

digital
egoism

Finding Virtue in the Virtual

Dr. Tom Chatfield

Perspectiva is a registered charity operating as a collective of scholars, artists, activists, futurists and seekers who believe credible hope for humanity's future lies in forms of economic restraint and political cooperation that are beyond prevailing epistemic capacities and spiritual sensibilities. We work to develop an applied philosophy of education for individual and collective realisation in the service of averting societal collapse; and to cultivate the imaginative and emotional capacity required to usher in a world that is, at the very least, technologically wise and ecologically sound.

systems-souls-society.com

Dr. Tom Chatfield

Dr. Tom Chatfield is a British author, educator and philosopher of technology. He has authored ten non-fiction books exploring digital culture - most recently: *Critical Thinking* and *Live This Book!*. He is also a guest faculty member at the Saïd Business School, Oxford, for its executive leadership program, a former associate at the Oxford Internet Institute, and a guest lecturer and educational advisor at institutions across Europe and the US. His debut novel, *This Is Gomorrah* (Hodder), was published worldwide in 2019, and was a Sunday Times thriller of the month, shortlisted for the CWA Thriller of the Year, and won the France's Prix Douglas Kennedy for the year's best foreign thriller.

Finding Virtue in the Virtual

Dr. Tom Chatfield

Part of the Digital Ego project.

What does it mean to grow and flourish, together, in a digital age?

The Digital Ego project is a Perspectiva initiative which seeks to speak at a systematic level to technology's mediation of modern life, asking what narratives and frames can connect personal experience to the global picture; can find virtue in the virtual realm; and can place a properly understood sense of self and soul at the heart of our mediated lives.

Finding Virtue in the Virtual

Dr. Tom Chatfield

What does it mean to place the ethics of technology upon firm foundations in the 21st-century? In this essay, I make the case that virtue ethics offers a practical, humane basis for doing so: that it can help us scrutinise the values entailed by the design, deployment, and regulation of technology; and that it can do so with a greater flexibility and faithfulness to lived experience than other overarching ethical accounts.

I'll also argue that virtue ethics can help us avoid certain category errors common to many discussions of technology: proffering ethical codes as a solution rather than a diagnosis; focusing too narrowly on data, code, and internal processes; and erasing social and political contexts via misleading claims of neutrality and inevitability.

1. The twin myths of tech neutrality and inevitability

There is no such thing as a neutral tool. To enter a vehicle is to transform your relationship with geography in particular ways. To lack a vehicle in a built environment expressly designed around their capabilities - to be unable to afford a car in Los Angeles, say - is to find yourself at the sharp end of a host of assumptions about freedom, space, and society.

Similarly, to pick up a weapon is to move through a world populated with potential targets. If I have a gun holstered on my belt, this changes me and my relationship with others in ways that can only be understood by analysing what the new entity 'me-and-my-gun' is capable of and disposed towards. As the philosopher Bruno Latour put it in his 1992 essay ['Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts'](#)¹:

The distinctions between humans and nonhumans, embodied or disembodied skills, impersonation or 'machination', are less interesting than the complete chain along which competencies and actions are distributed.

Why is this so significant? As slogans like '[guns don't kill people, people kill people](#)'² suggest, the seductive notion that technology itself is neutral - that a tool is simply a tool, and all that matters is how it's used - is all-too-frequently evoked in order to evade discussion of the assumptions and possibilities it embodies, not to mention the value-laden systems of regulation, power and profit surrounding it.

If technologies themselves are neutral, the people who make and maintain them have no particular responsibility towards the people who use them (and upon whom they're used) beyond ensuring certain standards of quality and functionality. If the most one can say about a town in which everyone walks around holding an assault rifle is that it's up to them to use their rifles responsibly, the question of what it means to live in a community where lethal force is a constantly visible prospect makes no sense.

All that can be expressed is a hope that people use their military-grade weapons ‘well’ - whatever that might mean in the context of an artefact designed expressly to kill in combat.

To talk about the possibilities, values, and preferences instantiated in technologies is to talk about what are often called their affordances: a term coined by the psychologist James J. Gibson in a [1977 paper](#)³ to describe the possibilities for action presented by a particular environment. As the philosopher Shannon Vallor notes in her 2016 book *Technology and the Virtues*⁴, acknowledging and analysing the affordances of technologies is an ancient challenge - but one with novel, urgent elements today:

The invention of the bow and arrow afforded us the possibility of killing an animal from a safe distance - or doing the same for a human rival, a new affordance that changed the social and moral landscape. Today's technologies open their own new social and moral possibilities for action. Indeed, human technological activity has now begun to reshape the very planetary conditions that make life possible... our aggregated moral choices in technological contexts routinely impact the well-being of people on the other side of the planet, a staggering number of other species, and whole generations not yet born. Meanwhile, it is increasingly less clear how much of the future moral labour of our species will be performed by human individuals.

In particular, exploring the gestalt nature of this moral labour - its diffusion of responsibility between those designing, regulating, using, and profiting from different technologies - is an important corrective both to the myth of technological neutrality and to a second, related error, embodied in what are known as deterministic accounts of innovation.

Technological determinism is based on the claim that new technologies more-or-less inevitably bring with them a set of fixed behaviours and outcomes, and that - to borrow a borrowed phrase from another philosopher, LM Sacasas - ‘resistance is futile’ when it comes to challenging these.

Sacasas himself borrows the phrase from no less an authority than *Star Trek: The Next Generation*, where it's the battle cry of the Borg collective, a cyborg civilisation whose mission is to assimilate all other life-forms into their hive mind. 'Resistance is futile!' its drones repeat as they try to extinguish every form of consciousness and freedom alien to their own. They're wrong, of course: the *Star Trek* universe wouldn't be much fun if resistance was indeed futile. But their sinister hubris is a handy (and gloriously heavy-handed) metaphor for all those mindsets that insist upon technology as a form of destiny. As [Sacacas notes](#)⁵, to identify and oppose what he calls the 'Borg complex' mode of tech analysis is to assert the ethical significance of intellectual freedom - and of taking responsibility for our creations:

One of the strangest things about the myths of technological neutrality and inevitability is that, even though they directly contradict one another, they're often articulated together.

