

Citizenship

in a Networked Age

An Agenda for Rebuilding Our Civic Ideals



TEMPLETON WORLD
CHARITY FOUNDATION



UNIVERSITY OF
OXFORD

Authors

Dominic Burbidge researches decentralisations (technological and governmental), citizenship, and social trust. He is a Research Director in the University of Oxford, an External Advisor to Templeton World Charity Foundation, and Director of the Canterbury Institute. Dominic is the report's lead author.

Andrew Briggs is Professor of Nanomaterials at the University of Oxford. He is Honorary Fellow of the Royal Microscopical Society, Fellow of the Institute of Physics, and Fellow of St Anne's College and of Wolfson College, Oxford. He has more than 600 publications, with over 25,000 citations. His book with Roger Wagner describing how science thrives in cultures concerned about the big questions of humanity has spawned a documentary film and a series for children. With Michael Reiss, he is writing a book that explores the complementary roles of scientific insight and spiritual wisdom in the promotion of human flourishing. From 2002-2009, he directed the UK Interdisciplinary Research Collaboration in Quantum Information Processing. Current work in his laboratory develops machine learning for measuring and tuning quantum devices, with results that greatly exceed what is possible for humans.

Michael J. Reiss is Professor of Science Education at University College London, Visiting Professor at the Universities of York and the Royal Veterinary College, a Fellow of the Academy of Social Sciences, the Society of Biology and the Royal Society of Arts, a member of the Nuffield Council on Bioethics and a Priest in the Church of England. He is President of the International Society for Science & Religion and of the International Association for Science and Religion in Schools and was previously Director of Education at the Royal Society and a Specialist Advisor to the House of Commons (education) and the House of Lords (bioethics). Much of his recent writing has been concerned with the aims of education and with how we can have curricula and approaches to teaching in schools that enable students to flourish.

Background

The report is the product of a two year research project at the University of Oxford, generously supported by Templeton World Charity Foundation. The report was published in 2020.

Consultation process

The research for this report benefited from an extensive international consulting process with leading experts of technological change and human flourishing. In addition, a local advisory group convened multiple times in the University of Oxford to refine the report's scope of inquiry. We are grateful to those who gave generously of their time at consultations conducted in the Expanded Reason Congress in Rome, Italy, in September 2018, at a group of selected education leaders attending the OECD 2030 Future of Education meeting in Paris, France, in October 2018, at the Global Innovations for Character Development conference in Nairobi, Kenya, in October 2018, and at the Fuller Theological Seminary in Pasadena, California, in October 2018. We also convened a dedicated conference on science, philosophy and religion, from the perspective of the Abrahamic faiths, in Aiyia Napa, Cyprus, in November 2018. We are additionally grateful for consultation meetings held during the writing of the report. These were hosted in the Cyber Policy Center of Stanford University, USA, the Legatum Institute, UK, the Abigail Adams Institute and Harvard University, USA, and the Center for the Study of Statesmanship, Catholic University of America, USA, all in 2019. A detailed review of the research ideas gathered through the initial consultation phase is available in the publication, 'Promoting Human Flourishing through the Best of Scientific Insight and Spiritual Wisdom: A Global Engagement' (2019).

~ Contents ~

Executive Summary	6
Introduction	12
1. Efficiency is a good aim but not enough	19
a. The networked age.....	20
b. Big data & big computation.....	25
c. Automation, algorithms & artificial intelligence	32
d. Human intention & civic ideals	39
2. Where citizenship starts.....	43
a. Our human space, going forward	44
b. The citizen vs slave distinction	49
c. Ethics through the navigation of competing desires	55
d. Rebuilding our citizenship	66
3. The changing relationship between time & attention.....	67
a. Being only one citizen	68
b. The friendship paradox & FOMO (Fear Of Missing Out)	75
c. Time & attention as scarce resources	86
d. New categories of community, platform & institution in a networked age.....	88
4. Redefining public & private	98
a. Why do we have a private realm?	99
b. What is our public realm?	104
c. End-to-end encryption	114
d. Privacy as conscience & space for personal moral development	129
5. Algorithmic vs democratic decision-making	134
a. Defining algorithmic decision-making.....	135
b. Defining democratic decision-making	140
c. Algorithms & justice	148
d. The distinct role of public reason in democratic deliberation	153
6. Conclusion.....	162
Bibliography	172
Acknowledgements	185
External reviewers.....	187

Citizenship in a Networked Age

An Agenda for Rebuilding
Our Civic Ideals



Executive Summary



THE CHARACTER OF CITIZENSHIP IS CHANGING IN THE NETWORKED AGE. Citizenship is not just a legal nicety—it encompasses the relationship between an individual and the society of which they are a member. It entails a spirit of public service, justice, neighbourliness, democratic participation, and moral reasoning. Participation in decision-making is central to our collective citizenship, both as a privilege and a responsibility. Digital networking, however, is irreversibly changing how we interact with each other, and in turn machine learning is becoming inseparable from its development. *Search optimisation*, for example, has evolved into *influence optimisation*. Learning what a user's preferences are has changed to targeting what a user's preferences should become. Machine learning is moving from learning how to *classify* to learning how to *decide* and *optimise*. And yet in the midst of all this, human participation in decision-making endures as a necessary and core feature of civic engagement, and therefore our citizenship itself. Citizen involvement in collective decisions remains essential to progress, both because progress requires principles of justice, and because it needs the kind of shared commitment that can only be earned through our civic engagement.

HUMAN LIFE IS CIRCUMSCRIBED BY MORTALITY IN A WAY THAT MACHINE LIFE IS NOT, AND HUMANS ARE SELF-AWARE OF THEIR MORTALITY IN A WAY THAT MACHINES ARE NOT. This, together with our conscience, gives humans a distinct capacity for moral decision-making which no machine can replace. Mortality is an end point that is hard to reconcile with the deeper human aims of happiness or living a life well-lived, which means humans have developed a way of detecting moral fundamentals even though they might die at any time. Machines lack this tension of seeking deeper goods while knowing one is going to

die. They do not develop moral reasoning as such. In this networked age, the challenge is to protect and nurture existing human methods of moral decision-making, something that can benefit greatly from machine optimisations of technical aspects *insofar as they are in service of human judgment of the moral whole*. For the pursuit of human flourishing in the networked age, the choice is not between citizens and machines. It is about identifying and protecting human uniqueness for moral decision-making.

Recommendation 1

Identify and protect human
uniqueness for moral
decision-making

Recommendation 2

Nurture the complementary
skills of humans and machines
for collective decision-making

AT THEIR BEST, THE TECHNOLOGIES OF THE NETWORKED AGE WILL CONTRIBUTE TO BETTER DECISION-MAKING IN PURSUIT OF THE IDEALS OF GOOD CITIZENSHIP. The outcomes will only be as good as the values which motivate the process. Civic engagement is needed because there are few goals worth pursuing on which everyone agrees. The effect of new technologies is generally not to raise new moral and social issues which did not previously exist, but rather to give a new focus and urgency to norms of behaviour which long predate modern communications. It can sometimes look as though the distinction in former societies between citizens and slaves has a modern parallel in the distinction between data companies and consumers. The asymmetry is not absolute, but there does seem to be a shift of influence from those whose data is harvested to those who then make use of this data for their own purposes, whether commercial or political. The economic success of big data industries is bound up with the ownership and privileged use of that data. Regulation will need to be based, however, not on the relationship between corporations and consumers but on the relationship between democratic institutions and citizens. Humans are making rapid progress on the methods of machine learning for multi-objective optimisation. In order to use these methods in support of collective decision-making, there needs to be sufficient agreement about the values that these optimisations are oriented around. This provides a new context and motivation for developing our principles as a society and defining what human flourishing means.

Recommendation 3

*Engage in consensus-building
about civic ideals for a
networked age*

TECHNOLOGY CHANGES HOW PEOPLE INTERACT. If there was once a time when most people knew few others outside their own village, now it is often hard to tell from where in the world a given message comes. The distance between an individual's private persona and public projection can be stretched almost completely. The fear of missing out on the glamorous lifestyle and vast numbers of friends others seem to enjoy can be as overwhelming as it is deceptive. And while these networks seem to encourage a certain togetherness, they are not naturally conducive to the moderation of viewpoints. On any issue where there is incipient polarisation, machine learning can target a user with a news story or opinion slightly more extreme, thereby reinforcing the polarisation and capitalising on bias.

NEW KINDS OF COMMUNITIES AND INSTITUTIONS ARE BEING FORMED IN THE NETWORKED AGE. The pressures are to connect people of the same tastes, the same biases, the same political leanings, and the same generations. Human attention has become a precious resource, and where your attention is, there will your desires be also. In the midst of this new market for attention, we have to realise that giving quality attention to others is the most important form of self-giving we can engage in as citizens. In other words, the civic burden is not so much on being a good speaker but on being a good listener—finding the way to understand others and what they say, and in so doing go against the tide of our polarisation.

Recommendation 4

*Teach listening as a
civic virtue*

PRIVACY PROVIDES A SPECIAL SPACE FOR CITIZENS TO REFLECT ON WHAT MATTERS MOST TO THEM. Privacy is widely recognised as a fundamental human right, based on principles of human dignity. It is linked to other rights such as equal treatment and free expression. The networked age gives unprecedented opportunity and even encouragement for individuals to voluntarily forgo much of their privacy. In turn, the distinction between the private realm and the public realm becomes increasingly blurred. End-to-end encryption gives some protection to individual privacy, rather like not reading other people's letters, but the metadata can still be harvested. Machine learning can help in the fight to filter out fake news, but ultimately the reader must be aware of the need to form their views independently and thoughtfully, often in private first. In developing civic ideals, digital privacy provides space, therefore, for society's moral development. In particular, it protects against the demand for instant response. In private, individuals and groups are able to self-initiate their search for truth, meaning, and purpose.

Recommendation 5

*Maintain distance between
thought and speech*

Recommendation 6

*Promote the value of privacy
for personal moral development*

DEMOCRATIC DECISION-MAKING IS ALL ABOUT AGREEING ON A WAY FORWARD IN THE CONTEXT OF DIVERSE AND OFTEN CONFLICTING INTERESTS. These decisions involve both technical and moral components. Machine learning surpasses human abilities in an increasing range of technical decisions. Democratic decision-making, however, involves so far as is possible building consensus among people as part of the process of deciding what is just and moral. A crucial ingredient in democratic decision-making, which machines will never be able to replace, is the moral basis to society-wide decision-making. At its best, under wise and inclusive leadership, democracy facilitates shared commitment to the agreed course of action. Good democratic decision-making needs to be based on sound technical and moral reasoning, with shared ownership of the outcome.

Recommendation 7

Revalue democracy in terms
of the ability to bring about
social unity and trust



In the conclusion of this report, we summarise how these recommendations build from the evidence and arguments of the report's individual chapters.



Introduction

Citizenship 
in a Networked Age

Introduction



The networked age can be defined as that period of history where interpersonal distance was eroded and decision-making became human-machine cooperative. This report asks the following question: in the networked age, what is the space for civic engagement? Both components of the networked age have profound implications for the political, civic and social dimensions to what counts as human flourishing and challenge the fundamentals of our civic engagement. As interpersonal distance is reduced and decision-making becomes human-machine cooperative we lack a robust account of *good citizenship aimed at human flourishing*, and that is the gap this report seeks to fill. The report's contribution is to diagnose current social and political changes to our decision-making landscape in the light of rapid technological transformations and to make specific, practical recommendations to re-orientate the networked age towards human flourishing.

