PROMOTING CHARACTER EDUCATION AS PART OF A HOLISTIC APPROACH TO RE-IMAGINING THE DIGITAL AGE

Ethics and the Internet Webinar
9th December
14.30 – 16:00 (UK time)

Organised by
Professor James Arthur
Dr. Tom Harrison
Dr. Gianfranco Polizzi
Introduction
p. 2

1. Character and the Computer Scientist
Professor Kirsten Martin - p. 6

2. Algorithmic Explainability and Practical Wisdom
Dr. Christopher Burr - p. 9

3. Virtue and Vice in the Design and Deployment of Digital Technologies for Education
Dr. Simon Coghlan, Henrietta Lyons, Dr. Tim Miller and Professor Jeannie Marie Paterson - p. 12

4. Justice & Social Credit
Dr. Marcelo Thompson - p. 16

5. Digital Well-Being: A New Frontier in Human Flourishing
Dr. Matthew Dennis - p. 19

Dr. Alison Powell - p. 22

7. Re-Imagining the Digital Age through Digital Ethics
Professor Simon Rogerson - p. 25

Big Questions and Future Directions for Research and Practice
p. 29
INTRODUCTION

For quite a few years now, the extent to which the belief that the internet is not only a force for good but can also undermine human flourishing - that is, our ability to thrive and live well collectively - has been on the rise. We live in an age that is saturated with digital technologies. But while the internet presents a range of opportunities - e.g., for learning, socialising and participating in society - it also poses a number of challenges, from issues of privacy, financial safety and misinformation, to cyberbullying and online abuse. On the one hand, these challenges undermine individuals’ ability to use the internet safely. On the other hand, and more problematically, the extent which online abuse can undermine public debate, the spread of cyberattacks and misinformation during elections, or the illegal harvesting of users’ data for political purposes (as in the case of Cambridge Analytica), suggest that these challenges have ethical implications and wider consequences for our societies and for democracy.

In order to discuss these challenges in more depth, the Jubilee Centre for Character and Virtues of the University of Birmingham organised and held a webinar titled ‘Promoting Character Education as Part of a Holistic Approach to Re-Imagining the Digital Age’. The webinar, which took place on 9th December 2020, was an opportunity for a selected number of academics from multiple disciplines to submit, prior to the webinar, and to present, during the webinar, short positions papers on a number of different, but interrelated, themes.

The Jubilee Centre is currently undertaking a project titled ‘Cultivating Cyber-Phronesis’, which examines whether, how and to what extent a targeted approach to character education - one that is primarily grounded in the cultivation of different virtues such as honesty and compassion - can help adolescents aged 13-16 make better ethical decisions online, particularly when interacting and communicating with others. However, many of the challenges presented by the internet, which are exacerbated by the ways in which internet corporations run search engines and online platforms, cannot be tackled exclusively through education. That is why policymakers and tech experts in the United Kingdom and beyond have been grappling with the question of how to ensure that the digital environment operates and is managed and designed in ways that are both empowering and safe for users.
As a result, the idea behind the webinar was to explore the ways in which we can promote character education in the digital age as part of a more holistic approach to human flourishing, one that requires a commitment to re-imagining the digital age, from how users engage with digital technologies to the responsibilities of governments and internet corporations. Under this overarching theme, a total of seven academics working from multiple disciplines ranging from philosophy, law and media studies, submitted position papers focusing on the political economy and socio-technical features of the digital environment, on internet governance, on virtue ethics and character and moral education.

The first paper below, submitted by Professor Kirsten Martin from the University of Notre Dame in the United States, starts with an analysis of some of the ethical constraints that characterise the digital environment. As Martin argues, such an environment is driven by internet corporations’ economic interests, which are rooted in practices of commodification and data tracking. The challenges that these practices raise for the individual user, and the extent to which the profits of internet corporations are prioritised regardless of the different forms of online harm and abuse that are typical of the digital age, require a new solution. What Martin suggests is that computer scientists should be trained through virtue-based character education so that they can be better equipped to re-shape the design of digital products in ways that are both mindful of digital inequalities and based on principles of integrity.

In his paper, Dr. Christopher Burr from the University of Oxford and the Alan Turing Institute in the UK, argues that algorithmic opacity, which refers to the lack of transparency that underpins the functioning of algorithms, undermines users’ ability to make wise decisions online insofar as it creates an “epistemic barrier”. According to Burr, besides promoting digital literacy education, it is imperative to redesign the digital environment in ways that facilitate algorithmic explainability – which is based on delivering explanations to users of how algorithms function and how they are programmed to make decisions – over algorithmic opacity.

The third paper, submitted by Dr. Simon Coghlan, Henrietta Lyons, Dr. Tim Miller and Professor Jeannie Paterson from the University of Melbourne in Australia, focuses on the ethical challenges that result from how ed-tech corporations operate – i.e., from how they use facial recognition to how they rely on algorithms and AI to predict students’ behaviour or to assign grades. As Coghlan and colleagues suggest, the
education system needs to be re-imagined in ways that promote virtue ethics. According to them, this applies 1) to educational institutions, which should prioritise virtues such as compassion, honesty and fairness in their approach towards students, 2) to students themselves, who need to develop the skills and awareness necessary to use digital technologies in ways that are driven by different virtues, and 3) to educators, who need to incorporate elements of virtue-based moral and character education into their practices.

The paper submitted by Dr. Marcelo Thompson from the University of Hong Kong focuses on governance by framing the Chinese social credit system as an abstract system that lies at the intersection of identity, trust and law. Thompson draws on such a system as an example to argue that modern societies need regulatory mechanisms designed to respond to the challenges that are inherent in how data can be processed at a government level for the purpose of identifying identity attributes both among citizens and corporations. According to Thompson, these kinds of mechanisms should be implemented in ways that are mindful of any potential distortions, and in line with virtues that range from harmony and self-cultivation to self-respect.