Marshall McLuhan once said, 'There is absolutely no inevitability as long as there is a willingness to contemplate what is happening'. The handwaving rhetoric that I've called a Borg Complex is resolutely opposed to just such contemplation when it comes to technology and its consequences. We need more thinking, not less, and Borg Complex rhetoric is typically deployed to stop rather than advance discussion. What's more, Borg Complex rhetoric also amounts to a refusal of responsibility. We cannot, after all, be held responsible for what is inevitable.

One of the strangest things about the myths of technological neutrality and inevitability is that, even though they directly contradict one another, they're often articulated together. To say that a tool is neutral is to say that its users bear sole responsibility for what's done with it, presumably on the basis that this is their free choice. By contrast, to say that technology has an internal logic dictating certain outcomes is to say that people cannot ultimately choose whether or how to use it - and that dissent is the province of Luddite fools. Yet this deterministic rhetoric often dovetails with rhapsodies upon user empowerment. As the CEO of Evernote, Phil Libin, put it in a [2012 interview](#)⁶ (highlighted by Sacacas in his writings):

I'm actually very optimistic about the Google Glasses - and those by other companies who will make them... I've used it a little bit myself and - I'm making a firm prediction - in as little as three years from now I am not going to be looking out at the world with glasses that don't have augmented information on them. It's going to seem barbaric to not have that stuff. That's going to be the universal use case. It's going to be mainstream. People think it looks kind of dorky right now but the experience is so powerful that you feel stupid as soon as you take the glasses off...

It's all too easy to play the game of digging up predictions that didn't come true. But what's telling about Libin's line of argument is its treatment of human desire and technological possibility as twin sides of the same coin. Google Glass offers such a great experience that anyone who uses it will, seemingly inevitably, want to keep on using it. To do otherwise will become 'barbaric': it will mean existing outside the grand progress of technological civilisation.

In the best neo-Darwinian style, this framing suggests that technology's powers will sooner or later make its offerings synonymous with the outcome of a free choice (and that such choice is thus an illusion when it comes to aggregated human behaviours over time). People are being gifted more opportunities than ever before by products and platforms whose dominance is pre-ordained: a reading of history that's only plausible if you ignore the chaotically branching possibilities, debates, rethinks, and repercussions surrounding every innovation.

These myths of neutrality and inevitability matter not only because they deny both agency and responsibility when it comes to any choice more fundamental than 'which app shall I install next?' but also because, by doing so, they negate any basis for an ethics of technology that isn't based upon either expert condescension (*please invent the great innovation that will inexorably save us!*) or the decontextualised idealisation of personal responsibility.

In each case, what purports to be ethical engagement is little more than wishful buck-passing: the pretence that we live in a world where the complexities of our 'aggregated moral choices in technological contexts' can be palmed off as non-issues or personal preferences.

What's the alternative? It begins with paying close attention to what's actually going on.

2. Technological affordances and moral labour

Near the start of their 2018 book *Re-Engineering Humanity*⁷, law professor Brett Frischmann and philosopher Evan Selinger explore an example of what they term 'techno-social engineering' at Oral Roberts University in Tulsa, Oklahoma. In 2016, the university introduced a requirement for students to purchase and wear Fitbit tracking devices for a physical education class. Previously, students had self-recorded their daily activities in a journal. Now, these activities would automatically be recorded by their devices.

A minor controversy ensued concerning how far students had given informed consent to this tracking, how data would be stored, and so on. This controversy faded once it became clear that the university had provided adequate safeguards. One kind of monitoring had simply been replaced by another: the technology of pens and paper by automated tracking and recording. Who, in this day and age, would seriously suggest things should be different? Indeed, who would deny that Fitbits provide more detailed and more reliable data than journals, and do so more conveniently?

Frischmann and Selinger aren't in the business of mourning pens and paper. But, by digging into the different affordances of old and new approaches, they unearth some significant complexities. For a start, they argue that there are profound psychological differences between actively recording observations and passively being monitored:

Students who record their daily physical activities in a journal find the analog medium affords several steps that require time and effort, planning and thinking. It can orientate students to record fitness data in ways that automated and unreflective inscription machines could never do. The medium directs student attention inwardly and outwardly and the recorded data can reveal more than meets the eye.

For Frischmann and Selinger, it's this active/passive distinction, not the presence or absence of any particular technology, that matters. What's at stake is a certain ethic or set of values:

Think-and-record activities inspire self-reflection, interpersonal awareness, and judgement. These activities are valuable because they're linked to the exercise of free will and autonomy... The key to techno-social engineering better humans just might lie in taking these slower tools more seriously.

Within the space of two paragraphs, we have moved from a description of students scribbling in journals to a discussion of values associated with being a 'better' human being. Is this move justified? The answer, I would suggest, is an emphatic yes - and one that's all the more important for the starkness of placing such an ethically charged claim alongside what might more often be treated as a minor example of tech-enabled efficiency.

To see why, we need to consider not only students' actions and options, but also the obligations and expectations accompanying them. To ask someone to use a wearable device is to ask them to consent to a process of observation that will automatically generate exhaustive data about their daily activities. Once they agree, they will become part of a system that, if it works as intended, requires little from them beyond acquiescence. By contrast, asking them to record their own actions means asking them to embark upon a process of self-observation - and trusting them to do so diligently. This second scenario requires not only practical effort but also the kind of moral labour highlighted by Vallor: undertaking to perform a task accurately and honestly while resisting the temptation to distort or fabricate its results.

