping see what is at stake and how our current sense of citizenship as a natural, normative fact has accrued, but the distinction does not provide clear navigation on whether or not an even higher type of citizenship exists out there, and whether the technological improvements we are currently witnessing amount to completion of our social nature through *participatory pursuit of the computerised optimal*, or whether they are in fact a reduction of us to slave-like conditions. Lingering concerns encourage us to turn to alternative ways of looking at the normative basis to our citizenship.

g. Ethics through the navigation of competing desires

Adrian Weller argues that artificial intelligence systems need to be developed with three principles in mind if they are to perform well and in line with human flourishing:

1. Transparency: making interpretable the reasons for artificial intelligence's predictions or decisions.
2. Reliability: safely scaling probabilistic reasoning to unforeseen settings.
3. Trustworthiness: ensuring artificial intelligence can reliably initiate acts of kindness through the proper inferring of the beliefs and goals of other agents.[47]

These are good ways to regulate artificial intelligence and avoid it becoming an unknowable, damaging force, but they do not show us good ways of being citizens. It may be that good citizenship is simply no longer necessary, but as pointed out, all data-driven analysis and execution relies at some level on human-generated purpose. If, in a democratic society, we view human-generated purpose as best realised collectively and in mutual trust, the way we organise that collectivity and mutual trust is of permanent relevance for the shaping of our future society. Our mutual trust—our oneness—is dependent on our citizenship, that normative fact about who we are able to be when we come together politically. Point (3) of Weller's prescription is therefore in danger of clouding the solution, for it asks for trustworthy artificial intelligence, assuming that it can be obtained through alignment with human goals. But not all human goals are trustworthy, and so artificial intelligence trained to align with them may prove destructive. The Leverhulme Centre for the Future of Intelligence gives further description of what is to be understood by 'trustworthiness':

Human studies indicate that a theory of mind may be essential to build empathetic trust, and for reliable initiation of acts of kindness. Equipping AIs to infer beliefs and goals of other agents (such as humans) may improve human-machine collaborations; yet such cognitive insight may prove a double-edged sword, allowing deception and even manipulation.[48]

What is meant here is that there is a danger that through learning a theory of the mind, artificial intelligence may gain the ability to deceive and manipulate others. It is a fair warning, but the irreducibly interactive—perhaps even co-evolving—relationship between humans and machines makes the ethical dilemma more two-way than even this double-edged sword suggests. Trustworthy artificial intelligence is still untrustworthy if it pursues the beliefs and goals of untrustworthy agents.

The level of our human-to-human trustworthiness dictates our overall trustworthiness in using tools to help us achieve our goals. In other words, the trustworthiness of our tools or machines are only of value insofar as we are trustworthy users and can trust each other so to be. Nuclear power, for example, can be used for energy or for weapons—the ethics of who possesses nuclear power is therefore additionally a question of human trustworthiness in using something for a good purpose, not simply a question of how trustworthy nuclear power is itself as a resource. In similar manner, the

[47] Weller, A., 'Where is AI going, and how will it promote flourishing?' Presentation at the conference 'Science, Philosophy, Religion & Human Flourishing', Ayia Napa, Cyrpus, (Nov 2018). See also Leverhulme Centre for the Future of Intelligence, 'Trust and Transparency' (2019a). http://www.lcfi.ac.uk/projects/ai-trust-and-society/trust-and-transparency/.

[48] Leverhulme Centre for the Future of Intelligence, 2019a.

extent to which we as citizens can pursue common goals ethically carries over into the likely use we will make of new opportunities and new technologies. It may be that an army is not ready to pass through a land of gold, for instance, if it means it will likely destroy the local population and steal the gold. Another army may be reliable enough ethically to do so (an assessment that is separate from the question of whether gold itself is good or bad). Evaluating our societal trustworthiness for *participatory pursuit of the computerised optimal* is therefore of the essence, and forces the question of what kind of moral rules and norms we require from ourselves as the opportunities afforded by these new technologies develop. Are we up to the mark? Will we exercise civic virtue? What strategies for helping ensure coordination and the pursuit of common goals will be required? What ideals of citizenship must be rebuilt for the networked age?