As with the paper submitted by Thompson, the last three papers also focus on governance, but they do so in ways that are connected with questions of education and how the digital environment is, and should be, redesigned, thus reprising some of the topics introduced above. In his paper, Dr. Matthew Dennis from Delft University of Technology in the Netherlands, suggests that we, as a society, can only face the ethical challenges that digital technologies pose as long as we take a character- and virtue-based ethical approach to technology. According to Dennis, such an approach depends on promoting not just character education in the digital age, but also regulatory mechanisms and the design of digital products that problematise, in ways that are both ethical and driven by virtues, both issues of commodification and the extent to which users may be subject to manipulation.

The paper submitted by Dr. Alison Powell from the London School of Economics and Political Science also focuses on virtue ethics. But as Powell argues, a virtue ethical approach to technology is not sufficient on its own in taking account of the socio-cultural and political economy structural constraints that characterise the digital environment. That is why, for Powell, such an approach needs to be complemented 1) with a
capabilities approach, which recognises that individuals are not equal in their power to make ethical decisions, and 2) with a care ethics approach. The latter, according to Powell, has the potential to address both power asymmetries and the conflicting values of different actors, especially within contexts of decision-making processes, inasmuch as it is based on the idea that individuals should care for one another in the face of future uncertainties and as part of broader networked collectivities.

Finally, the last paper, submitted by Professor Simon Rogerson from De Montfort University in the UK, outlines the foundations for a new vision of the digital age, one that is fully grounded in virtue ethics. Rogerson argues that such a vision is needed in order to address and overcome the ethical challenges that result from digital inequalities and from the ways in which the digital environment operates. For Rogerson, the task of re-imagining the digital age depends 1) on promoting digital ethics education, so that students can learn how to use digital technologies both virtuously and responsibly, 2) on embedding virtues and ethics into the design of digital products, and 3) on ensuring that internet governance, in order to be virtuous, is based on the participation of multiple stakeholders, including users themselves, in decision-making processes.
1. Character and the Computer Scientist

The goal of this position paper is to criticize how and for whom the current digital environment operates – and how character education of computer scientists is a way forward to creating the digital environment that values all stakeholders. The current focus of Internet companies on the easy-to-quantify consequences miss marginalized stakeholders whose interests and rights are diminished in the current environment.

Digital Environment - Who Matters?

Powerful actors and their interests dominate the digital environment. The ISPs that route the traffic; social media sites that give a platform for individuals and companies to connect; multinational corporations who sell their products and services; the billion-dollar companies that host the internet’s content on their servers. These actors command the bully pulpit as to the norms of the digital environment including what information is collected, who gets access to that information, and what type of content is available and monetized.

The current digital and data ecosystem is designed around these dominant players, and their interests are embedded in how the internet works in two ways. First, customer facing firms make money by engagement, therefore keeping users online is critical to being profitable. A social network companies ‘customers’ are the advertisers they solicit to buy ad space on their platform; users are the supply for those advertisers. In many ways, social network platforms should be seen as a way for advertisers to connect to users; these customer facing firms are the honeypots designed to lure users to their site. Because of this structure, content that is inflammatory or false, or just maddening, is more likely to keep people engaged.[1]

Second, the more hidden data ecosystem, comprised of ad networks, data aggregators, and digital marketers, relies on a steady stream of intimate knowledge about consumers in order to conduct hyper targeted manipulation later. These third parties create a second incentive to customer facing firms to lure users to website with content in order for those users to be tracked and surveilled.[2]

Combined, we have a digital environment where third party marketers create an unhealthy demand for consumer engagement and information.[3]

**Internet Corporations - Who Matters?**

For Internet corporations, this environment creates a dangerous incentive for poor behavior. An important tension for corporations is whether to focus on short-term profit maximization versus the long-term value for the company - where the long-term value of the company is best achieved by managing the firm for all stakeholders.[4] The digital environment exacerbates this tension in that powerful, dominant actors are current disproportionately prioritized and less powerful users are marginalized as stakeholders.

Consider gamergate, revenge porn, deep fakes, cyberbullying, disinformation campaigns, or doxing, a quick search shows zero management or business ethics scholarship on these topics. No scholar, whose job is to critically evaluate the actions of business, currently has the marginalized stakeholders who are disproportionately impacted by these phenomena: the users, the subjects of the bullying, the victims of the disinformation campaigns. Instead, we focus on ad network revenues and bitcoin. It is within this layer of tensions - the internet corporation and the digital environment - that we ask computer scientists to create new services and products and is why we need to focus on the development of the computer scientists in order to navigate these tensions.

**Computer Scientists - Who Matters?**

What makes a good computer scientist? Given our current digital environment, a ‘good’ computer scientist may be one that fits or survives in the digital environment; a person that is successful at helping internet corporations achieve their goals. Given the current pressures of the digital environment, this would be a mistake.

Developing the identity of a computer scientist as a computer scientist is critical to develop a separate set of obligations and virtues apart from the current digital environment. Similar to civil engineers who design the infrastructure of our physical world, computer scientists develop our digital world and should have a similar set of virtues instilled in them as they develop in the classroom. In other words, we need to cultivate the identity of a computer scientist as one who cares about the larger community, lives with integrity, designs with an understanding of who is marginalized, and sees excellence as designing solutions for a better digital environment.[5]

Curriculum for the computer scientist would address:
1) Who gains power from this program? Who is marginalized by the program’s implementation?

---


2) What does it mean to be a computer scientist? Who should be considered in designing and developing a program/Al?

3) In choosing an algorithm to fit the data, who benefits and who is marginalized by that choice of algorithm? Whose data is ‘seen’ accurately by that algorithm and whose data is ‘missed’?

4) What voices are missing or inaccurate in the data? Whose voices are loud in the data?

5) Who deserves to be categorized by a computer program and who deserves to be treated like an individual?

6) Whose identity is being overlooked by the digital environment and internet corporations? How is the design of the AI/program reinforcing who is seen, heard, and considered?