Especially in the context of education, it's reasonable to ask what kind of a student each of these approaches encourages someone to be - and what standards it suggests they'll be assessed by.

Is a good student someone who can be trusted to take responsibility for a sustained self-assessment; or is it someone whose comfort and convenience are best served by unobtrusive automatic monitoring (and who no longer has the option to skip their daily exercise)? You might reply that the most realistic answer is ‘a bit of both’ - but it’s not obvious that both options are on offer.

What kind of a person are students being encouraged to grow into by an education system that suggests constant, automated monitoring is a necessary feature of the world?

The implications of choices like this extend well beyond their immediate context. What kind of a person are students being encouraged to grow into by an education system that suggests constant, automated monitoring is a necessary feature of the world? What might it mean for a society to integrate such surveillance into the fabric of education; for students to perform all their schoolwork on devices that automatically report on their actions or inactions; or for facial recognition systems to track attentiveness in classrooms in real time?

None of these scenarios are hypothetical. Here’s how Todd Feathers and Janus Rose reported for Vice magazine’s [Motherboard](#)⁸ website in September 2020 on the growing use of ‘digital proctoring’ software to monitor students in some US colleges:

The software turns students’ computers into powerful invigilators - webcams monitor eye and head movements, microphones record noise in the room, and algorithms log how often a test taker moves their mouse, scrolls up and down on a page, and pushes keys. The software flags any behaviour its algorithm deems suspicious for later viewing by the class instructor.

Dystopian though it may sound, there are clear reasons for the widespread adoption of such tools. The Covid-19 pandemic has led to rapid increases in remote learning and assessment. This has in turn left colleges struggling with what it means to monitor students working from home, to prevent copying and cheating on a mass scale, and to come up with measurable proxies for attendance and participation.

So long as software is deployed responsibly, you might say, surely the diligent and the innocent have nothing to fear? As Motherboard's account suggests, this defence starts to founder once the affordances of remote technologies are more closely scrutinised. In the case of proctoring software designed to monitor online exam-taking, for example, a factor that should be entirely irrelevant to any assessment - the colour of someone's skin - can become a major obstacle thanks to the fact that some facial recognition systems repeatedly classify those with darker skin as being too poorly lit to recognise. Similarly, students with unreliable internet connections, disabilities, anxiety, ADHD, or who live in close quarters with dependents, are more likely to be flagged up as 'suspicious' thanks to the patterns of their gaze, their keyboard and mouse use, their physical environment, logon timings, and so on.

In such cases, automated systems' assumptions about what is desirable and 'normal' can't be separated from larger questions about the nature of 21st-century education, or indeed about membership in a 21st-century society. As Shea Swauger, Librarian and Senior Instructor at the Auraria Library, put it in an April 2020 article for [Hybrid Pedagogy](#)⁹:

Cheating is on the rise, we can't trust students, and the best strategy to protect academic integrity is to invest in massive surveillance systems. At least, that's the narrative that ed-tech companies catering to higher education are selling based on their products and marketing campaigns... If I take a test using an algorithmic test proctor, it encodes my body as either normal or suspicious and my behaviours as safe or threatening. As a cisgender, able-bodied, neurotypical, white man, these technologies generally categorise my body as normal and safe, and because of this, they would not endanger my education, well-being, employment, or academic standing. The majority of the students on my campus don't share my identities and could have a very different experience being read by test proctoring algorithms.

As its vendors have pointed out, colleges are under no obligation to use such software in any particular way, or indeed at all. But its very existence embodies a powerful set of incentives and assumptions around trust, privacy, and what it means to study and succeed as a student in the 21st-century. And - crucially - it's not the only model out there, either for education or technology. Alternative practices, approaches, and attitudes exist; and many students and educators spent 2020 [asserting their ethical and practical superiority](#).¹⁰

Even if a surveillance system can be made to work seamlessly, effectively, and impartially ([which seems unlikely](#)¹¹), what does it mean for a society to make submission to such monitoring a model for education, employment, or civic life? As Evan Selinger and the philosopher Evan Greer put it in a February 2020 [article](#)¹² warning against the move to deploy facial recognition technologies on university campuses (a warning that soon proved [prophetic](#)¹³):

Given the many ways [on-campus facial recognition] technology can be used and the ease of adding its functions to existing cameras, any deployment will normalise the practice of handing our sensitive biometric information over to private institutions just to get an education... [moreover] facial characterisation tends to be underwritten by junk science and integrating it into education risks dehumanising students and favouring overly-reductive approaches to teaching... Indeed, the mere prospect of widespread facial surveillance will have a chilling effect on campus expression. Students who are afraid to be themselves and express themselves will pull back from crucial opportunities to experience intellectual growth and self-development - and students from marginalised communities will be the most affected.

Societally, such software is of a piece with moves the Covid-19 pandemic has accelerated everywhere from business and leisure to governance and administration: towards the [normalisation of surveillance](#)¹⁴ and algorithmic data processing in the name of security and convenience; towards offers of efficiency and simplicity behind which under-examined prejudices or explicitly exploitative motives may lurk; and towards a fundamental asymmetry between what users themselves understand versus what others understand about them.

Indeed, the prospect of entire nations introducing regimes of total technocratic surveillance is now not so much speculative fiction as well-documented reality¹⁵. Is what Human Rights Watch has termed the ‘automated tyranny’¹⁶ of China’s pandemic response a foretaste of all our futures?

The right questions can only begin to be asked, however, if technology’s affordances are borne in mind, together with the values and purposes they embody.