Changes currently afoot as to what our human community *is* and *is for* require us to look anew at our civic ideals. Each of us must take seriously the dilemma of being *only one citizen* as the human community grows and becomes ever more interconnected. As my community grows, I may not feel I am growing personally; as my community becomes more interconnected, I may not feel as interconnected as everyone else. Analysis of the internet, digital technologies and artificial intelligence has so far tended to look at how they change society and the world as a whole, underappreciating the fact that in the midst of any change I remain just one person, just one citizen. From that individual citizen point of view, progress in the expansion of the internet, digital technologies and artificial intelligence may mean that

my own contribution and *my own sense of self* reduce, and that reduction accelerates in line with accelerated erosions of interpersonal barriers. To the lone citizen there is, as yet, no guaranteed shape to our networked citizenry; we must reach into the wisdom of the ages in the midst of these upheavals to give a better account of what our human community *is* and *is for*.

The networked age is not yet fully emergent, making us at times appear more like a set of specific users of services and platforms and, at other times, more like a genuine collectivity or citizenry. We entered the networked age on *transactional terms* with digital technologies that appeared to make our lives easier, bit by bit. We stay on *more permanent terms* with digital technologies that share aspects of our being and our identity, and make numerous promises about future efficiency gains and data consolidations. The *permanent* starts membership of a new kind of collectivity. We are each travelling the road from *transactional* to *permanent* at different speeds, with some far behind and some far ahead, though the general direction is clearly one-way. This report, in setting out *an agenda for rebuilding our civic ideals in the networked age*, joins a wider discourse pointing towards our emergence as a collectivity in the midst of technological advancements. What the report advocates is that our coming of age be an *emergence as a citizenry*, guided not just by regulations but also civic ideals that give effect to long-term principles for our new habits and habitat.

An important test case is the introduction of algorithms and artificial intelligence into our public sphere decision-making, which is set to affect deeply our spirit of public service, justice, neighbourliness, democratic participation and moral reasoning. While numerous institutions and experts have focused on the changes technology and artificial intelligence are having on our mental development, our economy and jobs, and even the survival of the entire human race, there has been less analysis of what the changes mean for society's moral framework (the *normativity* of our citizenship).

Our civic ideals are being tested in the networked age, an age which threatens to take decision-making out of the hands of individuals and to break up older understandings of community. Increasingly, we are following the decisions and recommendations of algorithms and artificial intelligence to understand what would be in our interests. We communicate in siloed groups that are purpose-built around particular aims or views rather than the common good. Our networks of communication can be created or collapsed at any time, and have reduced regard for members' wellbeing over time. While there are, of course, clear and often very substantial benefits to technological developments for efficiency of service delivery, ease of communication and availability of information, it is imperative to keep an eye on the ethical dimension to these changes. We need to offer ideas on complementary human-machine collective decision-making that prioritises humans' unique faculty for moral reasoning.

A number of changes to our civic engagement are already becoming apparent:

- a. A growing commitment to algorithmic decision-making in public administration due to its presumed objectivity, expertise and rapidity. In turn, this can frame public participation by citizens as relatively sentimental, inept and slow, leading to greater outsourcing of moral and judicial decision-making to algorithms, which are often presumed to be impartial. Citizen decision-making is presumed to be biased and individualistic.
- b. Reduced habit among citizens of suspending judgment or developing political views in closed settings before sharing with the wider public.
- c. Declining natural forgetfulness for bad or incorrect things said.

- d. Curtailment of one's social network away from those with alternative viewpoints, especially during important political moments.
- e. A weakening ability to boycott online services that hold one's data and history, and do not charge money for the service.
- f. Increased quantification of public goals and application of utilitarian reasoning for balancing competing interests, as opposed to democratic deliberation over human values.
- g. Growing disconnect between global, cosmopolitan views and local, traditional living, accompanied by increased anxiety on both sides over who the ultimate decision-makers really are.
- h. A burgeoning plurality of cybergovernance institutions and a weakening of the authority of national laws, in turn making the target of democratic advocacy less clear.
- i. Increased hate speech and misinformation, with civic engagement narrowly focusing on chastising those who have spoken hatefully or falsely.

These changes in some ways build on 20th century marketization and urbanisation of human relations, but they also deviate in significant ways through extraordinary personalisation of the laws of supply and demand and a decentring of the urban experience through the creation of multiple overlapping social networks. Crudely-speaking, 20th century industrialism helped everyone buy the same things; 21st century personalisation of the laws of supply and demand gives everyone their own, individualised economy, in line with their online data profiles. In threatening our ability to be citizens and act in pursuit of a common good, these trends risk a breakdown in social solidarity and a growth in arbitrary authority. In this sense, distrust between the “haves” and the “have nots” is not simply an economic question of

appropriate distribution but a growing existential question about who the decision-makers really are.

To help navigate these trends, the report brings to bear the following unusual questions to the more usual concern over technological change:

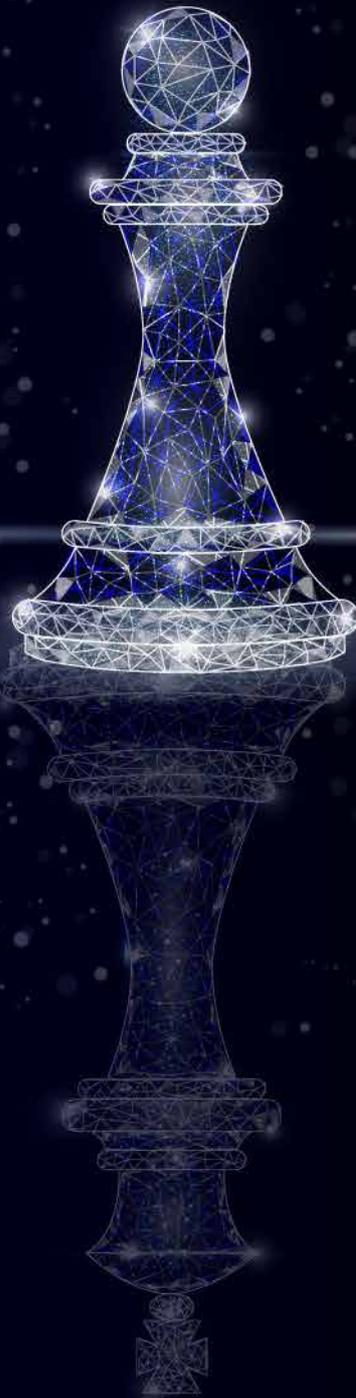
- a. Do we have a basic understanding of human nature and human flourishing that can help us decide if technological changes are good for our moral reasoning or not?
- b. Can we distinguish between those uses of algorithms, machine learning and digital technologies that are helpful practical optimisations, and those that unhelpfully emasculate genuine moral and democratic deliberation?
- c. What civic virtues are required for the 21st century? How can I be self-giving towards others in the networked age?
- d. What is the link, if any, between a call for greater privacy and one's personal moral development? Or is a call for greater privacy solely about individual autonomy to do whatever one likes?
- e. What kind of internet would promote humanity's moral conscience?

In starting to respond to these questions, the report argues that the feared displacement of human decision-making through algorithms, machine learning and digital technologies should not be considered a threat to democracy because human decision-making remains unique in the way it is able to order pursuit of the common good through interpersonal consensus.

Democratic societies are able to aim at social unity as a good in itself. This grants their permanence, despite all manner of technological developments, which instead specialise at efficiency in pursuing already established objectives. Indeed, a desired goal of interpersonal unity lies in a sense behind all our social engagements and should be factored into evaluations of whether decision-making structures are efficient or optimal. Key to this is identifying the civic ideals that will release trusting and trustworthy democratic energies for the networked age. To this we now turn.

Chapter 1

Efficiency is a good aim
but not enough



Efficiency Is a Good Aim But Not Enough



a. The networked age

The networked age plays host to two overlapping phenomena: the erosion of interpersonal distance through communicative technologies, and the transformation of decision-making to become human-machine cooperative. The report asks what these two phenomena mean for our civic engagement and moral reasoning as a community. As interpersonal distance is eroded and decision-making becomes human-machine cooperative we lack a robust account of *good citizenship*¹ aimed at *human flourishing*²—this is the debate the report hopes to spark. The present chapter outlines the technological innovations that give the networked age its specificity. These developments, such as machine learning and big data, present novel difficulties for how we understand the space for human decision-making going forward. In particular, these developments can sometimes make it look as if human reasoning is no longer needed in progressing towards a more flourishing society. And if human reasoning is no longer needed, it becomes harder and harder to reintroduce human moral thinking into decision-making structures. In responding to this challenge, the chapter points out key differences

¹ Throughout this report, the term “citizenship” is used interchangeably with “civic engagement”. There is a difficulty in this, however, in that British use of “citizenship” tends to avoid such connotations and takes a more strictly legal understanding (a legal citizen, contrasted with a resident or migrant). American use of “citizenship” instead tends to appeal to community engagement and community duties as well (a good citizen as one who serves his or her community). For the avoidance of doubt, the use throughout is the American sense—we are interested in what the networked age means for our civic engagement and political participation.

² For insightful mapping of what “human flourishing” means, see VanderWeele, T. J., ‘On the promotion of human flourishing’. *PNAS*, Vol. 114, No. 31 (2017), pp. 8148-8156; VanderWeele, T. J., ‘Measures of Community Well-Being: a Template’. *International Journal of Community Well-Being*, Vol. 2 (2019), pp. 253-275.

between human and machine problem-solving and decision-making. It is then possible to demonstrate aspects of human moral reasoning that will not only endure but, in fact, increase in relevance in the midst of technological change. Humans are unique in their moral reasoning because they are able to seek universal, timeless goods while being self-aware of their future death. This habit produces in humans a unique capacity to form an overall sense of goodness and moral purpose despite individual goods being hard to optimise on a single spectrum. Humans' unique ability to engage in moral reasoning will only increase in importance as the networked age reaches its peak.

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

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

seems the advent of human-machine cooperation means specialists are themselves partially cut-out from giving full explanations of what progress is and where it will lead.

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

A common defence against the accusation that decision-making is being taken out of the hands of ordinary people is that scientific and technological progress simply helps realise the self-determination of individual people, meaning they remain ultimately in charge and will not have to obey arbitrary authority in the use of these technologies. This fails on ethical, existential and technological grounds. *Ethically* speaking, it requires consensus on a libertarian position of everyone's self-defined morality being tolerable, and such consensus is not apparent. Libertarianism enjoys little popularity and morally incompatible options cannot actually be tolerated on grounds that truth itself is self-determined. Then, *existentially* speaking, the defence that scientific and technological progress serves individual self-determination also does not hold. Humans are social by nature and realise their aims as part of groups (families, schools, social classes). This means that pursuing one's self-determination as an individual does not existentially represent what society is or the fact that we often seek group-determination, not individual self-determination. *Technologically* speaking the self-determination justification also fails because machine-human cooperation is leading to detachment from human agency in the execution of tasks and the optimisation of solutions.

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

Henry Kissinger, Eric Schmidt and Daniel Huttenlocher write that the Enlightenment replaced Christianity's emphasis on the divine with an emphasis on individual reason. They further that the age of the internet and artificial intelligence threatens to displace that emphasis on individual reason as artificial intelligence automates the drawing of conclusions from data and the action that follows from those conclusions.³ Over the ages there have been changes in the coordinating goods that society can be said to be seeking, though such paradigms are near impossible to fully demarcate—from seeking the divine (oneness with the creator), to sound individual reason (oneness with oneself), to efficient data-gathering and automated execution (oneness with doing). Alongside these shifts are concurrent shifts in our conceptualisations of the opposites to the goods that we seek—those evils we try to avoid. Vaguely, one can outline the following transitions through history on what counts as bad or evil, with our current networked age driven by the avoidance of inefficiency, a commitment begun in the industrial revolution:

³ Kissinger, H. A., Schmidt, E. & Huttenlocher, D., 'The Metamorphosis'. *The Atlantic* (Aug 2019).
<https://www.theatlantic.com/magazine/archive/2019/08/henry-kissinger-the-metamorphosis-ai/592771/> (emphasis in original).