The goal would be to have a set of professional obligations independent of the firm where the computer scientist works. We need to explicitly design curriculum to arm computer scientists with the identity and tools to navigate the tensions within the current digital environment.
2. Algorithmic Explainability and Practical Wisdom

Introduction

Our online interactions are increasingly mediated by algorithmic systems. Natural language processing algorithms analyse sentiments expressed in social media posts; recommendation systems filter and select products, news, and videos for us to consume; and ubiquitous sensors implemented in smart devices record and analyse biometric data to nudge us towards purportedly healthier habits.

While there are many opportunities with these sociotechnical developments, the functionality of algorithmic systems can often be highly complex and difficult to interpret. We can refer to this as the problem of algorithmic opacity, and identify several (non-exhaustive) reasons why it occurs.

First, the vast data repositories that store digital records of our online interactions, which are used to train the algorithmic systems (e.g., machine learning algorithms), often contain complex features that bear little resemblance to the ordinary categories we employ to describe our daily behaviours.[1]

Second, the inner workings of the algorithm itself may be inherently uninterpretable, preventing users from understanding how the system reached a particular decision outcome (e.g., a film recommendation based on the user’s prior viewing habits). This is often known as the “black-box” problem.

In both of these cases, digital literacy is important. Ultimately, the degree to which an algorithmic system is opaque, whether because of the data it relies upon or the mechanisms by which it operates, is in part determined by the knowledge and understanding of the individual assessing the system. However, even if the individual is an expert in machine learning, there may be a third (non-technical) barrier—it may not be practically feasible or possible to access the necessary explanatory information, perhaps due to legal restrictions designed to protect commercial interests (e.g., intellectual property rights).

[1] Automated facial recognition systems that rely on convolutional neural networks are a good example of this issue, as the features they rely on typically bear little resemblance to ordinary categories, such as ‘nose’ or ‘eyes’ (for a non-technical introduction see [1]).
Practical Wisdom

Each of these issues of algorithmic opacity have ethical significance because of the impact they have on an individual’s ability to develop and exercise practical wisdom. If we understand practical wisdom as the knowledge or understanding that enables its possessor to “do the right thing” in a given context [2], then opaque algorithmic systems can clearly create epistemic barriers for moral agents. Such barriers inhibit the situational appreciation of moral agents by a) challenging their ability to identify salient moral properties inherent in a wide variety of online interactions, and b) placing informational constraints on the rational evaluation of different actions. As an example, consider the choice of a user to share sensitive information about their mental health with close friends online, without the knowledge that it could be used for the purpose of advertising or to assess their risk of self-harm [3].

This is not merely an issue for individual decision-making. The choices we make online also affect the choice behaviour of others. This is most clear in the case of recommendation systems that power social media platforms. For instance, when a user clicks on a video, likes a post, or spends longer looking at a particular photo, these behaviours provide relevance feedback to the system that is used to tailor future recommendations. However, the feedback is also used to shape recommendations made to other similar users, based on the platform’s proprietary definition of ‘similar.’ Although an individual decision, such as positively rating a film, may be judged as trivial in terms of the impact it has on shaping the choice behaviour of another user, the possible effects that emerge from the collective dynamics of these sociotechnical systems can be substantive [4].

Algorithmic Explainability

What can we do to promote more reflective and informed decision-making in online environments increasingly mediated by opaque algorithmic systems?

While there is no single answer to this question, one promising approach comes from the field of algorithmic explainability. In May 2020, the Information Commissioner’s Office (ICO) and the Alan Turing Institute released joint guidance, primarily targeted at developers and policy-makers, on how to deliver explanations to users affected by the outcomes of algorithmic systems (e.g., automated decisions about which adverts users are shown) as well as how to document and explain the processes that lead to the outcomes (e.g., the training and evaluation of the algorithmic model) [5]. The guidance identifies six explanation types:

- **Rationale explanation**: the reasons that led to a decision, delivered in an accessible and non-technical way.
- **Responsibility explanation**: details of who was involved in the design, development, and deployment of an algorithmic system.
- **Data explanation**: an account of the data that was used in a particular decision and how, and the data used to train and test the algorithmic system.
• Fairness explanation: the steps taken throughout the design, development, and deployment of the system to mitigate the effects of algorithmic bias.
• Safety and performance explanation: steps taken to ensure the accuracy, reliability, security and robustness of the system’s behaviour.
Impact explanation: the impact that the use of the system has or may have on an individual, and on wider society.

Each of these possible types of explanation could plausibly have prudential value, depending on how and to whom it is delivered. For instance, providing a typical user with a rationale explanation about why an algorithmic system reached a particular decision is very different from, say, providing a professional auditor with a technical safety and performance explanation.

Therefore, while there exists a moral responsibility on individuals to reflect upon and rationally evaluate their choice behaviour online (e.g., which data they choose to share), if they face epistemic and legal hurdles in doing so, then there is a limit on how much responsibility they can exercise. If the developer of a system chooses an unnecessarily complex algorithm when a simpler (more interpretable) alternative exists, or fails to adequately report on how the algorithm was developed and evaluated, they have failed to act in a responsible manner. In a similar vein, improving the digital literacy of students requires education providers to be in a position to be able to deliver appropriate educational material, pitched at the right level and prioritising the most important types of explanation (e.g., rationale and data explanations). While technology developers could of course help support such initiatives, it is unlikely that many will do so without appropriate market incentives, clear regulatory standards, or other policy interventions.

The contours of our digital environment are constantly shifting as a result of the mediating power of algorithmic systems. Navigating this new terrain in an ethically mindful and reflective manner requires new skills, capabilities, and a cooperative approach to the cultivation of practical wisdom. An increased attention to the phronetic potential of algorithmic explainability could help us deliver on this collective responsibility.