Frischmann and Selinger touch on all of these concerns in *Re-Engineering Humanity*. Yet they don’t end their opening chapter with a jeremiad. Instead, having analysed the affordances of old and new approaches in the case of Oral Roberts University, they suggest some modest positive steps that might be taken based on such an analysis:

...the university could combine the fitness tracking tools. It could require students to use a fitness tracking device that collects data, while also expecting them to write reports about the collected data in a journal. This two-step process would be more comprehensive and accurate than journaling alone. It also gives students an opportunity to reflect on their performance and freedom to define how and what to communicate to their instructors and peers...

Once the right questions have been asked, in other words, a negotiation can in principle take place between different systems and approaches, animated by a clear discussion of what human ends the result should be directed towards - and what might need to be mitigated along the way.

The right questions can only begin to be asked, however, if technology’s affordances are borne in mind, together with the values and purposes they embody. This in turn demands an explicitly ethical understanding of the assumptions embodied in a technology’s design and deployment - and the permission and the will to turn such investigations into action.

3. Towards a meta-ethics of technology

Near the beginning of *Technology and the Virtues*, Shannon Vallor coins the term ‘technosocial opacity’ to summarise the depths of uncertainty that characterise the present’s visions of the future - and the depths of ambivalence surrounding technology’s place in it:

Our present condition seems not only to defy confident predictions about where we are heading but even to defy the construction of a coherent narrative about where exactly we are. Has the short history of digital culture been one of overall human improvement or decline? On a developmental curve, are we approaching the next dizzying explosion of technosocial progress as some believe, or teetering on a precipice awaiting a calamitous fall ... Our growing technosocial blindness, a condition that I will call acute technosocial opacity, makes it increasingly difficult to identify, seek, and secure the ultimate goal of ethics - a life worth choosing a life lived well.

If, in such a context, we wish to invoke such ideas as ‘ethics’ and ‘purpose’, where can and should we look for guidance as to what they mean? This question concerns what’s known as meta-ethics. To discuss meta-ethics is to discuss how we define fundamental concepts such as right, wrong, goodness, and morality: to ask what it means to offer a coherent, compelling account of ethics for our times. As the title of this essay suggests, I believe that the answer lies in a version of the approach known as virtue ethics. Before we consider such an ethics in depth, however, it’s important to consider two other major schools of meta-ethical thought in western philosophy - deontological and utilitarian ethics - and why they may be less fruitful. My analysis, it should be noted, is indebted to Vallor’s foundational work.

Deontological ethics is interested in questions of moral duty, and the rules of right action that might define such duty. Perhaps the most famous of these is Immanuel Kant’s categorical imperative: the argument that each individual should ask of each of their actions, ‘is the principle upon which I am acting one that should also govern the actions of all other people in similar situations?’ In other words, an action is only right if it flows from a moral rule that any right-thinking person would wish to be universal.

Kant's rule offers a powerful riposte to the prospect of people picking and choosing personal definitions of right and wrong, as well as to the view that no universal ethical standards can be asserted purely based on human experience. As Vallor points out, however, its very universality also renders it curiously impotent in the face of present uncertainties:

Consider the dutiful Kantian today, who must ask herself whether she can will a future in which all our actions are recorded by pervasive surveillance tools, or a future where we all share our lives with social robots... How can any of these possible worlds be envisioned with enough clarity to inform a person's will? To envision a world of pervasive and constant surveillance, you need to know what will be done with the recording, who might control them, and how they would be accessed or shared...

In other words, the contingent questions begged by any such future scenario render the formulation of universal duties incoherent. Unless, of course, we're willing to embrace precisely the uncertainty that deontological ethics seeks to dispense with: to frame the future's duties in terms of what we may owe to one another in specific instances, and to ask what different moral questions we might wish to ask of each emerging situation.

The other major meta-ethical school of utilitarian thought similarly founders on opacity. Utilitarianism - and the broader ethical category to which it belongs, consequentialism - is based upon the powerfully pragmatic principle that right actions are those aligned with the best possible outcome for the greatest possible number of people. This approach can also be framed in terms of harm and risk reduction, as seen in the work of philosophers like Peter Singer and Nick Bostrom. Right actions, in this context, are those which do most to reduce preventable human (and animal) suffering, and/or which make catastrophic future events less likely.

While deontological ethics is interested primarily in an individual's sense of duty - and thus the ways in which personal intentions map onto generalisable moral rules - utilitarian ethics is interested in the achievement of particular worldly states of affairs.

To paraphrase one of the most famous arguments from Singer's 2009 book *The Life You Can Save*¹⁷, almost anyone would naturally leap into a shallow pond in which a child was drowning if the only cost were replacing their brand new trainers afterwards. Yet, for less than the price of such a pair of trainers, almost everyone living in some degree of comfort can transform the lives of several people suffering elsewhere by, say, donating to a charity like the [Against Malaria Foundation](#)¹⁸. Thus, everyone should either do so, or seek to undertake similarly impactful actions.

For me, arguments such as Singer's are simultaneously compelling, of immense ethical significance, and inadequate. They offer a pragmatic guide to maximising certain desirable outcomes from certain resources - and have been influential on attempts at establishing rigorous utilitarian frameworks such as the [Effective Altruism movement](#)¹⁹ - without at any stage constituting a systematic account of human ethical relations. Once we have agreed that certain outcomes are desirable, the reasoned calculus of maximising these outcomes is hugely valuable. But the ethical reasoning supporting such a calculus must inexorably have taken place elsewhere, in contexts within which even an appeal as seemingly self-evident as that of reducing suffering cannot offer clear guidance. Where, to echo Bernard Williams's [critique](#)²⁰ of utilitarianism, are the non-subjective moral sentiments to which we might appeal when searching for some 'impersonal' perspective from which to make our assessment?

At the other end of the scale to Singer's focus on immediately preventable suffering - a divergence that itself suggests the difficulty of reconciling rival utilitarian framings - thinkers like Bostrom suggest a series of criteria and caveats in key domains, aimed at avoiding civilisational disaster. These criteria are typified by the convergence of ethical frameworks for AI around such principles as transparency, justice and fairness, nonmaleficence, responsibility, and privacy.