A helpful resource for answering these questions can be found in the rich debate on ethics, normative heuristics and incentives for solving collective action problems. The philosopher Alasdair MacIntyre states, 'From Hobbes onward the psychological problem had been posed, Why should men do other than act to their own immediate advantage?'[49] Answers to this question are evident not only among philosophers but also in the many ways in which humans have gone about solving collective action problems practically: founding institutions and setting rules of engagement to establish boundaries on what is allowed, regularising incentives and punishments to encourage certain types of conduct and dissuade others, nurturing cultures and customs that narrate and explain the link between good actions and good outcomes.[50] While it is true, however, that institutions and cultures in this sense work, they are nevertheless fragile and often fall short in achieving what is required. Human civilization is an almost constant process of rule-setting and rule-evaluation in tandem with critical reflection of our overall purpose and goal. Because the networked age challenges the very constitution of our social fabric, it upsets these methods of rule-setting and rule-evaluation and asks us to think afresh on the kind of collectivity that we are pursuing the good for. The networked age revolutionises our way of being and way of belonging, requiring us to look deeper into the foundations of our ability to engage in institutional design for the common good. That is, ultimately, a requirement to look into the way we generate our ethical systems, a sort of meta-ethics on how we become the kind of people that produce strong ethical systems over time and in new settings—in this case generating ethical systems able to withstand rupture to our ontology as a networked community.

With this in mind, it is possible to source discussion of our civic ideals not just through a citizen vs slave distinction but also through taking a fresh look at the inevitable tensions in ethics, namely, the way we value things as good for our flourishing. Broadly speaking, there are five reasons humans value things:

1. They pertain to our basic appetites (food, drink, shelter).

[49] MacIntyre, A., *A Short History of Ethics: A history of moral philosophy from the Homeric Age to the twentieth century* (London: Routledge, 1966), p. 179.
[50] On institutional design as a means for solving collective action problems, see Ostrom, E., *Understanding Institutional Diversity* (Princeton: Princeton University Press, 2005).

2. They establish our status (gold watches, tattoos, academic titles).
3. They improve us and those around us (exercise, advice, giving, justice).
4. They are enjoyable in themselves (play, art, contemplation).
5. They lead us as a journey to higher goods (religion, study, travelling).

All these can be subsumed under the more general reason that we seek to be happy,[51] but the differentiation here is nevertheless useful for showing that there can be tensions between the choices we make, a fact too easily dismissed when the goods we seek are thought reducible to a single utility framework.[52]

Economic and technological developments narrow the distance in obtaining these things that we value. They do not, however, help us prioritise what of them should be most valued, nor do they provide a sounder basis for our method of collective self-direction. They often help with our speed of communication and the certainty of our memory. They do not establish a hierarchy of values, but rather respond to the hierarchy of values we set.

Satisfaction of our basic appetites can often provide a natural hierarchy of valuation (today my thirst is more urgent than my hunger, but my hunger more urgent than my need for shelter). Greater productivity and automation means, however, that the satisfaction of basic appetites can in some parts of the world increasingly be taken as a given. This places relatively greater emphasis on the other four forms of valuation in determining a hierarchy to order individual lives, the market economy, and the common good. The four additional reasons for valuing things are, however, potentially never-ending. The shift of production towards them increases the anxiety in deciding among their relative importance. Capitalism becomes, in this sense, an engine of anxiety-creation over our valuation priorities by making choices and options multiply. It becomes harder to be sure of how our choices fit an authentic overall pursuit of happiness. Objective valuation of things as "needs" becomes less and less easy to assume as the economy specialises its marketing towards what we, as individuals, are likely to mistake as needs given our particular habits, tastes, experiences and psychologies. Personalisation of data-gathering techniques accelerates the project of making wants look like needs and the idea of a "need" turns into "fear of what you are in danger of missing out on" if you make the wrong consumption choices (see also Section 3.b).[53] Individual notions of status, improvement and enjoyment encourage ever more personalisation of the laws of supply and demand, escalating the pressure on, and evolutionary relevance of, each person's choices and valuations, and deescalating interpersonal moral correction and advice.

[51] Aristotle, 2004.

[52] As in Bentham, J., *An Introduction to the Principles of Morals and Legislation* (London: W. Pickering, 1823).

[53] In addition, see interesting discussion by Cal Newport about how social media platforms depend on a notion of "missing-out" rather than "need" in soliciting permanent participation. Newport, C., *Deep Work: Rules for Focused Success in a Distracted World* (London: Piatkus, 2016), pp. 184-8.