References
3. Virtue and Vice in the Design and Deployment of Digital Technologies for Education

Introduction

Thinking about digital technologies and ethics often focuses on popular digital platforms, such as Facebook and Twitter. Risks to privacy and autonomy posed by these platforms and by big data are often discussed. These are important topics. However, it is also important to think about ethics in less high-profile applications of digital technology. In this presentation, we shall use virtue theory to reflect on ethical questions about digital technologies in education.

Virtue and Vice

Virtue is the possession of good or excellent character traits; vice the possession of morally flawed ones (Ameriks & Clarke, 2000). Virtue involves a combination of multiple states of affairs in a person, including having the right feelings, intentions, and motives. It also requires having practical wisdom or phronesis (Harrison, 2016) - the capacity for discerning or judging the right actions (Hursthouse, 1999). Vice occurs when any of these components are lacking. Instead of focusing entirely on moral rules or on consequences of actions, virtue theory focuses on character traits, such as compassion, benevolence, fairness, and honesty. Reflecting on these character traits can provide insight into the proper moral responses of the entities deploying technologies and of the individuals subject to them (Vallor, 2016). To explore this, we have chosen an example close to our own core business: the use of digital technologies in education to supervise, monitor, and assess students.

Education Technologies

Digital technologies are very much part of the classroom at schools and universities. Increasingly, we are seeing the use, or proposed use, of these technologies not merely to provide educational content but also to perform more controversial functions. Consider three examples. First, some grading algorithms use data in a predictive way to assign students grades for assessments they have not done (Hern, 2020). These predictive grades can affect students’ educational prospects. Second, facial recognition tools are
being developed to monitor school attendance, classroom attention and student movement around a campus (Alba, 2020; Lovell, 2019). Third, online exam proctoring technologies use AI to closely monitor, record, and analyse student behaviour during exams (Satariano, 2020). Though online proctoring is increasingly used, some students and educators condemn it as Big Brother-like (Coghlan et al., 2020). We think there are three perspectives from which to think about virtue theory arising from the use of these technologies.

**Educational Institutions**

The first perspective focuses on the educational institutions deploying digital education technologies. Virtues we might expect from a school or a university include compassion, honesty, and fairness towards students. While educational institutions may act with good intentions, good intentions alone are not sufficient for virtue. To be virtuous, a number of things need to align. For example, good character requires not only the right intentions but also the practical wisdom to discern the right use of technology. Practical wisdom here would include being able and disposed to examine carefully the possible impact of technology on all the students.

Consider the grading algorithm used by the British grading agency, Ofqual, to predict the results of secondary students in the UK in circumstances where they could not sit their final exams (Hern, 2020). The agencies using the algorithm may have had good intentions – perhaps they wanted to treat students fairly by giving them the mark they deserved and to avoid undue grade ‘inflation’ (Adams, 2020). Some system was needed to do this. But by basing the mark on historical patterns of results, the algorithm treated individual students from lower income schools more harshly than students from higher income schools (“A-levels and GCSEs,” 2020). There was, moreover, no opportunity for students or their teachers to appeal their grades (Adams & Stewart, 2020). Subjecting already vulnerable students to the risk of poorer outcomes and without the right of appeal demonstrated flawed character. Whether or not the decision-makers from Ofqual dismissed or simply failed to consider these unequal impacts, they showed poor judgment of a kind that resulted in unfairness, which is also a vice.

**Students**

A second perspective focuses on the students. How should students think about these technologies? This question requires us to consider the character traits students need if they are to adequately understand when they are being wrongly treated by the use of education technologies, and how they should respond to them (Harrison, 2015). For example, virtue is needed to know whether students should speak up for themselves and their fellows against the use of these technologies. This requires phronesis or practical wisdom (Harrison, 2016). If these technologies do not treat them unfairly, then it does not manifest virtue to be strongly critical of the education authorities that use them. An attitude of condemnation may even manifest vice.
Again, the use of these technologies may have been motivated by good intentions, such as improving teacher insights, protecting students, and ensuring high standards of honesty in the student body. Yet these technologies also hold risks of bias, discrimination, and individual privacy invasion.

Students need judgement to recognise the importance to their educational institutions of the institution taking measures to, for example, keep them safe. They also need insight into feelings of discomfort and anxiety that they may have in response to such surveillance. Thus, students also need to develop skills of character in order to ascertain and balance the merits against the risks of these uses of technology. A student who has developed such practical skills will not be blind to the fact that they and their fellows have been unjustly treated. They may even choose to vigorously protest against the injustice (Busby, 2020). Because virtue involves such practical skills, it can, as Aristotle taught us, be practiced, nourished, and developed over time (Ameriks & Clarke, 2020). Virtue can also be thwarted or undermined by means outside influences.

**Educators, Students, and Moral education**

The idea that virtue can be nourished or undermined brings us to the third perspective on virtue and new digital technologies in education. This perspective focuses on both educators and students taken together. Educators do not merely teach content; educators inevitably also provide a moral education. Intentionally or unintentionally, they teach their students about virtue. These virtues include honesty, trust, and trustworthiness. We might think here about the use of exam proctoring technologies. Online proctoring using AI could, for example, be praised as nourishing student honesty, or criticized as tending to undermine the development of the virtue of trust.

Teaching academic honesty as a virtue should, we think, be supported. But encouraging students to think that it is justified to minutely inspect every expression that exam-takers make without their true consent may, even inadvertently, tend to teach a vice. AI-based exam invigilation may accustom students, who are often in the process of developing their characters, to excessive surveillance and invasions of privacy. Alternatively, exam proctoring technologies may teach students about the importance of academic honesty. And so on. Arguably, then, morally assessing exam proctoring technology is not a straightforward matter. On the contrary, making judgments of this complex type may turn not merely on knowing moral rules or calculating consequences, but on having and developing virtue, including sound practical judgment.