There is much to admire (and heed) in warnings against worst-case scenarios for our species. In their applications, however, such frameworks start to more closely resemble the practical wisdom virtue ethics aspires towards than Kantian or consequentialist commandments. As Anna Jobin, Marcello Ienca & Effy Vayena argued in '[Nature Machine Intelligence](#)'²¹ in September 2019, when it comes to the future of AI there is:

...substantive divergence in relation to how these principles are interpreted, why they are deemed important, what issue, domain, or actors they pertain to, and how they should be implemented.

Such ethical codes are much less like computer code than their creators might wish. They are not so much sets of instructions as aspirations, couched in terms that beg at least as many questions as they answer.

Despite the power and importance of utilitarian analyses within certain domains, in other words, there is no one great societal test to be passed, no single consensus or paradigm to be shifted - and no way of imposing alleged solutions into such spaces without silencing many of those voices that most need to be heard. There is, rather, the unfolding collective challenge of finding ways of flourishing under conditions of technosocial opacity - and, incrementally, imperfectly, of creating virtuous cycles of technology's development, interrogation, and deployment.

4. Virtue in practice

A central contention of virtue ethics is that, given the profound uncertainties surrounding each unfolding life, no one trajectory is guaranteed to provide purpose or contentment - but that it is possible to describe the kind of conditions and aptitudes compatible with the fulfilment of human potential. Such fulfilment is termed, in the Aristotelian virtue ethical tradition, *eudaimonia*. What does *eudaimonia* entail? The philosopher and classicist Edith Hall teases out some of its complexities in her 2018 book [*Aristotle's Way*](#)²²:

The eu- prefix (pronounced like 'you') means 'well' or 'good'; the daimonia element comes from a word with a whole range of meanings - divine being, divine power, guardian spirit, fortune, or lot in life. So eudaimonia came to mean well-being or prosperity, which certainly includes contentment. But it is far more active than 'contentment'. You 'do' eudaimonia; it requires positive input. In fact, for Aristotle, happiness is activity (praxis). He points out that if it were an emotional disposition which some people are either born with or not, then it could be possessed by a man who spent his life asleep, 'living the life of a vegetable...'

Aristotle is, Hall notes, 'usefully gregarious and concrete as a model for virtue in practice' - which isn't the same thing as being timelessly correct. Aristotle was wrong about plenty of things (gender politics and slavery among them). In bequeathing the world a view of ethics that insists upon their concrete, contingent quality, however, he provided a framework well-suited for addressing the tensions and interdependencies I've anatomised so far - not to mention a philosophy compatible with a host of other traditions committed to purposeful self-development.

In particular, virtue ethics is committed to the idea that moral character lies at the heart of ethics; and that, paradoxically, it is primarily by working on our own character that we become able to treat others well. Moral character is a capacious concept. It relies not on fixed rules of wrong and right action, but rather on practising virtuous behaviours in day-to-day life - and the psychological significance of role models upon behaviour and beliefs. Every action, no matter how small, is potentially a precedent.

Similarly, inactions and happenstance are of great significance. To be disadvantaged, abused, or unfortunate is to be confronted by obstacles to thriving that it may prove impossible to overcome. In this sense, civic virtues such as respect for justice, fairness, and liberty - and the communal cultivation of these - can be of greater weight than purely personal achievements.

Perhaps above all, virtue ethics is determinedly modest in its ambitions.

Perhaps above all, virtue ethics is determinedly modest in its ambitions. It sees thriving and goodness alike as lifelong journeys with no final destination, and even the best of us as only too human. As the philosopher Julian Baggini put it in his 2020 book [*The Godless Gospel*](#)²³, an exploration of Jesus's reported words and deeds as a model for secular ethics:

One neglected feature of Jesus's example is that he models the need for work on the self. The supposed divinity of Christ tends to make us think of his goodness as being inherent, but this is not how he is portrayed in the Gospels. For sure he had a precocious wisdom.... And yet he did not begin his ministry until he was thirty. Even someone as morally gifted as Jesus needed time for his wisdom to grow, and that wisdom needed constant nurturing.

It's useful at this point to consider a concrete example of virtue in practice when it comes to tech; and, in particular, what it means to align the development and deployment of a technology with the growth, freedom, and empowerment of those affected by it.

In November 2016, the researcher Joy Buolamwini - then a grad student at MIT - spoke at [TEDxBeaconStreet](#)²⁴ about facial recognition systems and race. When she was an undergraduate at Georgia Tech studying computer science, Buolamwini explains, she used to work on so-called social robots - and soon discovered that the robot she was using couldn't 'see' her because of the colour of her skin. In a pre-emption of the problem with some proctoring software discussed earlier in this essay, she found that she had to 'borrow' her (lighter-skinned) roommate's face in order to complete a project. Soon after this, she visited Hong Kong to take part in an entrepreneurship competition and paid a visit to a local start-up that was demonstrating one of its social robots. 'You can probably guess', Buolamwini says, what happened next:

The demo worked on everybody until it got to me... It couldn't detect my face. I asked the developers what was going on, and it turned out we had used the same generic facial recognition software. Halfway around the world, I learned that algorithmic bias can travel as quickly as it takes to download some files off of the internet.

As a recent stream of examples has emphasised - from Zoom calls 'cutting off'²⁵ the heads of those with dark skin, to Twitter algorithms automatically placing white faces²⁶ at the centre of cropped images - Buolamwini was being excluded by default from such categories as 'normal', 'significant' and even 'human'. Importantly, however, she was also far from a passive victim.