And yet in these burgeoning choices over status, improvement, enjoyment and journeying we are able to grow our moral reasoning, which is our ability to establish a hierarchy of valuations. As Sir John Templeton wrote:

Making choices has often been considered a means of human evolution. Each choice we make stems from our perspectives or intentions and from the quality of consciousness that we bring to our thoughts, feelings, and actions. Conscious evolution, through making responsible and positive choices, can be a beneficial path.[54]

Amartya Sen concurs, arguing that it is in the ranking of our preferences that we express our moral judgments.[55]

How do we build our moral reasoning, and how can it be built at the level of a networked community? MacIntyre provides a suggestion in his most recent book *Ethics in the Conflicts of Modernity* by pointing to the way in which competing desires act as a starting point for ethical reflection by forcing a person to attempt a prioritisation of certain types of goods over others: 'What small children desire they try to get. But […] as they grow older they learn to delay satisfying some of their desires and develop desires that can be satisfied only at some time, even some distant time, in the future.'[56] The original "marshmallow test" was designed to evaluate the change with age in children's capacity for deferring gratification.[57] Over the course of our lives in which 'objects of desire have multiplied' we build up a kind of history of desires, some of which are 'transformed, others replaced.'[58] At certain times, when ordinary life is radically disrupted,

it requires little reflection to recognize that if I am to answer the question "What shall I do?" I had better first pause and pose the question "What is it that I want?" Somewhat more reflection is needed to recognize that I also need to think critically about my present desires, to ask "Is what I now want what I want myself to want?" and "Do I have sufficiently good reasons to want what I now want?" and still further reflection to recognize that I will be likely to go astray in answering these questions if I do not also ask how I came to be the kind of person that I now am, with the desires that I now have, that is, to ask about the history of my desires.[59]

---

[54] Templeton, J., *Wisdom from World Religions: Pathways toward Heaven on Earth* (Philadelphia: Templeton Foundation Press, 2002), p. 261.

[55] Sen, A., 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory'. *Philosophy & Public Affairs*, Vol. 6, No. 4 (1977), pp. 317-344, p. 337. See also Murphy, J. B., 'Nature, Custom, and Reason as the Explanatory and Practical Principles of Aristotelian Political Science'. *The Review of Politics*, Vol. 64, No. 3 (2002), pp. 469-495.

[56] MacIntyre, A., *Ethics in the Conflicts of Modernity: An Essay on Desire, Practical Reasoning, and Narrative* (Cambridge: Cambridge University Press, 2016), p. 5.

[57] Mischel, W. & Ebbesen, E. B., 'Attention in Delay of Gratification'. *Journal of Personality and Social Psychology*, Vol. 16, No. 2 (1970), pp. 329-337.

[58] MacIntyre, 2016, p. 3.

[59] Ibid, p. 4.

There are ways in which this process of critical self-reflection over our deepest desires are being altered in the networked age. We often receive direct communicative feedback about our desires in a way we did not used to (e.g. if I were to post a picture of a crocodile-skin jacket on Instagram and write, "Thinking of buying one…"). Humans have always received feedback about their intentions, but the feedback is now possible from many more people simultaneously, like speaking in the town hall with everyone possibly interested but possibly not. At times the feedback is more cursory in nature because it requires less time per engagement. All this changes the speed and type of communication in forming our desires, though does not supplant those desires nor necessarily alter our underlying human values.

The *history of desires* is also changing in our networked age in that our desires are being memorialised more accurately through technology. This is most apparent among what Shannon Vallor describes as 'devotees of the Quantified Self', who 'employ mobile, wearable, and/or biometric sensors such as the FitBit and Jawbone devices, smartphone apps such as Moves and Chronos, video cameras, and a range of other devices to measure, track, analyse, and store volumes of recorded data concerning an ever-expanding list of personal variables.'[60] The idea is that one can collect streams of data about oneself and then take ownership of that data in using it to help make better choices. In her book *Technology and the Virtues*, Vallor debates whether this amounts to a kind of moral self-cultivation, under the idea that the good life is, in part, an examined life, a point agreed by many philosophers, theologians and ethicists. Can a collection of data about oneself through apps help offset the danger that 'the unexamined life is not worth living'[61]? Vallor's view is that it cannot, unfortunately. While she is sympathetic to the attempt, she believes that attention towards moral goods is about focusing-in on those goods through reducing the noise of everything else going on. 'As any philosopher of perception or cognitive scientist knows,' she writes, 'attention is as much about the ability to *screen out* information as it is about taking it in; in fact, the former capacity enables the latter.'[62] If I use an app to count my steps, but my number of steps does not usefully correlate with a life well lived, in what way does that data help my moral development? Could it even distract me from asking the bigger questions of life?