**Conclusion**

Digital technologies can both benefit and harm society generally and education specifically. Virtue theory guides us in thinking about the character traits of those deploying the technologies and those subject to them. Educators and educational institutions should seek to develop certain aspects of good character in students and to avoid promoting vice, including by modelling the right behaviour and attitudes. Therefore,
educators themselves should, to some extent, also look inwards, thinking seriously about their
own characters and attempting to practice and develop virtue and practical wisdom.

References


China’s Social Credit system has often been portrayed as a project of totalitarian oppression and discipline; one that aims to displace liberal forms of normative ordering centered on the idea of rule of law and replace these with authoritarian forms of ordering based on trust or reputation. Here the SCS is presented, alternatively, as a natural evolution of the ways in which the normative repertoires of liberal societies instantiate justice commitments under conditions of what Anthony Giddens characterizes as ‘high modernity’ — that is, conditions of extreme accentuation of the consequences of modernity.

In effect, for all the ways the SCS bears on aspects of human subjectivity — indeed, for how it seeks to appropriately recognize and represent individual attributes, and to foster the development of virtues of a civic nature —, there is nothing in the System’s conceptualization that cannot, as a matter of principle, be reconciled with liberal purposes. It is rather the case that, in conditions of high modernity, commitment to such purposes requires something like the SCS to emerge.

That is so, in particular, due to a phenomenon inherent to high-modern societies, which here will be called ‘the fact of articulation’ — and which can be divided into three aspects. The first is that the reflexive relationship abstract systems at the root of modern societies — institutional and technological systems in particular — hold with individual identity increasingly leads to the articulation of identity attributes in some more-or-less explicit material form. Second, in conditions of high modernity, this growing materiality of individual identity progresses exponentially, in terms of its pace and extension, as well as of the depth of its entanglement with the actual human attributes it — often distortedly — refers to. The third and final aspect has to do with the role that, particularly in Western societies, private technological corporations have been allowed to play with little restraint by State authorities: namely, one of leveraging the first two aspects above to extend their dominance over how people develop their individual identities in the information environment — in turn determining what values, virtues, and modes of life will flourish or flounder in contemporary societies.

As a fact, the articulation of identity attributes is inexorable, an unavoidable product of the working of large-scale abstract systems. One of the central characteristics of modern societies is indeed the emergence of such systems — systems, that is, which increase the complexities of contemporary life while also enabling people to navigate such complexities. This is so since trust in abstract systems enables people to bracket out concerns they would otherwise have. Yet, in their material instantiations, abstract
systems also constrain regardless of trust — from the network effects of social media platforms to the entrenched biases of artificial intelligence systems. As they do so, such systems encroach most fundamentally on the development of human subjectivity, making one’s identity attributes the domain of someone else.

Inexorable though the fact of articulation is, however, its detrimental effects need not be — if only regulatory mechanisms are in place to ensure the appropriate recognition of individual attributes. Interestingly, such mechanisms are a two-way route. Not only they are needed to protect individuals against the misrecognition of identity attributes by the material instantiations of abstract systems, but trust in abstract systems itself depends, in turn, on the appropriate articulation of attributes — of the virtues — of actors in charge of such systems. At a deeper level, trust is always trust in people. In understanding how central this reciprocal relationship between identity and trust is for the governance of modern societies, one can appreciate the importance of regulatory mechanisms that safeguard its proper development.

Now, in modern societies, the ultimate repository of trust is the state; being such a repository is, in effect, the defining role of the modern nation state. In stabilizing normative expectations concerning abstract systems — a role performed by the law — the state provides assurances that enables trust in such systems to be established. But this trust-validating role can only take place where the state and the law are themselves trusted. Interestingly, thus, the legitimacy of state power and the authority of law also depend on regulatory mechanisms that ensure the appropriate articulation of identity attributes. It was John Locke himself who observed the fiduciary nature of the relation between citizens and governments and noted that prerogative — the discretionary power to act for the public good — is “always largest in the hands of our wisest and best princes”. Yet, rarely does sovereign virtue, or at least the virtue of prospective sovereigns, lend itself to easy identification.

To date, the regulation of identity attributes has fallen chiefly to traditional, rules-based institutional forms, typically through the laws of defamation and privacy, but also more generally through the recognition of individual rights. The appraisal of merit and excellence — the recognition of virtues — has taken place largely outside of law’s empire. In high-modern societies, however, the fact of articulation calls for the revision of traditional institutional legal forms — while it also calls for new, embedded regulatory mechanisms, capable of responding in kind to the growing materiality of identity processes and the distortions thus portends. With regard to the state itself, such mechanisms could open new avenues for a certain democratization of identity processes at the root of state authority. But these are avenues that extend far beyond the state, paving the way for values that underpin Chinese society — values of harmony and self-cultivation, of correlativity between power, merit and self-respect — but that are not foreign to the political culture of Western liberal societies themselves.

It is in such a context that one needs to understand China’s Social Credit System — i.e.
as a much-needed response to the fact of articulation, and one whose complexity defies simplistic characterizations of the System as a rule-of-trust or rule-of-reputation antithesis to the rule of law. What the Social Credit System does is to institutionally reimagine a relationship — that between trust, identity and the law — which lies at the very heart of modern social processes. And it does so not to supersede but to rejuvenate the role legal institutions have traditionally performed in validating trust and the identity claims on which trust is based.

In doing so, the Social Credit System responds to the central problem of justice of our time, posed by the fact of articulation — namely: how to instantiate the relationship between trust, identity and the law in institutional mechanisms that enable people to appropriately evaluate and articulate the truth about their attributes, and the different forms of action on which these are based.
Challenges to character-based ethics have been put forward in recent years, most notably by empirical psychologists who have contributed to the so-called ‘situationist’ debate (Doris 2002, 1998; Alfano 2013). In what follows, I sketch how online technologies present a new and unprecedented ethical challenge, one that has the potential to shake character-based philosophy to its very core. Despite the difficulty of outlining a character-based ethics of digital well-being (Burr & Floridi 2020, Burr et al. 2020), I propose that the situation is far from hopeless. Character-based ethicists can – and I propose should – respond to the challenge of emerging technologies in two ways: first, they must be suitably ambitious in estimating the power of online technologies to transform and complicate current conceptions of human flourishing. Second, they must anticipate the seismic practical consequences of this re-evaluation, especially for those who are now coming of age in our new digital world (for example, 13 to 16-year-old adolescents). Both these tasks, I contend, will involve redrawing the existing parameters of character-based ethics, in an analogous way to how virtue ethicists convincingly responded to their situationist critics at the turn of the 21st century.