In order for a computer to 'see' anything, a machine learning algorithm must be trained by exposing it to samples of whatever it is supposed to recognise: in this case, hundreds of thousands of examples of both faces and things-that-are-not-faces. If only certain types of faces are included in the training set, those who deviate too far from their norm will be harder to detect. All of this, Buolamwini notes, embodies not so much the implacable verdict of an automated system as the explicit product of a series of human choices:

Training sets don't just materialise out of nowhere. We actually can create them. So there's an opportunity to create full-spectrum training sets that reflect a richer portrait of humanity... we can start thinking about how we create more inclusive code and employ inclusive coding practices. It really starts with people. So who codes matters. Are we creating full-spectrum teams with diverse individuals who can check each other's blind spots? On the technical side, how we code matters. Are we factoring in fairness as we're developing systems? And finally, why we code matters. We've used tools of computational creation to unlock immense wealth. We now have the opportunity to unlock even greater equality if we make social change a priority and not an afterthought.

Why, how, who: for all the complexities of the answers they demand, the questions that unlock the black box of encoded injustice couldn't be simpler. And this in turn suggests some of the most fundamental things we can say about the biases, prejudices, and injustices latent in tech systems: that all of these are only ever latent or invisible to somebody; and that it's only a narrowly deterministic narrative that allows this somebody to plead ignorance on behalf of humanity as a whole.

It is now almost five years since Buolamwini's talk, time in which she has helped build [one among a growing number of movements](#)²⁷ advocating for equitable, accountable AI. Yet the very flaw she identified continues to create divisions and disadvantages - as do countless other inequities, exclusions, and injustices (consider the [ongoing scandal](#)²⁸ of Google's dismissal of the two co-leads of its AI ethics team, the world-renowned researchers Timnit Gebru and Margaret Mitchell).

What's going on; and what can be done about it? The answer, I would suggest, is as much about the people and priorities present (and absent) in boardrooms and workplaces as it is about data or code. And it points towards the heart of the problem for tech ethics itself. When it comes to technology, it's not enough that we seek either virtuous tools or virtuous people. Rather, we need to ask what it means for the ongoing process of designing, debating and deploying a technology to itself be a virtuous one

5. What is to be done?

If prejudice and injustice are inscribed in the data we feed into machines, then scrutinising this data presents a profound ethical opportunity: a chance simultaneously to recognise and redress structural inequalities and exclusions. Importantly, however, it will never be ethically adequate to focus only (or even primarily) upon data itself. Why? As the researchers Alex Hanna, Emily Denton, Andrew Smart, Hilary Nicole, and Razvan Amironesei argued in a December 2020 essay for [Logic](#)²⁹ magazine:

A particularly pernicious consequence of focusing solely on data is that discussions of the 'fairness' of AI systems become merely about having sufficient data. When failures are attributed to the underrepresentation of a marginalised population within a dataset, solutions are subsumed to a logic of accumulation; the underlying presumption being that larger and more diverse datasets will eventually morph into (mythical) unbiased datasets. According to this view, firms that already sit on massive caches of data and computing power - large tech companies and AI-centric startups - are the only ones that can make models more 'fair'.

There is, in other words, a gaping absence at the heart of any argument that ethical issues can be resolved solely by relying on big companies to build up bigger and better datasets. For much the same reason as there's no 'neutral' ethical perspective from which a utilitarian can weigh the world in their scales, 'unbiased' datasets are mythical artefacts predicated upon an impossibility: a world in which no value-laden choices or preferences exist around a technology's research, development, deployment, governance, and regulation.

For me, a great gift of virtue ethics is that it requires us to address precisely this context through the lens of each life's potentials and dignity: that we acknowledge the explicitly ethical interdependencies of a society's norms, inclusions, exclusions, and the weighty individual and collective demands made of us by hopes of growth and thriving.

Indeed, perhaps the weightiest of all these demands is that we acknowledge the depths of our fallibility, vulnerability, and dependency, both upon one another and upon the systems surrounding us. In his 1999 book *Dependent Rational Animals*³⁰, the philosopher Alasdair MacIntyre makes the case that discussing human existence in terms of the 'normal' capabilities of healthy, seemingly autonomous adults is itself a profound ethical category error. This is not only because to do so is to ignore the arbitrariness of the world's inequalities, but also because our existence is defined in the most fundamental sense by dependency: by our species' extended infancy and childhood; by sickness, infirmity, and age; by tools, trade, and technology, without which there is no such thing as a human society.

If we are meaningfully to discuss life as it is lived, MacIntyre suggests, we must begin not with a snapshot of some notionally independent adult, but rather by acknowledging that each life's interwoven trajectory demands:

...that those who are no longer children recognise in children what they once were, that those who are not yet disabled by age recognise in the old what they are moving towards becoming, and that those who are not ill or injured recognise in the ill and injured what they often have been and will be and always may be.

I have two young children and, like many parents, one of the earliest lessons I've struggled to master is that my children's desires are an imperfect guide to their wellbeing; and that making their lives easier is not always the best way to prepare them for life.

It also matters, MacIntyre continues, that this recognition of mutual dependency is not couched in terms of fear or rejection. To be human is to be born into utter helplessness, in circumstances beyond our choosing. It is to grow and change, constrained by these circumstances and biological inheritance. It is to achieve some measure of independence, for a time, in the context of society's vast networks of exchange and competition. And it is to seek not only survival but also - so long as the body's basic needs are met - some form of flourishing or contentment. There is no final victory, no guarantee of success, and no infallible guidance. There is only the contingent business of trying, together, to live and to know ourselves a little better.

All of the above entails, to repeat a phrase I've used several times already, moral labour whose difficulty and significance are inextricably linked. I have two young children and, like many parents, one of the earliest lessons I've struggled to master is that my children's desires are an imperfect guide to their wellbeing; and that making their lives easier is not always the best way to prepare them for life. Much like the students Frischmann and Selinger describe, it's more important for me gradually to help them develop a measure of self-control, fairness, and ambition - and to show them that trust can be earned - than it is for me constantly to monitor and intervene in everything they do.