A number of ethics apps try to explicitly improve one's moral development. Evan Selinger and Thomas Seager explain that most are about ethical advice for particular purchases (e.g. whether a product is environmentally friendly), or else seek to direct one's behaviour towards a sense of the good through a combination of nudging, quantification and gamification.[63] An interesting example is the app 'GPS for the Soul', which tracks one's heart rate as an indicator of stress level. When stress is

---

[60] Vallor, S., *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (New York: Oxford University Press, 2016), pp. 198-9, 201.
[61] Plato, *Apology*, 38a5. In Plato, *Complete Works* (Indianapolis: Hackett Publishing Company, 1997), edited by John M. Cooper, p. 33.
[62] Vallor, 2016, p. 200 (emphasis in original).
[63] Selinger, E. & Seager, T. P., 'Digital Jiminy Crickets: Do apps that promote ethical behaviour diminish our ability to make just decisions?' *Slate* (13 Jul 2012). https://slate.com/technology/2012/07/ethical-decision-making-apps-damage-our-ability-to-make-moral-choices.html.

too high, the app will 'connect you to whatever you need to get to a place of balance'[64]—the images, music and files that help you recentre. As the founder argues, 'the solution to the problems created by technology isn't anti-technology, but more and better technology.'[65] The app seeks to relativize the surrounding noise that is negatively affecting one's peace of mind. It does not aim at sparking critical self-reflection, which might of course make the person feel worse about themselves.

What, then, if I am a serial killer with a list of enemies, some of whom I have already knocked off and some of whom, unfortunately, are still around. When I get stressed, and my heart rate goes too high, I open my app to bring me back to my list of those I have already eliminated, perhaps with tranquil music playing in the background, and it helps restore my peace of mind.

MacIntyre's approach has very little to do with obtaining peace of mind: tension over competing desires and our yearning for a more coherent narrative in our history of desires provides the basis for critical moral reflection, *precisely because it is uncomfortable*. The issue is not whether there is more noise or less noise, but being willing to ask if I have sufficiently good reasons to want what I currently seem to want. For the serial killer, this is not about accumulating data but about examining the reasons given for wanting such data in the first place—critical self-reflection over the personal development that led up to this point.

Just as personal ethics can, in this way, be built through navigating one's competing internal desires, so too can civic morality be built through navigating competing social desires. While clarity over the nature of politics is something one may find hard to draw out of MacIntyre's own work,[66] it is easy enough to make the case more generally that conflict over desires happens at the societal level too, and that such tension and interpersonal competition provides an important justification for institutional rule.[67] Ultimately, institutions 'are the humanly devised constraints that structure political, economic and social interaction.'[68] They provide these constraints as a way of managing at a more macro level our competing desires, directing our energies towards common goals, common goods.

---

[64] Huffington, A., 'GPS for the Soul: A Killer App for Better Living'. *HuffPost* (16 Apr 2012). https://www.huffpost.com/entry/gps-for-the-soul_b_1427290?guccounter=1.

[65] Ibid.

[66] See the criticism of how MacIntyre's ethics fails to relate sufficiently to a theory of politics in Duff, A., 'The Problem of Rule in MacIntyre's Politics and *Ethics in the Conflicts of Modernity*'. *Politics & Poetics*, Vol. 4 (2018), pp. 1-21; Sigalet, G., 'Waldron's Challenge to Aristotelians: On the Political Relevance of Moral Realism'. *Politics & Poetics*, Vol. 4 (2018), pp. 1-23.

[67] See, for example, North, D. C., *Structure and Change in Economic History* (New York: W. W. Norton, 1981); North, D. C., *Institutions, Institutional Change and Economic Performance* (Cambridge: Cambridge University Press, 1990).

[68] North, D. C., 'Institutions'. *Journal of Economic Perspectives*, Vol. 5, No. 1 (1991), pp. 97-112, p. 97.

h. Rebuilding our citizenship

How can these traditional sources for discussing ethics and citizenship help us find civic ideals in our networked age? At play in both these accounts is a sense that challenges to our human dignity can be reversed to provide renewed moral ideals. In the case of the citizen-slave distinction, the horrors of slavery and slave-like conditions are rejected in favour of their opposite: equal status among human beings and common participation in the body politic. For the meta-ethics of human nature involving a constant struggle with competing desires, self-reflection on our truest desires and the extent to which our choices often fall short in fulfilling them help guide us towards a better hierarchy of values, a better ethical system.

The hoped-for contribution here is to argue that the much-cited tensions, difficulties and struggles brought about by the networked age can, by extension therefore, provide material for an opposite formation of civic ideals that act as antidotes. What does this mean? It means the empirically-grounded problems in mutual development of humanity and technology enjoys within itself remedies for those problems, in the form of virtuous opposites to current vices. A rich discussion of civic ideals can and should, therefore, accompany each and every conversation about civic collapse. That, at least, is the claim here.