Cultivating Cyber Phronesis identifies three key dimensions in the ethics of online technology debate, and I want to start by offering some thoughts on how we could understand these dimensions as relating to one another. Each theme (1–3) focuses on three potential levers to advance our ability to live well with online technologies. The first concerns how we can learn to improve our ability to function in online environments, specifically how specific character traits can be instilled in young people to improve their digital well-being. The second concerns the design of online technologies, a theme that theorists have recently shown – prefigured by the situationist debate – how easily our ethical selves can be influenced by persuasive technologies (Dennis 2020a; Frank 2020; Lanzig 2019, 2018). The third lever is governance, that is, the power of policy makers, politicians, and other regulators to guide our interactions with online technologies. All of these levers of change offer ways to improve our online lives, but clarifying the relations between them illuminates how they can contribute to a theory of digital well-being that we so urgently need.

From the viewpoint of ethical urgency, it is surely right to focus on how to teach young people the skills they need to practically cope with the impact of today’s online technologies (Polizzi & Harrison 2020; Polizzi 2020a, 2020b). This is a battle that many of us confront on a daily basis. Furthermore, it is easy to imagine how these problems are most acutely felt by 13 to 16-year-old adolescents (the target demographic of the Cultivating Cyber Phronesis project), many of whom are in need of ethical guidance on
how to align their online activities with a flourishing life. Nevertheless, in my view, the second and third themes merit special attention. Most obviously, regulators are best placed to shape the future of online technologies through legislation, as was perhaps seen by the introduction of the GDPR in 2015. Aside from the actions of regulators, however, in my view the most informative literature in the ethics of technology consistently emphasises how the design of online environments is the most effective way to affect how we live online. This is especially important, I contend, because design features of online technologies have the potential to undermine any potential advances in a character-based approach to online technologies.

Over the last decade, Netherlands-based ethicists from the value-sensitive design tradition have proposed that our relationships with online technologies can be most effectively improved by rethinking of how we design digital environments (Brey et al. 2012; van de Poel 2012; van den Hoven 2019, 2015; van den Hoven et al. 2017). Many of these scholars have criticised existing approaches to the design of online architecture. Like Tristan Harris from the Center for Humane Technology, these scholars view many of the current design imperatives of online technology companies as directed at the capture of user attention and data extraction, often in order to maximise corporate profit. Moreover, today’s online environments are suffused with manipulative e-choice architecture, hyper-nudges, and gamification, all of which are designed to keep us hooked online. I have suggested elsewhere that these are in tension with our capacity to flourish, which should motivate us to think critically about how to design online environments (Dennis 2020a, 2020b, under review a & b).

Understanding the relationship between the three research themes outlined above suggests various connective affinities (practical, logical, and sequential). I have suggested that education is rightly viewed as a practical priority, governance holds most direct power, whereas design has greatest potential to improve our online lives. In addition to these general affinities, the connection of character education and design tracks key contours of the situationist debate. From this, I contend, we have much to learn. Just as situationists such as John Doris employed Darley and Batson’s infamous Good Samaritan study to undercut the importance of character-based concepts (1973: 100), so too recent work on the power of persuasive technologies puts pressure on the central concepts of a character-based ethics of technology (choice, autonomy, virtue/vice, etc.) While this presents obvious challenges, it remains to be seen whether understanding the tension between persuasive technologies and our capacity for online virtuous behaviour is possible. If it is possible, there is a chance it may follow the familiar trajectory of the situationist debate. Despite how psychological studies such as Darley and Batson’s were initially invoked to criticise the work of character-based ethicists, these ethicists then incorporated these criticisms, improving the plausibility and coherence of their subsequent work.
References


van de Poel, I. (2012). Can We Design for Well-Being? In P. Brey, A. Briggle, & E. Spence (Eds.), The Good Life in a Technological Age (pp. 295-306), Routledge.


Character education is often based on encouraging and cultivating personal virtues, in the Aristotelian tradition. It is important to remember that every individuals’ virtues are enacted in a social context, as virtue ethicists like Alisdair Macintyre remind us. Increasingly, this social context is the context of the digital, networked and interconnected world. Technology is no longer separable from social, or even inner psychological lives. We live in, and with, technology, and therefore the technological is also part of our social milieu. Therefore, as technology is imagined, built and institutionalized, these actions also become part of the social space in which we encounter ethics. As young people begin to consider how they build and create their own social world around them, they should consider how their participation in constructing and maintaining the digital world influences not only their own internal moral state but also serves to build a world in which virtue also connected with care of others, and development of their own and others’ capabilities. Beginning with these considerations and with the idea that technology is built as part of a social world or milieu, this brief paper suggests ways to consider ethics as values in action, focused constructively on the development of social world rather than only on the narrow consequences of technological developments.

Values in Action

It is helpful to think of ethics as values in action. This means that there is no particular point where you should start or stop thinking about ethics, because ethics is embedded in most of the decisions we take. In general terms, ethics provide a set of principles that help us to think clearly about rights, obligations and responsibilities. They are not only to do with the consequences of an action or a development, but unfold in relation to everything that we encounter in the social world. We distinguish these from ethical frameworks - ways of structuring decisions in the process of design and development of technology. In considering how ethics is practiced in digital societies, we suggest combining perspectives of virtue, capability and care.