Also like many parents, a second lesson I'm still trying to learn is that the other person who all-too-often needs to improve their self-control is me. To love and to nurture other human beings brings pain as well as joy; frustration and exhaustion as well as delight; the prospect of devastating loss alongside the gain of consuming love. And these satisfactions and sacrifices can't tidily be separated. To withdraw your care from any relationship is to make yourself less vulnerable, for a price: it's to diminish what you risk and give, but also what you can receive and gain.

I could make my life easier by outsourcing my children's education, discipline, and nurture to the nudges of expert systems, much as a government might choose to reward or punish its citizens' actions via implacable, ubiquitous surveillance.

In each case, however, the fantasy of an optimised existence is one that hollows out not only people's relations with each other, but also the value of most other things worth pursuing. It seeks to impose an empty vision of perfectibility in place of the purposeful, mutual struggles through which human dignity and potential are asserted and sustained.

6. Virtues for the virtual

Crucially, the moment those designing and deploying a technology start seeking out others' experience rather than making assumptions on their behalf - the moment they start embodying open questions like why, how, and who in a design process rather than declaring certain technocrats' preferences to be synonymous with the 'logic' of technology itself - they begin, for the first time, to see technology as it actually is. That is, they begin to see the human-made world as one that its creators and maintainers at once bear responsibility for, and are constantly instantiating this responsibility within.

Where does this lead when it comes to this essay's promise: of finding virtue in the virtual? It begins with the fact that all deployments of technology imply a certain ethic or set of values. There is no such thing as a neutral tool - which makes it vital to pay attention to the affordances both of technologies and the contexts they exist within.

In particular, we need to beware of the boosterist rhetoric of convenience, ease and efficiency, and its connection to two interrelated myths: of technology's neutrality and of the inevitability of the changes it brings about. Against these, we must pay particular attention to the nature of the moral labour entailed by different situations - and what it implies for such labour to be outsourced to or via information systems.

This necessary attentiveness takes the form of questions; and of the time, space, and will to ask and address them. What kind of a person - what kind of a citizen, a student, a worker, a friend - do such systems encourage us to be. How do they encourage us to relate to others? What assumptions around normality, desirability, and excellence are we automating within them?

As the world buckles beneath the pressure of the Covid-19 pandemic, it is becoming all too easy for surveillance to infiltrate ever further into our lives - and to do so in the name of maintaining standards, preventing deceit, ensuring fairness, and providing support. Such claims are hollow at the core: not because they are ineffective (it's their putative effectiveness and efficiency that makes them so seductive) but because they are too often corrosive of the very possibility of earning or bestowing trust; of the private spaces within which self-knowledge, self-authorship, and rich mutual engagement can occur.

Against this, what's required is an explicitly ethical understanding of the assumptions embodied in a technology's design and deployment: one alive to the complexity, opacity, and interdependencies of the 21st-century context; one empowered to address and redress structural injustices at the institutional as well as the technological level; one able to define and defend the ethical and legal frameworks within which the proportionate, accountable collection, retention and processing of information can take place.

Neither expert condescension nor the decontextualised praise of personal responsibility is adequate for such tasks - and nor can universalised accounts of moral duty or utilitarian calculus provide a sure ethical foundation.

In the virtue ethical tradition, however, there is something sufficiently modest and humane to speak to our times: something that begins by acknowledging our limitations, our interdependencies, and the significance of our circumstances; that embraces the plurality of routes to human flourishing; and that understands the necessarily contingent, communal nature of the practices such flourishing might arise from.

Central to the idea of virtue is its practical cultivation over the course of each life and, in parallel with this, a belief in the human potential to grow beyond our beginnings: to follow role-models, and potentially to become one; to seek self-authorship within the context of a meaningful community in a manner closely aligned with the German concept of [Bildung](#)³¹.

In particular - in the context of contemporary societies within which technology is implicated in every facet of life - an ethics of technology founded on the attentive interrogation of a plurality of experiences is required. This interrogation should take its direction from the dismantling of embedded injustices and inequalities around 'normality' and desirability; of exploitative and manipulative forms of surveillance; and of the loss of human dignity and potential that comes with the outsourcing of education, work, and governance to opaque, unchallengeable systems.

As a final philosophical aside, the work of the philosopher Luciano Floridi has provided an important and vivifying context for these reflections. Floridi's [informational ontology](#) is ecological in its concern for the health of the information environments within which we exist alongside countless human-made entities, all of which bear some minimal moral weight. This is Kantian in its scope, and deeply informed by information theory, but in its emphasis on mutually dependent thriving, it also offers a paradigm for the architecture of a networked world aligned with human dignity and freedom.

If much of the above sounds abstract, its implications - as befits a philosophical tradition emphasising the importance of *praxis* (thoughtful action) and *phronesis* (practical wisdom) - are only too tangible. As the social psychologist Shoshana Zuboff articulates in her critique of '[surveillance capitalism](#)'³², one of the information age's most significant frontiers for power and profit entails algorithmic systems at once predicting their users' aggregated actions and conspiring to make these predictions come true. That is, it entails the deployment of behaviourist models preoccupied above all with keeping their users 'stuck' in certain predictable patterns. For the author and technologist [Jaron Lanier](#)³³, such a model constitutes nothing less than [addiction by design](#), with all the losses and diminishments this suggests:

*The algorithm is trying to capture the perfect parameters for manipulating a brain, while the brain, in order to seek out deeper meaning, is changing in response to the algorithm's experiments....
As the algorithm tries to escape a rut, the human mind becomes stuck in one.*

This is where the [Digital Ego](#) project³⁴ - which I have developed for Perspectiva alongside its founding director, Jonathan Rowson, and the writer and researcher Dan Nixon - comes in. To cite the first of its foundational principles, the project is devoted to 'defining and advocating for what it means to be free in the digital era', as opposed to becoming 'stuck' within systems explicitly engineered to resist such freedom. Following from this, the project focuses on models of online community predicated around freedom and autonomy; on challenges to the endorsement of optimisation, efficiency, and novelty as somehow inherent to technology; and, reflecting Perspectiva's cross-level focus on systems, souls, and society, upon a fundamentally plural account of paths to human flourishing.