Changing conceptualisations of evil



Figure 1.1 Key differences in conceptualisations of evil through the ages

Kissinger writes:

The Enlightenment sought to submit traditional verities to a liberated, analytical human reason. The internet's purpose is to ratify knowledge through the accumulation and manipulation of

ever expanding data. Human cognition loses its personal character. Individuals turn into data, and data become regnant.⁴

In the face of this chilling assessment, and the apparent struggle between efficiency and democracy, the report seeks to make a starkly positive contribution: we can and will find space for citizenship in the networked age and it is a space that completes, rather than competes with, our human flourishing. Given that one of the two core components of the networked age is the advent of human-machine cooperative decision-making, this chapter gives an initial overview of these dynamics to help show exactly what is at stake. By reviewing areas of technological progress one can more easily point out the essentially human contribution that can nevertheless endure in shaping the networked age.

b. Big data & big computation

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

⁴ Kissinger, H. A., 'How the Enlightenment Ends'. *The Atlantic* (Jun 2018).

<https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>.

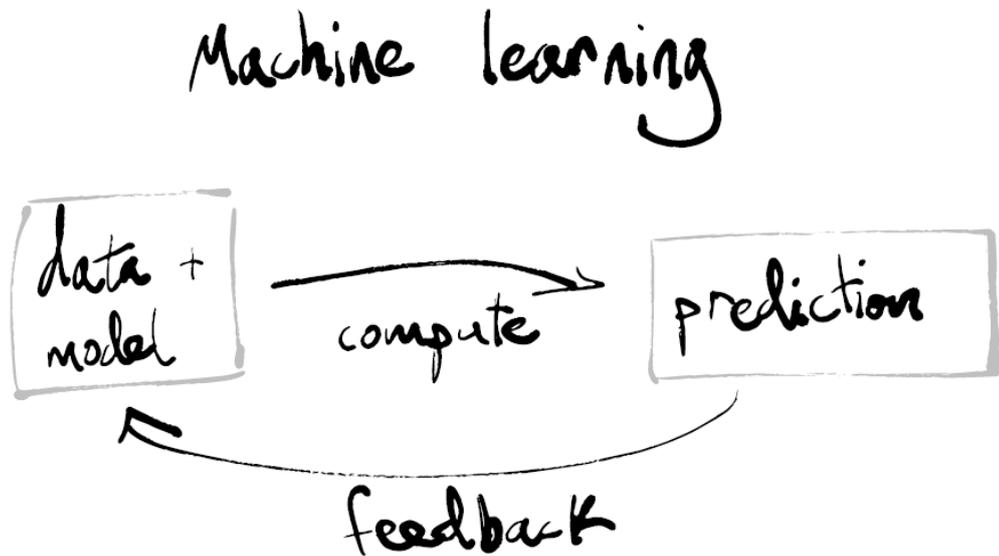


Figure 1.2 A simple diagram of machine learning

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

These advances in capacity are mainly in terms of big data and big computation. Alongside this, however, there is additional advancement in method because machine learning can refine the explanatory model over the course of the investigation. This automates the traditional scientific approach of writing down a hypothesis and then conducting an experiment, because in a sense the hypothesis is constantly re-written while computation evaluates the data and provides feedback on what model would best fit.⁵

⁵ Udrescu, S. & Tegmark, M., 'AI Feynman: a Physics-Inspired Method for Symbolic Regression' (27 May 2019). arXiv:1905.11481.

In a controversial article entitled ‘The End of Theory’, Chris Anderson pushes the point still further by arguing that big data means we no longer need scientific models.⁶ His view is that the amount of data now available qualitatively shifts what sorts of methodologies we should use to understand the world, and as such debunks the idea that we need to rely on the causal models scientists traditionally have in their minds when they carry out experiments. Instead, data itself can frame the research design and lead the way in identifying consistencies and inconsistencies and tell us directly about the way the world works. He explains:

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

But faced with massive data, this approach to science—hypothesize, model, test—is becoming obsolete. [...] Petabytes allow us to say: “Correlation is not enough.” We can stop looking for models. We can analyze the data without hypotheses about what it might show. We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot.⁷

Of course, in terms of logic, Anderson’s view that theory has therefore ended is, itself, a theory, which makes him part of the old guard that need to be transcended. But even when stepping aside this minor difficulty, we run into the problem that there are, clouded

⁶ Anderson, C., ‘The End of Theory: The Data Deluge Makes the Scientific Method Obsolete’. *Wired* (23 Jun 2008).
<https://www.wired.com/2008/06/pb-theory/>.

⁷ Ibid.

within his argument, two distinct ways in which big data can of its own accord be thought able to overcome the shortcomings of previous methods of analysis, and these two ways need to be unpicked.

The first is that big data enhances and accelerates capacities for identifying correlations.⁸ Via this route, through more data an analyst can be surprised at a correlation that was unexpected, like if he or she were to run a multilinear regression on the demographics that tend to watch YouTube videos of cats and find—surprisingly—that they are mostly watched by those in the financial sector. In a sense, the theory that the finance sector loves cat videos did not exist in the mind of the researcher prior to the observed correlation, but this does not amount to an elimination of the need for a model because, notwithstanding the surprise, such a model is latent in the choice of demographic factors included in the study, as well as in the more general assumption that correlations identified through this method indicate meaningful relationships. It may be that machine learning increases the number of parameters for consideration during the process of feedback and computation,⁹ but this does not mean no model is needed at all, just that the model is being constantly refined. Continuous machine choice of what parameters or factors to include in the model is still based on and linked back to the original criteria of the research design—seeking explanation or optimisation of a particular dependent variable. That enduring link with the *explicandum* (what is to be explained), means there will always be a model, however buried.¹⁰

⁸ On the way artificial intelligence is often limited to statistical regressions to identify correlations, see Askonas, J., ‘How Tech Utopia Fostered Tyranny’. *The New Atlantis* (Winter, 2019). <https://www.thenewatlantis.com/publications/how-tech-utopia-fostered-tyranny>.

⁹ Ghahramani, Z., ‘Probabilistic machine learning and artificial intelligence’. *Nature*, Vol. 521 (2015), pp. 452-459, pp. 454-5.

¹⁰ This point is applicable to the formation of research designs throughout the sciences, and is made here to demonstrate that machine learning likewise does not escape the need for human-instigated modelling. For further debate on this in biology over the

The second distinct way in which big data might be considered a resource so comprehensive that it amounts to a new scientific method is in the way it allows a personalisation or individuation of prediction hitherto unprecedented.¹¹ In other words, big data means small modelling—so small, in fact, that to some it does not look like modelling at all. It is a bit like saying that because I know so much about you as an individual, I can predict when you will go for your morning coffee without the aid of any model suggesting what times people like you tend to go for their morning coffee. What is going on here? The idea is that, through big data, I know you inside out and no longer need to collate studies of other people like you to get at your daily routine.

Again, however, there is a simplification at play that does not do justice to the underlying scientific methods. The mistake in logic is the idea that because my prediction works, I no longer need reference a model, even though there was a model involved in the first place and there continues to be one as I go about making correct predictions. Dropping explicit reference to hypotheses when certain of a

application of systems theory, see Leyser, O. & Wiseman, H., 'Integrative Biology: Parts, Wholes, Levels and Systems'. Ch 2 of Reiss, M. J., Watts, F. & Wiseman, H. (eds.), *Rethinking Biology: Public Understandings* (New Jersey: World Scientific, 2019); Gatherer, D., 'Modelling versus Realisation: Rival Philosophies of Computational Theory in Systems Biology'. Ch 3 of Reiss, M. J., Watts, F. & Wiseman, H. (eds.), *Rethinking Biology: Public Understandings* (New Jersey: World Scientific, 2019).

¹¹ Explaining the method, Zoubin Ghahramani writes, 'many large data sets are in fact large collections of small data sets. For example, in areas such as personalized medicine and recommendation systems, there might be a large amount of data, but there is still a relatively small amount of data for each patient or client, respectively. To customize predictions for each person it becomes necessary to build a model for each person—with its inherent uncertainties—and to couple these models together in a hierarchy so that information can be borrowed from other similar people. We call this the personalization of models, and it is naturally implemented using hierarchical Bayesian approaches such as hierarchical Dirichlet processes, and Bayesian multitask learning.' Ghahramani, 2015, p. 458.

property's causal dynamics is nothing strange in science. If we know how the cell works, for example, we know how the cell works; by this we mean we know cells in reality, not just in our imagined models. We can drop reference to what is hypothesised about the cell in front of us and speak with confidence about what we know. Likewise, if it is the case that I did actually manage to know enough about you to be certain of when you go for your morning coffee (unlikely though that is, given the way big data is not yet that big¹²), it would be because my model of you corresponds to reality and is true, not because a hefty amount of data means I can dispense with modelling. Though I may cease to speak in hypothetical terms, I am still committed to a behavioural model. It is just that in this case the model happens to be accurate. When it is not sufficiently accurate I revert to talking about what I thought would be the case, which reveals once again the predictive model I had in mind all along.

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

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

¹² Graham, M., 'Big data and the end of theory?' *The Guardian* (9 Mar 2012).
<https://www.theguardian.com/news/datablog/2012/mar/09/big-data-theory>.

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

There's no reason to cling to our old ways. It's time to ask: What can science learn from Google?¹³

Despite the above reservations about this conclusion, let us nevertheless suppose it is accurate and ask: what would a world where correlation supersedes causation look like? The hypothetical is important because the view may—even if untrue—be sociologically relevant for the way we go about shaping the world we are in, if believed. A world where correlation supersedes causation would be one where all impetus is placed on collecting and codifying data rather than thinking through the meaning behind how data points relate to one another. Data would accumulate and the speed of computation advance, but the quest to make sense of the process—either in the form of predictions or models—would recede in importance. Jonathan Zittrain makes the added point:

Intellectual debt accrued through machine learning features risks beyond the ones created through old-style trial and error. Because most machine-learning models cannot offer reasons for their ongoing judgments, there is no way to tell when they've misfired if one doesn't already have an independent judgment about the answers they provide.¹⁴

¹³ Anderson, 2008.

¹⁴ Zittrain, J., 'The Hidden Costs of Automated Thinking'. *The New Yorker* (23 Jul 2019). <https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking>.

Interest and respect would switch to those studies able to amass large datasets over those that develop independent lines of causal reflection.

This point has increasing relevance for society as a whole because the extent to which the search for truth needs human direction dictates our need for civic participation over the long-term. In rejecting Anderson's proposed approach of letting data correlations obviate theories of causation, one instead commits to the position that no matter the advances in big data and big computation, there will always be a space for goal-setting and the identification of dependent variables and explicanda—a space that is uniquely human.

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

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

c. Automation, algorithms & artificial intelligence

Algorithmic decision-making can be defined as optimised responses following pattern identification. Algorithmic decision-making can be automated computationally and therefore executed at great speeds, but it essentially rests on these two core elements of pattern identification and response. All algorithmic decision-making involves some initial codification of data, and then an execution command based on the results of that data codification. It is an ongoing question whether this amounts to *intelligence*. Usually, commentators will begin by describing surprisingly effective examples of algorithmic decision-

making or machine learning (such as AlphaZero playing chess), but that risks putting the cart before the horse. Definitively answering whether such processes count as intelligence requires first starting with a definition of intelligence before specific examples are considered. Otherwise what is most interesting or humanlike in the examples one sees biases one's sense of what intelligence looks like. If intelligence is defined in terms of being able to give optimised responses following pattern identification, then yes there is a thing called artificial intelligence that is becoming increasingly dominant. If, instead, there is something else to intelligence, one has to see whether machines are capable of it. If not, and humans are capable, one has to conclude that there is no such thing as artificial intelligence properly understood.