Virtue Ethics

Virtue ethics offers an approach that is focused on augmenting and improving the self. To live well, we must develop our moral character and demonstrate virtues in our decision-
making and behaviour. Essentially, virtue ethics is concerned with questions such as "What is a good life?" or, "What does it mean to be a good person?". The familiar demands of "technologies for good" or "don’t be evil" speak to the idea that developers of technology might be acting to shape the technological world in a particular way – this is the premise of the recent documentary “The Social Dilemma”.

However, this may be too narrow a framing. Looking at virtue ethics alone ignores the social, cultural, economic and political structures that all of us embedded in. Consequently, this approach does not offer satisfactory answers to questions such as: What sort of practical reasoning or wisdom is necessary for developers to navigate the pressures and constraints of the broader economic, political and social contexts that they inhabit? Often principles have to be negotiated and at times compromised because things change and well, “that’s just life”. What might shape or constrain the choice of action?

**Capabilities Approach**

The capabilities approach recognises that personal principles may be compromised in order to cope with structural constraints. As such, it promotes the idea that ethical thinking is also a capability in itself and not a given for individuals – or intrinsic to some. The capabilities approach recognises that individuals are not equal in their power to make ethically consequential decisions or voice their concerns in the process.

The capabilities approach augments the internally oriented focus of virtue ethics on the moral capacities of the individual, by adding the importance of structural constraints. However, in both of these philosophical schools of thought, decisions are made by individuals and it is individuals that must take the responsibility and account for the constraints imposed by the broader social, political and economic contexts. This assumes, however, that moral virtues and capabilities map out a clear path that ought to be followed. The problem is that, even if we might have moral character, our own values and virtues might come into conflict in a decision-making process. This is because people are always entangled in a diversity of relations that hold contradictory values and conflicting demands and decisions are never taken alone.

**Care Ethics**

Hence the need for a complementary, third approach – care ethics. This school of thought pays attention to the value conflicts and contradictions, offering a way to deal with failures that does not entail merely accepting vast amounts of personal guilt. Care ethics recognises that relationships are central to being human because they enable individuals to face uncertainties of the future – together is better than alone. One of the main tenets of this approach is that it places the emphasis on an engaged, active agent who acknowledges that she is entangled in a broad web of relationships and is aware that she requires the constant negotiation of disparate and often conflicting demands and obligations.
Instead of thinking of actors as separate individuals that happen to form communities or other social arrangements, the logic of care acknowledges that we are never separate individuals, but are composed of our many memberships, relations and social commitments that span our lives.

It is important to ask “what are the conditions and possibilities of care?” “What should we care about?” and “How do we [even] begin to care?”. Since it is impossible to care about everything, care is necessarily a selective model of attention and the choice of what or whom to care about is a kind of expression of power. The asymmetric expressions of power, however, are balanced by the fact that caring for someone or something deeply makes us vulnerable. This vulnerability is a necessary component of a relation of care – by caring we expose ourselves to the demands of those we care for.

Thus, care-full decision-making requires reflection, negotiation and openness to discussion – no decision is made individually. There must be a collective process of deliberation, because if we are going to face uncertainties of the futures we are creating through innovation, these futures are best faced together – rather than alone. This summary of the related strands of virtue, capability and care ethics outlines how these different ethical action positions are experienced, opening out the space to consider how to exercise these together in a society where technology is wound into all spheres of life.
7. Re-Imagining the Digital Age through Digital Ethics

The world has changed. In May this year, Mobile App Daily published its 2020 technology trend forecast. It demonstrates the world is now digitised through, for example, 5G, clouds, AI, algorithms, augmentation, machine autonomy, data analytics, edge computing and the Internet of Things. This digitisation of everything requires a greater emphasis on, what we should now call, Digital Ethics. If not, then a very bleak, discriminatory world beckons. It would be a world of privileged digital natives and an underclass of digital outcasts, a world of danger, domination and despair.

There now exists global deep-seated dependency on digital technology. By way of illustration, look how the social glue has come unstuck during the pandemic and we have turned to digital technology to allow us to live and keep us connected (Rogerson, 2020). Communication channels provide information about the latest developments, advice and restrictions. Social media keeps social groups and families emotionally together. Online outlets provide the products and services we need in our everyday lives. For digital natives, the move to the virtual is plausible and possibly pleasurable but for digital outcasts the move is fraught and frequently frightening.

Everyone has moral obligations and responsibilities in ensuring the Digital Age is inclusive and empowering rather than exclusive and constraining. Established rules may offer some guidance as to the correct path, but such rules can easily become the instruments of blatant superficial compliance which at best is problematic and at worst immoral. It is virtuous action that promotes an ethical digital age. A person is not virtuous because their actions comply, perhaps mindlessly, with established rules, the action is virtuous because of the virtuous nature of the person who performs it (Wyatt, 2008).

Virtuous action must occur in different ways crossing traditional barriers and challenging established norms. Three drivers; top-down, middle-out and bottom-up, come into play (Rogerson, 2015). Top-down drivers are typically impositions by bodies of authority which dictate where resources should be placed to achieve some overall goal. Middle-out drivers involve all those within, for example an organisation, who are empowered to initiate change, support it, propose new ideas, and innovate. Bottom-up drivers emanate typically from grassroots collective-action resulting in a widespread change. Boyle (2009) suggests top-down drivers provide political direction, middle-out drivers are the focus of change teams, and bottom-up drivers are the voices of citizens.

Education and Learning
Churchland (1996) explains that the development of moral character in children takes time – “time to learn how to recognise a large variety of prototypical social situations, time to learn how to deal with those situations, time to learn how to balance or arbitrate conflicting perceptions and conflicting demands, and time to learn the sorts of patience and self-control that characterise mature skills in any domain of activity”. The formality of school and the informality of home are equally important, particularly in the early stages of this journey to moral maturity. It is a journey which starts the moment a child is born, continues through childhood into adolescence and finally into adulthood.