The project is, I hope, a gregarious and pragmatic undertaking, which recognises that these challenges should primarily be redressed through practices and communities rather than enumerations of principles; that there is no such thing as an analysis of technology that isn't also an analysis of its embedding in particular social and political circumstances; and that one of humanity's most important undertakings when it comes to technology is to resist and reject its ill-considered implementations.

From facial recognition systems to the normalisation of ubiquitous surveillance, from autonomous weapons to weaponised social media ecosystems, there has never been a stronger case for mindful delay, dissent, and disavowal - and for forms of ethical thinking that place such dissent upon firm foundations. As the philosopher Carissa Véliz puts it in her 2020 book *Privacy is Power*³⁵, to speak of virtue and lived experience in present times is necessarily to speak of righteous anger as well as cool consideration; of the fact that human growth and flourishing are sometimes best served by resistance:

Aristotle argued that part of what being virtuous is all about is having emotions that are appropriate to the circumstances. When your right to privacy is violated, it is appropriate to feel moral indignation. It is not appropriate to feel indifference or resignation. Do not submit to injustice. Do not think yourself powerless - you're not.

There is always a choice. My hope is that, together, we can more often make it a wise one.

Endnotes

1 Latour, Bruno, 'Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts', in *Shaping Technology-Building Society: Studies in Sociotechnical Change*, MIT Press (1992), p. 225 - 259.

2 Selinger, Evan, 'The Philosophy of the Technology of the Gun', *The Atlantic* (2012)

3 Gibson, James, 'The Theory of Affordances', in *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, Laurence Erlbaum Associates (1977)

4 Vallor, Shannon, *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*, OUP USA (2016)

5 Sacasas, L. M., 'Borg Complex: A Primer', thefrailestthing.com (2013)

6 Rundle, Michael, 'Evernote CEO Phil Libin Interview: Evernote Business, Coal Mines And 'The Nike Of The Well-Ordered Mind'', *Huffpost* (2012)

7 Frischmann, Brett, & Selinger, Evan, *Re-Engineering Humanity*, Cambridge University Press (2018)

8 Feathers, Todd, & Rose, Janus, 'Students Are Rebelling Against Eye-Tracking Exam Surveillance Tools', *Motherboard* (2020)

9 Swauger, Shea, 'Our Bodies Encoded: Algorithmic Test Proctoring in Higher Education', *Hybrid Pedagogy* (2020)

10 Kelley, Jason, 'Students Are Pushing Back Against Proctoring Surveillance Apps', *Electronic Frontier Foundation* (2020)

11 Gastschrijver, 'Online proctoring isn't just wrong – it's ineffective', *Mare Online* (2020)

12 Greer, Evan, & Selinger, Evan, 'How Facial Recognition Technology Could Change College Campuses Completely', *MTV News* (2020)

13 Evan Greer on Twitter, (October 14th 2020)

14 Tung, Liam, 'Microsoft 365's Productivity Score: It's a full-blown workplace surveillance tool, says critic', *ZDNet* (2020)

15 Human Rights Watch website – 'Mass Surveillance in China'

16 Wang, Maya, 'China: Fighting COVID-19 With Automated Tyranny', *Human Rights Watch* (2020)

17 Singer, Peter, *The Life You Can Save*, Penguin Random House (2009)

18 <https://www.thelifeyoucansave.org/best-charities/against-malaria-foundation/>

19 <https://www.effectivealtruism.org/>

20 Smart, J. J. C., & Williams, Bernard, Utilitarianism for and against, *Philosophia*, (1977)

21 Jobin, Anna, & Marcelo, Ienca, & Vayena, Effy, 'The global landscape of AI ethics guidelines', *Nature* (2019)

22 Hall, Edith, *Aristotle's Way: Ten Ways Ancient Wisdom Can Change Your Life*, Penguin (2019)

23 Baggini, Julian, *The Godless Gospel*, Granta (2020)

24 Buolamwini, Joy, 'How I'm fighting bias in algorithms', *TEDxBaconStreet* (2016)

25 Rose Dickey, Megan, 'Twitter and Zoom's algorithmic bias issues', *Tech Crunch* (2020)

26 Hern, Alex, 'Twitter apologises for 'racist' image-cropping algorithm', *The Guardian* (2020)

27 <https://www.ajl.org/>

28 Schiffer, Zoe, 'Google fires second AI ethics researcher following internal investigation', *The Verge* (2021)

29 Amironesei, Razvan, & Denton, Emily, & Hanna, Alex, & Nicole, Hillary, & Smart, Andrew, 'Lines of Sight', *Logic* (2020)

30 MacIntyre, Alisdair, *Dependent Rational Animals: Why Human Beings Need the Virtues*, Bloomsbury (2009)

31 Rowson, Jonathan, 'Bildung in the 21st Century: Why sustainable prosperity depends upon reimagining education', *CUSP* (2019)

32 Zuboff, Shoshana, *The Age of Surveillance Capitalism*, Profile Books (2019)

33 Lanier, Jaron, 'Book excerpt: Jaron Lanier's 'Ten Arguments for Deleting Your Social Media Accounts Right Now', *ABC News* (2018)

34 <https://systems-souls-society.com/realisation/digital-ego-project/>

35 Veliz, Carissa, *Privacy is Power: Why and How You Should Take Back Control of Your Data*, Bantam Press (2020)

PER
SPE
CT
IVA