It is a regular concern that in the enthusiasm for technological progress, commentators anthropomorphise artificial intelligence (talking of computers as if they are humanlike), believing computers to be smart because they seem to mimic human thought. Less detected is the extent to which this also *anthropocentrises intelligence* (talking of intelligence as if it is most perfectly featured by humanity). Intelligence is not an exclusively human faculty, and there is no reason why many other things cannot be understood as intelligent.¹⁵ There is a danger in confusing the question, "Is this intelligent?", with the closely related but distinct question, "Is this intelligent in the way humans are intelligent?" In this sense, the Turing Test does more to blur than enlighten. The Turing Test is an 'Imitation Game'¹⁶ of asking 'whether or not a computer is capable of thinking like a human

¹⁵ See Templeton World Charity Foundation, 'Diverse Intelligences' (2018). <https://www.templetonworldcharity.org/our-work/diverse-intelligences>.

¹⁶ Alan Turing's original description of the Turing Test. Christian, B., *The Most Human Human: What Artificial Intelligence Teaches Us About Being Alive* (London: Penguin Books, 2011), p. 10.

being'.¹⁷ This is a confusing anthropocentrising of intelligence because it is not clear whether one is testing for intelligence or testing for being humanlike. In principle, an especially intelligent human may not look and act like the average human and so might fail the Turing Test by failing to properly imitate what is thought averagely human. In this sense, the Turing Test sometimes works more like a test of social congruence.

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

Let us consider a definition of intelligence: 'the ability to learn, understand, and make judgments or have opinions that are based on reason'.¹⁸ It seems that under the general descriptions already outlined, machines are capable of this if, for the sake of argument, we further specify 'understand, and make judgments ... based on reason' in terms of *following what is logical*. The definition above of algorithmic decision-making as *optimised responses following pattern identification* can, in this sense, count as making judgments based on reason in that the outputs logically follow from the inputs.

Giving the benefit of the doubt in this way, let us turn to the more specific question of whether machines are intelligent like humans are,

¹⁷ Rouse, M., 'Definition: Turing Test'. *TechTarget* (2019).

<https://searchenterpriseai.techtarget.com/definition/Turing-test>.

¹⁸ Cambridge Dictionary (Cambridge University Press, 2019).

<https://dictionary.cambridge.org/dictionary/english/intelligence>.

bearing in mind that humans are not necessarily the most intelligent beings around and so the fact that something is very intelligent does not make it humanlike, necessarily.

Michael Jordan, of University of California, Berkeley, explains that those narrating the rise of artificial intelligence often compound two different things, and this has direct relevance for whether machines are becoming intelligent like humans are. One area of progress is the development of software and hardware that seeks to approach or copy human-level intelligence, what he calls ‘human-imitative AI’. That was the original use of the captivating acronym A.I. as coined in the 1950s. The second is what he terms ‘intelligence augmentation’, something that has enjoyed great progress over the past two decades and is where ‘computation and data are used to create services that augment human intelligence and creativity.’¹⁹ Intelligence Augmentation flips the acronym to IA, and is in some ways an opposite to AI. It puts human intelligence first, with computation used merely to accelerate the execution of human ideas and goals. Jordan points out that IA is very unhelpfully subsumed under the general heading of AI because an increased ability to execute human ideas often looks like growth in non-human intelligence, but it is not. Jordan explains the confusion:

While related academic fields such as operations research, statistics, pattern recognition, information theory and control theory already existed, and were often inspired by human intelligence (and animal intelligence), these fields were arguably focused on “low-level” signals and decisions. [...] “AI” was meant to focus on something different—the “high-level” or “cognitive” capability of humans to “reason” and to “think.” Sixty

¹⁹ Jordan, M., ‘Artificial Intelligence – The Revolution Hasn’t Happened Yet’. *Medium* (19 Apr 2018). <https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7>.

years later, however, high-level reasoning and thought remain elusive. The developments which are now being called “AI” arose mostly in the engineering fields associated with low-level pattern recognition and movement control, and in the field of statistics—the discipline focused on finding patterns in data and on making well-founded predictions, tests of hypotheses and decisions.²⁰

In this way, optimisation or statistics researchers ‘wake up to find themselves suddenly referred to as “AI researchers”²¹ even though their work is more IA than AI. Statistical methods have very little to do with the human-imitative AI that is doing so much to heighten expectation in future capabilities for Artificial General Intelligence (AGI) and machine-led decision-making,²² and yet there endures a general assumption that IA and AI are the same thing. The distinction between IA and AI holds, however, even though machine learning systems are able to make ‘predictions which sometimes far exceed the capabilities of the top humans and competing computer programs in the world.’²³ It is not about processing speeds so much as the original reason why the calculations are being done in the first place that dictates whether the process is properly described AI or IA. If the computation is being done to automate goals that human creators originally set, it is IA. If the computation is to artificially recreate human behaviour or thought, it is AI. Success in human-imitative AI is neither sufficient nor necessary to accelerate progress in IA, and vice versa. After all, we already are humans, so why would creation of human-imitative AI augment our intelligence?

²⁰ Ibid.

²¹ Ibid.

²² On the creation of human-imitative AI that heightens expectations of AGI, see the work of anthropologist Beth Singler.

²³ Hawley, S. H., ‘Theopolis Monk: Envisioning a Future of A.I. Public Service’. Ch 14 of Lee, N. (ed.), *The Transhumanism Handbook* (Cham: Springer, 2019), part 2.

What about those who say machine learning is so advanced, or will be so advanced, that it produces outcomes that were not envisaged by the original human designers?²⁴ As should be clear from the preceding discussion, that framing already makes the mistake of *anthropocentrising intelligence* by taking human intelligence as the definition of what intelligence is, such that going past that limit in any respect gives us the impression that the machine has reached a higher level of being, potentially more competent in all respects. As human capacity is not a sovereign indicator of intelligence, we would do well to remember that going beyond what humans can do in one respect is limited to describing advancement in that respect only. So the fact that human designers did not envisage a particular action of the machine is insufficient for concluding that the machine is superintelligent. Indeed, accidents are often unexpected and sometimes they even give the impression that the thing that caused the accident has a life of its own, but that actually says precious little about whether the thing is intelligent.

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

1. Get as much data as you can and make sure it is of highest quality.
2. Distill your data into signals that will be maximally predictive—a process called feature engineering.
3. Once you have the most awesome data and tools for feature engineering, keep raising the capacity of your algorithms.²⁵

²⁴ Bostrom, N., *Superintelligence: paths, dangers, strategies* (Oxford: Oxford University Press, 2014); Tegmark, M., *Life 3.0: being human in the age of artificial intelligence* (London: Allen Lane, 2017).

²⁵ Bucher, T., *If... Then: Algorithmic Power and Politics* (Oxford: Oxford University Press, 2018), p. 25.

The key jump in logic in this seemingly straightforward explanation is in point two, where the researcher seeks to be ‘maximally predictive’. It begs the question on what it is that the machine is, ultimately, trying to predict. Although that step is hidden here, it is essential for our question of whether technological advancements remove the need for human input, and therefore civic input. As Taina Bucher explains:

Feature engineering, or the process of extracting and selecting the most important features from the data, is arguably one of the most important aspects of machine learning. While feature extraction is usually performed manually, recent advances in deep learning now embed automatic feature engineering into the modeling process itself. If the algorithm operates on badly drawn features, the results will be poor, no matter how excellent the algorithm is.²⁶

By omitting the human agency involved in selecting the thing that needs to be explained by step two, it is possible to give the impression that machines have the capacity to be self-learning and even self-determining for the whole process. However, the fact that human agency and prioritisation will always be present in step two—as a kind of limited first mover—means there will not be self-governing artificial intelligence in the sense of ultimate goals that are self-defined by machines. While it is accurate to describe machines as often displaying a form of intelligence by making judgments based on reason (in terms of following what is logical according to programming), there is a very different field of intelligence that humans have which will endure in setting the goals for the machine learning process. This enduring feature must be outlined and evaluated on its own terms if we are to fathom the true scope of the networked age’s human-machine cooperation.

²⁶ Ibid.

d. Human intention & civic ideals

Julie Cohen explains that ‘[i]nformation is never just information: even pattern identification is informed by values about what makes a pattern and why, and why the pattern in question is worth noting.’²⁷ The permanence of this feature in the human-machine learning process points to the long-term place of human values in our networked age. Rather than a force for regression, the inclusion of human thinking in the value part of our methodologies grants ‘a real opportunity to conceive of something historically new—a human-centric engineering discipline.’²⁸

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

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

²⁷ Cohen, J. E., ‘What Privacy Is For’. *Harvard Law Review*, Vol. 126 (2013), pp. 1904-1933, pp. 1924-5.

²⁸ Jordan, 2018.

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

The good life is not, therefore, about maximising a single variable and then terminating but about pursuing enduring goals in the midst of a difficult unknown. Leading a good life thus involves something of a mastering of one's relationship with death—being aware of death and its significance but not letting it paralyse pursuit of one's mission, duties and goals. From a machine's perspective, humans live very strangely indeed in seeking goals that are neither part of a loop nor steps towards their termination. From a human's perspective, machines live strangely in never enjoying the moment and always moving on to the next thing. Human life is about finding and pursuing the meaning of life despite knowing one's forthcoming death can come at any time and for seemingly unrelated reasons. Our reasoning process is therefore hard-wired in being able to navigate the incommensurable²⁹ push-and-pull of seeking atemporal goods despite temporal limitations.

²⁹ For discussion of the difference between the terms 'incommensurable' and 'intransitive', and why it is usually better to use the latter, see Burbidge, D., 'The Inherently Political Nature of Subsidiarity'. *American Journal of Jurisprudence*, Vol. 62, No. 2 (2017), p. 143-164, pp. 157-8. The former is used throughout this text simply because it is the more readily used phrase for making the basic point.

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

As stated, the end point and the goal are different in humans but in machines they are the same. This stark and unyielding difference also reveals the problem with applying a utility framework to trying to live well as a human, despite the fact that a utility framework is often found to be useful and appropriate when programming algorithmic decision-making. The big *advantage and disadvantage* of utilitarianism is that it is an ethical framework with no sense of time. This is an advantage because utility can be analysed and calculated in abstract, but it is also a disadvantage because utilitarianism must then rely on other accounts for interpretations of time and natural change.³⁰ All consequentialist justifications are troubled by the question, *when?* That is, *what is the time at which it would be proper to evaluate what is the greatest good for the greatest number?* Present suffering for future gain can always be a good decision if no limit is given on the future. Time, as we know, is potentially infinite, which means a utilitarian calculation of the greatest good for the greatest number has no in-built *number*. Actual application of utility frameworks requires an arbitrary demarcation of the timeline we must keep to—something not provided by the utility framework itself but through outside narration of how things naturally

³⁰ The occasional use in utilitarianism and economics of a “discount rate” that gives a rate at which future utility can be ignored is the exception that proves the rule. The discount rate attempts to redress utilitarianism's lack of a sense of time or natural change. For its estimation, it draws on external understandings of time, change, depreciation or likely future events that are distinct from the method of appraising the good in terms of utility.

change. In the same way as a utility framework is always placed on an already structured problem—with humans describing the temporal limits to the problem—so too are all plans of how to live a good life developed in conjunction with an assessment of the likely time of one's end. Humans engage in a praxis of working out their achievement of mission, duties and goals, in the face of the incommensurable fact of death.