The global population has a collective view of an acceptable digital age. This population will include those who have suffered directly from unethical situations. Examples include: the hesitant user of a web-based public service who is the victim of poor system design; the junior software engineer who is pressurised into unethical, yet commercially valuable, action by an internet organisation; and as well as the vulnerable young adult who is the victim of incessant cyberbullying. Digital ethics education and awareness must develop the individual’s confidence and skills, through lifelong learning, and so provide the tools to enable everyone to act responsibly and ethically. Discussion, dialogue, storytelling, case study analysis, mentoring, and counselling are examples of techniques that can be used to nurture practical wisdom and insight which will lead to virtuous citizens of the Digital Age.

The Digital Environment

The complex interrelated ethical and social issues within the digital environment must be addressed during the digital technology process and embedded within the digital technology product (Rogerson 2010). Process concerns the activities of digital technology professionals when undertaking planning, research, development and service/product delivery. The aim is for professionals to be virtuous in Aristotelian terms. In other words, a professional knows that an action is the right thing to do in the circumstances and does it for the right motive. Product concerns the outcome of professional endeavour and the potential impact of these products on people, society and the environment. The ethics focus of the product perspective is technological integrity from, for example, a Kantian or utilitarian perspective. This can be addressed by embedding ethics within digital products themselves.

For example, internet corporations could be proactive rather than reactive in promoting empowerment and safety for internet users. This virtuous action could lead to a new feature, information provenance, added to search engines. Information provenance would fix the origin and network of ownership thus providing a measure of integrity, authenticity and trustworthiness. It would provide an audit trail showing where information originated, where it has been and how it has been altered. In this way people would be able to consider how much credence they would give to a piece of information before acting upon it. In the Digital Age there is a moral obligation to address information integrity. Information provenance offers a normative instrument for turning the moral obligation of addressing information integrity into ethical practice.
Whilst the imperative for organisations must be virtuous action, this, as Wyatt (2008) explains, must sit comfortably with other ethical analyses. This is likely to lead to ethically justifiable actions and outcomes. This approach can be summarised by four statements:

- Consider if your acts are worthy of the ideal digital technology professional (virtue ethics)
- Consider who is affected by your work (utilitarian)
- Analyse how the least empowered will be affected by your decision (Rawlsian)
- Examine if other humans are being treated with respect (Kantian)

One way to promote virtuous action would be to establish ethics circles. An ethics circle is a means of providing support to individuals who are engaged in ethically sensitive decision making, raising general awareness of ethical issues and acting as informal staff development in this area (Rogerson, 1996). The circle comprises a small group of people who do similar or complementary work meeting on a regular basis to identify problems, to analyse the causes, to recommend solutions and, where possible, to implement those solutions themselves.

**Governance**

Governance must address both the process and product dimensions of digital technology development. It must promote a sense of obligation in professional developers, thus ensuring that digital technology products and services are fit-for-purpose. Governance implies a system in which all stakeholders have the necessary input into the decision-making process. A broader view needs to be taken in defining the stakeholders of the Digital Age. Stakeholders should include those:

- whose behaviour/work process will be affected by the development or delivery of digital technology
- whose circumstance/job will be affected by the development or delivery of digital technology
- whose experiences will be affected by the development or delivery of digital technology.

Governance should have its foundation as delivering ethical, efficient and effective digital technology. These three factors must be multiplicative rather than summative.

**Conclusion**

There is a need to develop a new vision for Digital Ethics which is theoretically grounded but pragmatic, so that industry and government will engage, accept and embrace this as a modus operandi. Digital Ethics can be defined as integrating digital technology and human values in such a way that digital technology advances human values, rather than doing damage to them. It therefore must include the formulation and justification of policies for the ethical use of digital technology, and carefully considered, transparent and justified action leading to ethically acceptable digital technology products and services.
(Rogerson, 2011). This form of Digital Ethics is transdisciplinary, drawing upon many disciplines including ethics, computing, psychology and education. It is the grounding which enables us to focus on the long-term needs of everyone rather than the short-term interests of the few. Digital education from an early age should engender virtue, wisdom and humility as well as instrumental skill and technological prowess. There will then be a digital age created by everyone for everyone. We will make a difference through challenging complacency, indifference and ambivalence regarding ethical digital technology by those involved in any aspect of planning, funding, researching, developing, implementing and using digital technology.

References


About the Author

Simon Rogerson became Europe’s first Professor in Computer Ethics in 1998 and in 2010 became lifetime Professor Emeritus in Computer Ethics at De Montfort University, UK. His early career was in industry as a technical software developer. He was the founding Director of the Centre for Computing and Social Responsibility (CCSR), launching it in 1995 at the first ETHICOMP conference which he conceived and co-directed until 2013. He was the founder and is the current editor of the Journal of Information, Communication and Ethics in Society. He sits on several international ICT-related advisory boards and has served on governing bodies in education and ICT. For his leadership and research achievements in the computer and information ethics interdisciplinary field he was awarded the IFIP-WG9.2 Namur Award in 2000 and the ACM SIGCAS Making a Difference Award in 2006.
BIG QUESTIONS AND FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

The Jubilee Centre would like to thank all the invitees who presented at the webinar. Their presentations, based on the position papers above, generated a lively and fruitful discussion. Below are a number of big questions that were raised at the webinar – questions that have research and practical implications that should be explored in the future:

- Can digital technologies promote and enhance virtue in society?
- What can we learn about character and virtues from the ethical challenges that digital technologies present for human flourishing?
- How can we challenge the structural constraints and ethical challenges that are inherent in how the digital environment operates and find ways of re-imagining such an environment in line with values of justice and equality?
- How can we ensure that the character development of computer scientists and tech developers is part of a more holistic approach to re-shaping the digital environment, one that will enable such professionals to act on, and design digital products with, wisdom and in line with ethical principles?
- What regulatory mechanisms, if any, should be implemented by governments to ensure that the digital environment is driven by virtues and ethical principles alongside economic interests?
- Can digital technological innovations improve current approaches to character education?