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

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

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

Chapter 2

Where citizenship starts



Where Citizenship Starts



a. Our human space, going forward

We need to find ideals of citizenship that work for value-based decision-making in our networked age. As the discussion in Chapter 1 made clear, the moral dimension to human decision-making will continue to be directional for the shaping of digital technologies, despite enthusiasm among some tech utopians to the contrary. Agreeing on and establishing ideals for civic engagement in our networked age will therefore be fundamental for healing democratic society. As Corrine Cath and colleagues write:

We are creating the digital world in which future generations will spend most of their time. [T]he design of a “good AI society” should be based on the holistic respect (i.e., a respect that considers the whole context of human flourishing) and nurturing of human dignity as the grounding foundation of a better world. The best future of a “good AI society” is one in which it helps the infosphere and the biosphere to prosper together.³¹

But before providing an account of any such good society we have to provide an account of what society is—the way we are bound together, our collective human space. Citizenship, as *a normative fact about who we are able to be when we come together politically*, fuses the basic matter of society with an account of the specifically democratic. It is a

³¹ Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M. & Floridi, L., ‘Artificial Intelligence and the “Good Society”: the US, EU, and UK approach’. *Science and Engineering Ethics*, Vol. 24, No. 2 (2018), pp. 505-528, p. 508.

mixed normative and empirical endeavour, saying something both about who we are (empirically) and the kind of people we want to be (normatively). Citizenship is hard to build up and hard to break down; something of a habit of the heart, resiliently playing out generation after generation.³²

These methods of civic togetherness are under attack. Who we are was broken up by an economics of individualism long before digital technologies came into existence,³³ and technologies have capitalised on the atomisation of cultures and societies through facilitating rapid personalisation of use and purpose. These technologies are not, therefore, optimising formation of common goals but, instead, optimising an individuation of goals. To some extent that was always the way with tools—they facilitate humanity’s greater specialisation of practices over time. Nevertheless, there is a difference here in the extent of personalisation and in the difficulty of assuming togetherness in social living, forcing us out of traditional questions on the economic effects of digital technologies and towards the relatively newer topic of their social and political effects. With the personalisation of human experience through digital technologies, do we still have the *common ground* that can work as the basis to a *common good*?

³² Julie Cohen remarks, ‘Well-functioning state and market institutions cannot be built in the span of a grant-funded research project or a military campaign. Their rhythms and norms must be learned and then internalized, bringing into being the habits of mind and behavior that democratic citizenship requires.’ Cohen, J. E., ‘What Privacy Is For’. *Harvard Law Review*, Vol. 126 (2013), pp. 1904-1933, p. 1912. See also Putnam, R. D., *Making Democracy Work* (Princeton: Princeton University Press, 1993); Putnam, R. D., *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon & Schuster, 2000); Uslaner, E. M., *The Moral Foundations of Trust* (Cambridge: Cambridge University Press, 2002).

³³ Manent, P., *A World beyond Politics? A Defense of the Nation-State* (Princeton: Princeton University Press, 2006); Elshtain, J. B., *Sovereignty: God, State, and Self* (New York: Basic Books, 2008); Siedentop, L., *Inventing the Individual: The Origins of Western Liberalism* (London: Penguin Books, 2015); Deneen, P. J., *Why Liberalism Failed* (New Haven: Yale University Press, 2018).

Through advances in augmented reality, for example, we will literally be seeing different worlds, depending on how good one's phone or glasses are; the world will appear differently to each person in accordance with how much they can afford these technologies and what options they choose within them. And seeing is believing. The dynamic is already at play more broadly in inequalities of data gathering, whereby the people around us online appear differently depending on our level of access to their data profiles. Trusted friends see me and my full timeline, loose acquaintances see my most recent post. Facebook, Inc. sees me nakedly. People become passive data generators, and organisations the active users of that data for predicting and generating future trends.³⁴ Legal scholar Julie Cohen coins the term “modulated society” to describe a world in which individuals are fed choices that suit their comfort level, rather like setting a thermostat to a preferred temperature. She details the move from a liberal to a modulated society through increased surveillance and data-gathering techniques:

Citizens of the modulated society are not the same citizens that the liberal democratic political tradition assumes, nor do their modulated preferences even approximately resemble the independent decisions, formed through robust and open debate, that liberal democracy requires to sustain and perfect itself. The modulated society is the consummate social and intellectual rheostat, continually adjusting the information environment to each individual's comfort level. Liberal democratic citizenship requires a certain amount of *discomfort*—enough to motivate citizens to pursue improvements in the realization of political and social ideals. The modulated citizenry lacks the wherewithal and perhaps even the desire to practice this sort of citizenship.³⁵

³⁴ Zuboff, S., *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power* (London: Profile Books, 2019).

³⁵ Cohen, 2013, p. 1918 (emphasis in original).

It may be that one day humanity will conform to Cohen's image of citizens modulated beyond all ability to engage in self-direction, but we are not there yet. If it were true, there would be no value in writing and reading this report; it is premised on the view that good ideas can help make a good citizenry.³⁶ The liberal democratic tradition has regularly worked with a view of citizenship slightly more active and competent than we tend to find among real citizens; in that subtle idealism we give effect to our aspiration for a freer world.³⁷ Admittedly, this is citizenship as *a normative fact about who we are able to be when we come together politically*, breaking the boundaries of the fact/value distinction set by David Hume and Immanuel Kant in the 18th century.

What we value as citizens need not necessarily be in conflict with what we are empirically. It is not hard to see that as humans we have become utterly dependent on mutual social commitment: just imagine trying to make your own ball point pen, let alone your own smart phone. Much of this cultural and social dependence relies on education to pass what is learned from one generation to the next. There is growing evidence that this kind of cultural evolution has occurred hand in hand with genetic evolution, to provide humans with the hardware in the form of brains and larynxes that facilitate cooperation.³⁸ Insofar as we

³⁶ Cohen describes the fear that '[s]timuli tailored to consumptive preferences crowd out other ways in which preferences and self-knowledge might be expressed, and also crowd out other kinds of motivators—altruism, empathy, and so on—that might spur innovation in different directions.' Cohen, 2013, p. 1926. While that is a relevant concern, if it really takes place there would be little point in writing about it. Academic writing in the humanities and social sciences is, ultimately, a project in social improvement through self-knowledge.

³⁷ A good example is Garton Ash, T., *Free World: Why a Crisis of the West Reveals the Opportunity of Our Time* (London: Penguin Books, 2005).

³⁸ Henrich, J., *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (Princeton: Princeton University Press, 2017); Christakis, N. A., *Blueprint: The*

believe that cooperation and education are good things for citizens to promote, there will be a connection between what is and what should be. But it is not automatic.

Just as our notion of citizenship does not fully rely on the facts of who we are, but also on the normativity of who we want to be, so the changes brought through digital technologies tell only part of the story on the kind of humanity we will likely become. Our civic ideals are thus *empirically influential* for many aspects of the society we will go on to inhabit, because while their realisation is frequently frustrated, they endure as an important causal factor for what gives us common direction.³⁹

We turn now to two main accounts of how civic ideals can be generated. These two accounts are by no means the only ones available for rebuilding our civic ideals for the networked age, but they do provide a helpful guide on the richness of prior ethical and political thinking. The first is the citizen-slave distinction, and the second the ethics of navigating problems of scarcity and competing desires. These are two alternative routes in the history of social thought for establishing norms and rules for what it means to be a person-in-community, a citizen. Both hold relevance for rebuilding our civic ideals for the networked age, and therefore the account that follows draws from both literatures. The concluding argument of this Chapter is that both accounts indicate how the empirically-grounded problems of the networked age can, in fact, provide the material for an opposite rebuilding of our civic ideals.

Evolutionary Origins of a Good Society (New York: Little, Brown Spark, 2019).

³⁹ See, for example, the role played by particular countries' tolerance for 'creative destruction' in their economic and political transformation over time, as accounted for in Acemoglu, D. & Robinson, J. A., *Why Nations Fail: The Origins of Power, Prosperity and Poverty* (London: Profile Books, 2012).

b. The citizen vs slave distinction

Citizenship has both legal and normative dimensions; the focus here is on the normative: what is a good citizen, and what should a good citizen do? Defined as *a normative fact about who we are able to be when we come together politically*, citizenship is being looked at here from the perspective of the person-in-community—social by nature, finding particular fulfilment in joint human efforts and partnerships.⁴⁰ In terms of a normative aspiration, there is an ideal citizen contribution to the life of the body politic that can be imagined and then realised, at least in degrees. For liberal democratic orders, ideal citizenship has tended to include a spirit of public service, justice, neighbourliness, democratic participation and moral reasoning. We often fall short in achieving these elements in their fullest, yet they continue to guide us, and even inform us on who is failing to contribute. In their achievement, we form an active and well-connected community in pursuit of a common good.

The normative dimension to our common understanding of citizenship was in part generated through 18th century reflection on the differences between citizens and slaves. The basic argument from Jean-Jacques Rousseau is that in order to participate fully in the life of the body politic, we must throw off the chains of our slave-like dependency on rulers and their institutions.⁴¹ For Rousseau, this did not only apply as a criticism of the aristocratic and royal elites but to the church too—whose hierarchy, symbolism and narrative of the afterlife kept people subjugated, he felt. A coming of age for the citizen was about breaking free from these chains: ‘What makes the work of legislation difficult is not so much what has to be established

⁴⁰ Aristotle, *The Politics* (London: Penguin, 1992); Aristotle, *The Nicomachean Ethics* (London: Penguin, 2004).

⁴¹ Rousseau, J. J., *The Social Contract* (Cambridge: Cambridge University Press, 1997 [1762]).

as what has to be destroyed'.⁴² Rousseau held there to be something naturally good in establishing harmony between state and society, such that governance was about citizenship by means of releasing the general will of citizens.⁴³ The novelty of his thought came in establishing a connection between the freedom of citizens and the goodness of citizens. A flourishing society is less about producing good people as having a state and political order that realises peoples' natural goodness through harmony with their self-direction.

The idea of fair inclusion of citizens in the life of the polis is, of course, much older. The Roman sense of citizenship made general appeal to fairness, for example, but it was in terms of a fairness in receiving contributions from the wealthy, for which non-citizens should not be entitled. The way in which early Christians dedicated themselves to almsgiving towards the non-citizen poor put them at odds with Roman ideas of citizenship, contributing to the sense that they must be to blame for the fall of Rome.⁴⁴ A "citizen" is a term wedded to that of "city" and yet the Christian idea of almsgiving worked against the idea that the rich should give back to their home cities, arguing that God wanted giving to those least able to give back, and regardless of where they were. The provision of infrastructure and public entertainment from the wealthy to the citizens of one's native home gradually gave

⁴² Ibid, p. 78.

⁴³ 'What makes the constitution of a State genuinely solid and lasting is when what is appropriate is so well attended to that natural relations and the laws always agree on the same points, and the latter as it were only secure, accompany and rectify the former. But if the Lawgiver mistakes his object, if he adopts a principle different from that which arises from the nature of things, if one principle tends toward servitude while the other tends toward freedom, one toward wealth, the other toward population, one toward peace, the other toward conquests, then the laws will be found imperceptibly to weaken, the constitution to deteriorate, and the State will not be free of turmoil until it is either destroyed or altered, and invincible nature has resumed its empire.' Ibid, pp. 79-80.

⁴⁴ Gianakon, S. E., 'Citizenship and the Holy in Late Antiquity'. Humboldt Universität, Fulbright essay (May 2018).

way, therefore, to monasteries offering hospitality, health care and spiritual direction, regardless of citizenship. While a causal connection between Christianity and the decline of the Roman Empire is difficult to maintain,⁴⁵ these changes did result in rupture between normative and legal understandings of Roman citizenship, with the legal retaining much of its sense of exclusivity and the normative instead leaning closer to an idea of universal human dignity.

Roman citizenship was therefore a privilege distinct from slavery, but Rousseau's citizenship is instead non-privileged and natural, getting closer to that Christian idea. Both Rome and Rousseau's ideas of citizenship look similar on one level in that they can both be contrasted with slavery, but in fact Rousseau makes a clear break with Roman tradition by removing the category of members of the public who are neither slaves nor citizens. For Rousseau, everyone is either a slave or a citizen: the former by malicious design, the latter by natural right. Indeed, any sub-groupings of citizens by means of third-party associations are, for Rousseau, damaging to the strength of the body politic.⁴⁶ To help explain his point in *The Social Contract*, Rousseau writes the following footnote to the word 'city':

The true sense of this word is almost entirely effaced among the moderns; most take a city for a City, and a bourgeois for a Citizen. They do not know that houses make the city but Citizens make the City. [...] Only the French assume the name *Citizen* casually, because they have no genuine idea of it, as can be seen in their Dictionaries; otherwise they would be committing the crime of Lese-Majesty in usurping it: for them this name expresses a virtue and not a right.⁴⁷

⁴⁵ Augustine, *City of God* (London: Oxford University Press, 1963 [426]).

⁴⁶ Rousseau, 1997, p. 60.

⁴⁷ *Ibid*, p. 51 (emphasis in original).

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

Citizenship as a normative ideal that compels itself from nature eventually translated into the civil rights movements of the 20th century.⁴⁸ Citizenship became inextricable from social solidarity and the refusal to be subject to arbitrary authority.⁴⁹ In its democratic form, it meant dynamic and participatory engagement with the decision-making that directs the people as a whole. All this was predicated on the naturalism of Rousseau: citizenship in its normative sense is inclusive of universalistic conceptions of human dignity, with citizens demanding equal say in the way society is directed in accordance with what is their natural right.

Those who defended slavery sometimes tried to present it as natural.⁵⁰ Citizenship is, of course, easy to imagine as legally constructed and therefore not so natural after all, by means of contrast. The contention—dismissed now—was that slavery is natural by virtue of some people always being able to dominate others, and citizenship unnatural in that it is an attempted legal affirmation of equality which is not there in nature. The naturalism for citizenship as *a normative fact about who we are able to be when we come together politically* won the day, and the counter-argument that slavery is evident in history and therefore natural is now largely confined to the classroom as a devil's advocate position.

⁴⁸ See, for example, Allen, D. S., *Talking to Strangers: Anxieties of Citizenship since Brown v. Board of Education* (Chicago: The University of Chicago Press, 2004); Nash, K., *Contemporary Political Sociology: Globalization, Politics, and Power* (Chichester: Wiley-Blackwell, 2010), 2nd Ed., pp. 131-2.

⁴⁹ Pettit, P., *Republicanism: A Theory of Freedom and Government* (Oxford: Oxford University Press, 1997).

⁵⁰ BBC, 'Ethics guide: Philosophers justifying slavery' (2014).
http://www.bbc.co.uk/ethics/slavery/ethics/philosophers_1.shtml.

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

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

There are, of course, many aspects to the growth of technology that meet with liberal democratic appreciation of rights and freedoms. On this reading, the liberal project is thoroughly naturalistic because it takes an a-historical account of human nature (what is timeless about our humanity) and asks that laws be harmonious with it. Such an approach is also a direct source of optimism for the networked age: the freedom granted by the internet was first thought something of a state-of-nature experiment, for which our regulatory institutions would be wrong to impose out-of-touch rules because freedom brings with it peace and harmony. While that optimism for the internet lasted for a time, two serious threats to the narrative arose. The first was that some online activities clash with human rights, which put the

naturalistic argument in a difficult position of needing to choose between protecting what it held to be, ultimately, “natural” rights or, instead, protecting the boundaries of a state-of-nature territory. A good example is the choice between prosecuting those who view paedophilia online or, instead, guaranteeing their anonymity for reasons of privacy, the former of which appeals to the natural rights of children while the latter appeals to the need to defend rules that will ensure freedom online. The second threat to the “natural freedom” narrative of the internet was that the controllers of the internet formed a new elite, setting its rules and providing corporate conditioning to what goes on. Interestingly, this new elitism excites a naturalistic alternative in the form of the “hacker”, who can “move fast and breaks things”—an adage Mark Zuckerberg had to abandon once Facebook’s size meant it had to take greater responsibility for what it was doing.⁵¹ The “hacker” can be both an ordinary programmer (as in “hackathons”) or an anti-system rebel/activist (as in “Anonymous”). Either way, the freedom to move fast and break things is felt to naturally lead to innovation and good outputs, even if socially and politically disruptive—a logical reasoning that can be traced back to Rousseau’s connection between the freedom and goodness of citizens.

While there are positives at play in Rousseau’s account in terms of equality (or at least affirmed equivalence) between citizens, it is important to keep in mind that his is ultimately a rejection of a perceived slave-like status quo, followed by appeals to less-specific naturalism. He provides little in terms of how one would purposefully restructure the relationship between citizens and authorities in the event the natural is not all that good, which is the situation we find ourselves in with the world wide web.

⁵¹ Taneja, H., ‘The Era of “Move Fast and Break Things” Is Over’. *Harvard Business Review* (22 Jan 2019). <https://hbr.org/2019/01/the-era-of-move-fast-and-break-things-is-over>.

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

But in trying to imagine the kind of society we want to be, simply asserting a citizen vs slave distinction does not do enough intellectual work. It has been useful in helping see what is at stake and how our current sense of citizenship as a natural, normative fact has accrued, but the distinction does not provide clear navigation on whether or not an even higher type of citizenship exists out there, and whether the technological improvements we are currently witnessing amount to completion of our social nature through *participatory pursuit of the computerised optimal* or whether they in fact reduce us to slave-like conditions. Lingering concerns compel us to turn to alternative ways of looking at the normative basis to our citizenship.

c. Ethics through the navigation of competing desires

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

1. Transparency: making interpretable the reasons for artificial intelligence's predictions or decisions.
2. Reliability: safely scaling probabilistic reasoning to unforeseen settings.

3. Trustworthiness: ensuring artificial intelligence can reliably initiate acts of kindness through the proper inferring of the beliefs and goals of other agents.⁵²

These are good ways to regulate artificial intelligence and avoid it becoming an unknowable, damaging force, but they do not show us good ways of being citizens. It may be that good citizenship is simply no longer necessary, but as Chapter 1 pointed out, all data-driven analysis and execution relies at some level on human-generated purpose. If, in a democratic society, we view human-generated purpose as best realised collectively and in mutual trust, the way we organise that collectivity and mutual trust is of permanent relevance for the shaping of our future society. Our mutual trust—our oneness—is dependent on our citizenship, that normative fact about who we are able to be when we come together politically. Point (3) of Weller’s prescription might be in danger of clouding the solution, for it asks for trustworthy artificial intelligence, assuming that it can be obtained through alignment with human goals. But not all human goals are trustworthy, and so artificial intelligence trained to align with them may prove destructive. The Leverhulme Centre for the Future of Intelligence gives further description of what is to be understood by ‘trustworthiness’:

Human studies indicate that a theory of mind may be essential to build empathetic trust, and for reliable initiation of acts of kindness. Equipping AIs to infer beliefs and goals of other agents (such as humans) may improve human-machine collaborations;

⁵² Weller, A., ‘Where is AI going, and how will it promote flourishing?’ Presentation at the conference ‘Science, Philosophy, Religion & Human Flourishing’, Ayia Napa, Cyprus, (Nov 2018). See also Leverhulme Centre for the Future of Intelligence, ‘Trust and Transparency’ (2019a). <http://www.lcfi.ac.uk/projects/ai-trust-and-society/trust-and-transparency/>.

yet such cognitive insight may prove a double-edged sword, allowing deception and even manipulation.⁵³

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

The level of our human-to-human trustworthiness dictates our overall trustworthiness in using tools to help us achieve our goals. In other words, the trustworthiness of our tools or machines are only of value insofar as we are trustworthy users and can trust each other so to be. Nuclear power, for example, can be used for energy or for weapons—the ethics of who possesses nuclear power is therefore additionally a question of human trustworthiness in using something for a good purpose, not simply a question of how trustworthy nuclear power is itself as a resource. In similar manner, the extent to which we as citizens can pursue common goals ethically carries over into the likely use we will make of new opportunities and new technologies. It may be that an army is not ready to pass through a land of gold, for instance, if it means it will likely destroy the local population and steal the gold. Another army may be reliable enough ethically to pass through without stealing (an assessment that is separate from the question of whether gold itself is good or bad). Evaluating our societal trustworthiness for *participatory pursuit of the computerised optimal* is therefore of the essence, and forces us to ask what kind of moral rules and norms we require from ourselves as the opportunities afforded by these new technologies develop. Are we up to the mark? Will we

⁵³ Leverhulme Centre for the Future of Intelligence, 2019a.

exercise civic virtue? What strategies for helping ensure coordination and the pursuit of common goals will be required? What ideals of citizenship must be built for the networked age?

A helpful resource for answering these questions can be found in the rich debate on ethics, normative heuristics and incentives for solving collective action problems. The philosopher Alasdair MacIntyre states, ‘From Hobbes onward the psychological problem had been posed, Why should men do other than act to their own immediate advantage?’⁵⁴ Answers to this question are evident not only among philosophers, political scientists and economists but also in the many ways in which humans have gone about solving collective action problems practically: founding institutions and setting rules of engagement to establish boundaries on what is allowed, regularising incentives and punishments to encourage certain types of conduct and dissuade others, nurturing cultures and customs that narrate and explain the link between good actions and good outcomes.⁵⁵ While it is true that institutions and cultures in this sense work, they are nevertheless fragile and often fall short in achieving what is required. Human civilization is an almost constant process of rule-setting and rule-evaluation in tandem with critical reflection of our overall purpose and goal. Because the networked age challenges the very constitution of our social fabric, it likewise challenges these methods of rule-setting and rule-evaluation and asks us to think afresh on the kind of collectivity we are. The networked age revolutionises our way of being and way of belonging, requiring us to look deeper into our ability to engage in institutional design for the common good. It is, ultimately, a requirement that we look into the way we generate our ethical systems, a sort of meta-ethics on how we become the kind of

⁵⁴ MacIntyre, A., *A Short History of Ethics: A history of moral philosophy from the Homeric Age to the twentieth century* (London: Routledge, 1966), p. 179.

⁵⁵ On institutional design as a means for solving collective action problems, see Ostrom, E., *Understanding Institutional Diversity* (Princeton: Princeton University Press, 2005).

people that produce strong ethical systems over time and in new settings—in this case generating ethical systems able to withstand rupture to our ontology as a community.

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

1. They pertain to our basic appetites (food, drink, shelter).
2. They help establish our status (gold watches, tattoos, academic titles).
3. They improve us and those around us (exercise, advice, giving, justice).
4. They are enjoyable in themselves (play, art, contemplation).
5. They take us on a journey to higher goods (religion, study, travelling).

All these can be subsumed under the more general reason that we seek to be happy,⁵⁶ but the differentiation here is nevertheless useful for showing that there can be tensions between the choices we make, a fact too easily dismissed when the goods we seek are thought reducible to a single utility framework.⁵⁷

Economic and technological developments narrow the distance in obtaining the things that we value. They do not, however, help us prioritise which of them should be most valued, nor do they provide a sounder basis for our method of collective self-direction. They often help with our speed of communication and the certainty of our

⁵⁶ Aristotle, 2004.

⁵⁷ As in Bentham, J., *An Introduction to the Principles of Morals and Legislation* (London: W. Pickering, 1823).

memory. They do not establish a hierarchy of values, but rather respond to the hierarchy of values we set.

Satisfaction of our basic appetites can often provide a natural hierarchy of valuation (for example, today my thirst is more urgent than my hunger, but my hunger more urgent than my need for shelter). Greater productivity and automation means, however, that the satisfaction of basic appetites can in some parts of the world increasingly be taken as a given. This places relatively greater emphasis on the other four forms of valuation in determining a hierarchy to order individual lives, the market economy, and the common good. The four additional reasons for valuing things are, however, potentially never-ending. And a shift in production towards them increases the anxiety in deciding among their relative importance. Capitalism becomes, in this sense, an engine of anxiety-creation over our valuation priorities by making choices and options multiply. It becomes harder to be sure of how our choices fit into an authentic overall pursuit of happiness. Objective valuation of things as “needs” becomes less and less easy to assume as the economy specialises its marketing towards what we, as individuals, are likely to mistake as needs given our particular habits, tastes, experiences and psychologies. Personalisation of data-gathering techniques accelerates the project of making wants look like needs, and the idea of a “need” turns into “fear of what you are in danger of missing out on” if you make the wrong consumption choices (see Section 3.b).⁵⁸ Individual notions of status, improvement and enjoyment encourage ever more personalisation of the laws of supply and demand, escalating the pressure on, and evolutionary relevance of, each person’s choices and

⁵⁸ In addition, see interesting discussion by Cal Newport about how social media platforms depend on a notion of “missing-out” rather than “need” in soliciting permanent participation. Newport, C., *Deep Work: Rules for Focused Success in a Distracted World* (London: Piatkus, 2016), pp. 184-8.

valuations, and deescalating interpersonal moral correction and advice.

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

Making choices has often been considered a means of human evolution. Each choice we make stems from our perspectives or intentions and from the quality of consciousness that we bring to our thoughts, feelings, and actions. Conscious evolution, through making responsible and positive choices, can be a beneficial path.⁵⁹

Amartya Sen concurs, arguing that it is in the ranking of our preferences that we express our moral judgments.⁶⁰

How do we build our moral reasoning, and how can it be built at the level of a networked community? MacIntyre provides a suggestion in his most recent book *Ethics in the Conflicts of Modernity* by pointing to the way in which competing desires act as a starting point for ethical reflection by forcing a person to attempt a prioritisation of certain types of goods over others: ‘What small children desire they try to get. But [...] as they grow older they learn to delay satisfying some of their desires and develop desires that can be satisfied only at some

⁵⁹ Templeton, J., *Wisdom from World Religions: Pathways toward Heaven on Earth* (Philadelphia: Templeton Foundation Press, 2002), p. 261.

⁶⁰ Sen, A., ‘Rational Fools: A Critique of the Behavioral Foundations of Economic Theory’. *Philosophy & Public Affairs*, Vol. 6, No. 4 (1977), pp. 317-344, p. 337. See also Murphy, J. B., ‘Nature, Custom, and Reason as the Explanatory and Practical Principles of Aristotelian Political Science’. *The Review of Politics*, Vol. 64, No. 3 (2002), pp. 469-495.

time, even some distant time, in the future.’⁶¹ The original “marshmallow test” was designed to evaluate the change with age in children’s capacity for deferring gratification.⁶² Over the course of our lives in which ‘objects of desire have multiplied’ we build up a kind of history of desires, some of which are ‘transformed, others replaced.’⁶³ At certain times, when ordinary life is radically disrupted:

it requires little reflection to recognize that if I am to answer the question “What shall I do?” I had better first pause and pose the question “What is it that I want?” Somewhat more reflection is needed to recognize that I also need to think critically about my present desires, to ask “Is what I now want what I want myself to want?” and “Do I have sufficiently good reasons to want what I now want?” and still further reflection to recognize that I will be likely to go astray in answering these questions if I do not also ask how I came to be the kind of person that I now am, with the desires that I now have, that is, to ask about the history of my desires.⁶⁴

There are ways in which this process of critical self-reflection over our deepest desires are being altered in the networked age. We often receive direct communicative feedback about our desires in a way we did not use to (e.g. if I were to post a picture of a crocodile-skin jacket on Instagram and write, “Thinking of buying one...”). Humans have always received feedback about their intentions, but the feedback is now possible from many more people simultaneously, like speaking in the town hall with everyone possibly interested and possibly severely

⁶¹ MacIntyre, A., *Ethics in the Conflicts of Modernity: An Essay on Desire, Practical Reasoning, and Narrative* (Cambridge: Cambridge University Press, 2016), p. 5.

⁶² Mischel, W. & Ebbesen, E. B., ‘Attention in Delay of Gratification’. *Journal of Personality and Social Psychology*, Vol. 16, No. 2 (1970), pp. 329-337.

⁶³ MacIntyre, 2016, p. 3.

⁶⁴ *Ibid*, p. 4.

critical. At times the feedback is more cursory in nature because it requires less time per engagement, but less time per engagement does not always translate into less of a critical edge. The speed and type of communication in forming our desires is changing, and yet these communicative changes do not supplant our basic desires nor necessarily alter our underlying human values.

Nevertheless, our *history of desires* is changing in our networked age in that our desires are being memorialised more accurately than ever before. This is most apparent among what Shannon Vallor describes as ‘devotees of the Quantified Self’, who ‘employ mobile, wearable, and/or biometric sensors such as the FitBit and Jawbone devices, smartphone apps such as Moves and Chronos, video cameras, and a range of other devices to measure, track, analyse, and store volumes of recorded data concerning an ever-expanding list of personal variables.’⁶⁵ The idea is that one can collect streams of data about oneself and then take ownership of that data in using it to help make better choices. In her book *Technology and the Virtues*, Vallor debates whether this amounts to a kind of moral self-cultivation, under the idea that the good life is, in part, an examined life. Can a collection of data about oneself through apps help offset the danger that ‘the unexamined life is not worth living’⁶⁶? Vallor’s view is that, unfortunately, it cannot. While she is sympathetic to the attempt, she believes that attention towards moral goods is about focusing-in on those goods and reducing the noise of everything else going on. ‘As any philosopher of perception or cognitive scientist knows,’ she writes, ‘attention is as much about the ability to *screen out* information as it is about taking it in; in fact, the former capacity enables the latter.’⁶⁷ So, if I use an app to count my steps, but my number of steps

⁶⁵ Vallor, S., *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (New York: Oxford University Press, 2016), pp. 198-9, 201.

⁶⁶ Plato, *Apology*, 38a5. In Plato, *Complete Works* (Indianapolis: Hackett Publishing Company, 1997), edited by John M. Cooper, p. 33.

⁶⁷ Vallor, 2016, p. 200 (emphasis in original).

does not usefully correlate with a life well lived, in what way does that data help my moral development? Could it even distract me from asking the bigger questions of life?

A number of ethics apps try to explicitly improve one's moral development. Evan Selinger and Thomas Seager explain that most are about ethical advice for particular purchases (e.g. whether a product is environmentally friendly), or else directing one's behaviour towards a sense of the good through a combination of nudging, quantification and gamification.⁶⁸ An interesting example is the app 'GPS for the Soul', which tracks one's heart rate as an indicator of stress level. When stress is too high, the app will 'connect you to whatever you need to get to a place of balance'⁶⁹—the images, music and files that help you recentre. As the founder argues, 'the solution to the problems created by technology isn't anti-technology, but more and better technology.'⁷⁰ The app seeks to relativize the surrounding noise that is negatively affecting one's peace of mind. It does not aim at sparking critical self-reflection, which might of course make the person feel worse about themselves.

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

⁶⁸ Selinger, E. & Seager, T. P., 'Digital Jiminy Crickets: Do apps that promote ethical behaviour diminish our ability to make just decisions?' *Slate* (13 Jul 2012). <https://slate.com/technology/2012/07/ethical-decision-making-apps-damage-our-ability-to-make-moral-choices.html>.

⁶⁹ Huffington, A., 'GPS for the Soul: A Killer App for Better Living'. *HuffPost* (16 Apr 2012). https://www.huffpost.com/entry/gps-for-the-soul_b_1427290?guccounter=1.

⁷⁰ *Ibid.*

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

Just as personal ethics can, in this way, be built through working through one's competing internal desires, so too can civic morality be built through working through competing social desires. While clarity over the nature of politics is something one may find hard to draw out of MacIntyre's own work,⁷¹ it is easy enough to make the case more generally that conflict over desires happens at the societal level too, and that such tension and interpersonal competition provides an important justification for institutional rule.⁷² Ultimately, institutions 'are the humanly devised constraints that structure political, economic and social interaction.'⁷³ They provide these constraints as a way of managing at a more macro level our competing desires, directing our energies towards common goals, common goods.

⁷¹ See the criticism of how MacIntyre's ethics fails to relate sufficiently to a theory of politics in Duff, A., 'The Problem of Rule in MacIntyre's Politics and *Ethics in the Conflicts of Modernity*'. *Politics & Poetics*, Vol. 4 (2018), pp. 1-21; Sigalet, G., 'Waldron's Challenge to Aristotelians: On the Political Relevance of Moral Realism'. *Politics & Poetics*, Vol. 4 (2018), pp. 1-23.

⁷² See, for example, North, D. C., *Structure and Change in Economic History* (New York: W. W. Norton, 1981); North, D. C., *Institutions, Institutional Change and Economic Performance* (Cambridge: Cambridge University Press, 1990).

⁷³ North, D. C., 'Institutions'. *Journal of Economic Perspectives*, Vol. 5, No. 1 (1991), pp. 97-112, p. 97.

d. Rebuilding our citizenship

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

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

Chapter 3

The changing relationship
between time & attention



The Changing Relationship Between Time & Attention



a. Being only one citizen

Changes currently afoot as to what our human community *is* and *is for* require us to look anew at the status of the person-in-community and the dilemma of being *only one citizen as the human community grows and becomes ever more interconnected*. As my community grows, I may not grow; as my community becomes more interconnected, I may not become as interconnected as everyone else. As Zizi Papacharissi writes, '[o]nline technologies thrive on collapsing public and private boundaries thus affording opportunities for expression that may simultaneously empower and compromise individuals.'⁷⁴ Analysis of the internet, digital technologies and artificial intelligence has so far tended to look at how they change society and the world as a whole,⁷⁵ underappreciating the fact that in the midst of any change I remain just one person. From that individual citizen point of view these areas of progress may mean that *my own contribution* and *my own sense of self* reduce, and that reduction becomes accelerated in line with the accelerated erosion of interpersonal barriers. As Taina Bucher remarks, 'participatory subjectivity is not constituted through the imposed threat of an all-seeing vision machine, but by the constant possibility of disappearing and becoming obsolete.'⁷⁶ It becomes

⁷⁴ Papacharissi, Z., *Affective Publics: Sentiment, Technology, and Politics* (Oxford: Oxford University Press, 2015), p. 94.

⁷⁵ Schmidt, E. & Cohen, J., *The New Digital Age: Reshaping the Future of People, Nations and Business* (London: John Murray, 2013); Zuckerman, E., *Rewire: Digital Cosmopolitans in the Age of Connection* (New York: W. W. Norton, 2013).

⁷⁶ Bucher, T., *If... Then: Algorithmic Power and Politics* (Oxford: Oxford University Press, 2018), p. 17.

harder to maintain one's individuality; one constructs a buffer with the outside world to protect a sense of selfhood.⁷⁷

An important test case for understanding our changing individuality and moral reasoning is the introduction of artificial intelligence and algorithms into our public sphere decision-making, which is set to affect deeply our spirit of public service, justice, neighbourliness, democratic participation and collective moral reasoning. While numerous institutions and experts have focused on the changes technology and artificial intelligence are having on our mental development,⁷⁸ our economy and jobs,⁷⁹ and even the survival of the entire human race,⁸⁰ there has been less analysis of what the changes mean for our civic ideals and the normative underpinnings to our citizenship, and in particular in terms that make sense for the individual as just one citizen. The Leverhulme Centre for the Future of Intelligence—invariably a resource for comprehensive insight into the latest technological developments—researched the challenges that the future of artificial intelligence pose for democratic politics, but

⁷⁷ With thanks to Angela Franks for explaining this point at a workshop on the draft report in Harvard University, 5 Nov 2019. See also, Taylor, C., *Sources of the Self: The Making of the Modern Identity* (Cambridge: Cambridge University Press, 1989); Taylor, C., *A Secular Age* (Cambridge, MA: Belknap Press of Harvard University Press, 2007).

⁷⁸ Newport, C., *Deep Work: Rules for Focused Success in a Distracted World* (London: Piatkus, 2016); Wilmer, H. H., Sherman, L. E. & Chein, J. M., 'Smartphones and Cognition: A Review of Research Exploring the Links between Mobile Technology Habits and Cognitive Functioning'. *Frontiers in Psychology*, Vol. 8, No. 605 (2017), pp. 1-16; The Royal Society, *iHuman: Blurring lines between mind and machine* (London, Sep 2019).

⁷⁹ Points, L. & Potton, E., 'Artificial Intelligence and Automation in the UK'. *House of Commons Library, Briefing Paper*, No. 8152 (21 Dec 2017).

⁸⁰ Vinge, V., 'The Coming Technological Singularity'. *Whole Earth Review* (Winter 1993); Bostrom, N., *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014); Drexler, K. E., 'Reframing Superintelligence: Comprehensive AI Services as General Intelligence'. Technical Report, No. 2019-1, Future of Humanity Institute, University of Oxford (2019).

their focus was on (i) how well-equipped democratic institutions are to deal with the risks and opportunities of long-term transition to artificial intelligence; and (ii) what kinds of political agency might be attributed to new forms of artificial intelligence.⁸¹ That is a focus on artificial intelligence and how it raises questions for democracy. What it does not include are the kinds of questions democracy will likely raise for artificial intelligence. That is, how citizenship becomes an interactive force for managing and redirecting technological progress through civic ideals that help dictate the relationship between humans and machines.⁸²

Putting citizenship first throughout this report hopes to stimulate wider inquiry into what it means to do citizenship and civic engagement successfully in a context of rapidly-changing decision-making through algorithms and artificial intelligence. Humans flourish in their citizenship, both as a collectivity and also as participatory individuals. The need is for greater examination of the ways these macro changes in what our community is and is for, affect who I am and what I am for.

Some of the changes taking place are already quite clear:

1. A growing commitment to algorithmic decision-making in public administration due to its presumed objectivity, expertise and rapidity, in turn framing public participation by citizens as relatively sentimental, inept and slow⁸³; greater out-sourcing of

⁸¹ Leverhulme Centre for the Future of Intelligence, ‘Politics and Democracy’ (2019b). <http://lcfi.ac.uk/projects/ai-trust-and-society/politics-and-democracy/>.

⁸² See, for example, Beyer, J. L., *Expect Us: Online Communities and Political Mobilization* (Oxford: Oxford University Press, 2014); Penney, J., *The Citizen Marketer* (Oxford: Oxford University Press, 2017).

⁸³ Garton Ash, T., *Free Speech: Ten Principles for a Connected World* (London: Atlantic Books, 2016), pp. 364-7.

moral and judicial decision-making to algorithms, which are often presumed to be impartial, framing citizen decision-making as biased and individualistic (see Chapter 5).⁸⁴

2. Reduced habit among citizens of suspending judgment or developing political views in closed settings before sharing with the wider public (see Chapter 4).⁸⁵
3. Declining natural forgetfulness for bad or incorrect things said.⁸⁶
4. Curtailment of one's social network away from those with alternative viewpoints, especially during important political moments.⁸⁷
5. A weakening ability to boycott online services that hold one's data and history, and do not charge money for the service.⁸⁸

⁸⁴ Dafoe, A., 'AI Governance: A Research Agenda'. Future of Humanity Institute, University of Oxford (Aug 2018), p. 37; Bucher, 2018.

⁸⁵ Karpf, D., *Analytical Activism: Digital Listening and the New Political Strategy* (Oxford: Oxford University Press, 2017).

⁸⁶ On the relationship between forgiveness and forgetfulness, see Volf, M., *The End of Memory: Remembering Rightly in a Violent World* (Cambridge: W. B. Eerdmans, 2006).

⁸⁷ Bello, J., 'The making and consequences of everyday political discussion networks'. DPhil thesis, University of Oxford (2011); Vallor, S., *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (New York: Oxford University Press, 2016), pp. 182-3; Papacharissi, 2015, p. 35. For counter-argument, see Dubois, E. & Blank, G., 'The echo chamber is overstated: the moderating effect of political interest and diverse media'. *Information, Communication & Society*, Vol. 21, No. 5 (2018), pp. 729-745.

⁸⁸ Zuboff, S., *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power* (London: Profile Books, 2019).

6. Increased quantification of public goals and application of utilitarian reasoning for balancing competing interests, as opposed to democratic deliberation over human values.⁸⁹
7. Growing disconnect between global, cosmopolitan moral views and local, traditional living; increased anxiety on both sides over who the ultimate decision-makers really are.⁹⁰
8. A burgeoning plurality of cybergovernance institutions and a weakening of the authority of national laws, in turn making the target of democratic advocacy less clear.⁹¹
9. Increased hate speech and misinformation; civic engagement narrowly focused on chastising those who have spoken hatefully or falsely.⁹²

These changes in some ways build on 20th century marketization and urbanisation of human relations,⁹³ but they also deviate in significant

⁸⁹ Urbina, F. J., *A Critique of Proportionality and Balancing* (Cambridge: Cambridge University Press, 2017).

⁹⁰ Gaventa, J., *Power and Powerlessness: Quiescence and Rebellion in an Appalachian Valley* (Urbana: University of Illinois Press, 1980); Goodhart, D., *The Road to Somewhere: The New Tribes Shaping British Politics* (London: Penguin, 2017); Collier, P., *The Future of Capitalism: Facing the New Anxieties* (London: Allen Lane, 2018).

⁹¹ Harari, Y. N., *Homo Deus: A Brief History of Tomorrow* (London: Vintage, 2015), p. 436; Owen, T., *Disruptive Power: The Crisis of the State in the Digital Age* (Oxford: Oxford University Press, 2015); Garton Ash, 2016, p. 352.

⁹² Baldwin-Philippi, J., *Using Technology, Building Democracy* (Oxford: Oxford University Press, 2015); Anderson, C. W., *Apostles of Certainty: Data Journalism and the Politics of Doubt* (Oxford: Oxford University Press, 2018); Nyabola, N., *Digital Democracy, Analogue Politics: How the Internet Era is Transforming Politics in Kenya* (London: Zed Books, 2018); Woolley, S. C. & Howard, P. N. (eds.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (Oxford: Oxford University Press, 2018).

⁹³ Polanyi, K., *The Great Transformation: The Political and Economic Origins of Our Time* (Boston: Beacon Press, 1944). See also the discussion in Cohen, J. E., 'Law for the Platform Economy'. *U.C. Davis Law Review*, Vol. 51 (2017), pp. 133-204, p. 135.

ways through extraordinary personalisation of the laws of supply and demand and a decentring of the urban experience through the creation of multiple overlapping social networks. Crudely-speaking, 20th century industrialism helped everyone buy the same things; 21st century personalisation of the laws of supply and demand gives everyone their own, individualised economy, in line with their online data profiles.⁹⁴ In threatening our ability to be citizens and act in pursuit of a common good, these trends risk a breakdown in social solidarity and a growth in arbitrary authority. To repeat, distrust between the “haves” and the “have nots” is not simply an economic question of appropriate distribution⁹⁵ but a growing existential question about who the decision-makers really are.

Question: “If artificial intelligence knows your political preferences better than you, should it cast your vote for you in the way most optimal?” With the instinctual answer “no”, do you commit to an ideal of citizenship that is sentimental, inept and slow? It seems whichever way you answer, you as *just one citizen* are no longer the optimal decision-maker for pursuit of the common good. Perhaps an algorithm should even answer this question for you.

The current chapter focuses on the dilemmas the networked age presents for *being only one citizen* as one’s community changes shape and as traditional methods of decision-making are cast in greater and greater doubt. Specifically, three issues are tackled. The first is the way in which the internet, in particular, reshapes one’s network to give radically divergent impressions about how, and who, one is, as compared to those one networks with. The internet has in-built bias in prioritising digitised data as an essential comparative metric,

⁹⁴ See the discussion of how personal data is amassed and used by Amazon in Kelion, L., ‘Why Amazon knows so much about you’. *BBC News* (Feb 2020).

<https://www.bbc.co.uk/news/extra/CLQYZENMBI/amazon-data>.

⁹⁵ As in Piketty, T., *Capital in the Twenty-First Century* (Cambridge, MA: Harvard University Press, 2014).

reinforcing the need that lone citizens generate digitisable outputs. The second issue is the way in which the relationship between time and attention is under alteration in the networked age. Specifically, attention is becoming a scarce resource, a precious commodity. At the same time, unlike any other precious resource in history, we each enjoy a certain equivalence in endowment by means of all being in the present, creating a spontaneous equality among us in the extent to which time can be made interactive with who we are through the giving of our attention. That makes attention quite an unusual type of resource to exploit, politically-speaking, because it lies in people, and equally so.⁹⁶ The equality of its endowment adds impetus to the participatory, resource-releasing interactions demanded from the lone citizen in the networked age.

The third issue to be tackled in this chapter is the way in which the institutional basis to economic participation is shifting as firms and companies give way to platforms and communities. These legal and business shifts are important not only at the macro level but also in connection to the micro level of ordinary people going about their everyday lives. It has long been established that economic activity helps shape social norms and manners,⁹⁷ and so it is important to

⁹⁶ Some, of course, have greater attention spans than others. But from the point of view of getting to know a person's preferences and habits as they are, the normal level of attention one tends to deploy is appropriate for evidencing the kind of person one is, whatever that level is. This equivalence in ability to represent oneself through one's behaviour and action is what gives the "equality of endowment" of attention for the networked age.

⁹⁷ See, for example, Jean Ensminger on the views of the Baron de Montesquieu, as well as her own empirical findings. Ensminger, J., *Making a Market: The institutional transformation of an African society* (Cambridge: Cambridge University Press, 1996); Ensminger, J., 'Market Integration and Fairness: Evidence from Ultimatum, Dictator, and Public Goods Experiments in East Africa'. Ch 12 of Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E. & Gintis, H. (eds.), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies* (Oxford: Oxford University Press, 2004), p. 357.

identify the ways in which these changing business models may bring about changes in the ways we interact as human beings.

In the midst of all this is the ongoing frustration that, as a citizenry, we act together haphazardly. We seem capable of neither perfect unanimity nor perfect individuality. We come together and crash apart, in waves; all the time our decisions about what to do next are erratically connected to our expectations of what everyone else will likely do next.⁹⁸ In discussing what the networked age is bringing about, it is therefore of the essence to close the distance in understanding what the macro changes make me feel, be and do, as only one citizen trying his or her best to be happy.

b. The friendship paradox & FOMO (Fear Of Missing Out)

A good beginning to this dilemma of being a lone citizen in the networked age is purely mathematical. Here is the problem: when you are only one person (as most of us seem to be), you quickly become enmeshed in a paradox that the most common interaction you have is not with the average person but with the more frequently encountered person. Mathematically-speaking, this works the same as the difference between the mode and the mean. The mode describes the single most popular option, while the mean describes the arithmetic average. The mode can look like the average because so many people go for it, but in fact the mean may be different to it, depending on the spectrum of choices. It can help to give an example. The modal salary bracket is typically lower than the mean salary bracket. The reason for this is that it is a minority of individuals who earn extremely large

⁹⁸ Granovetter, M., 'Threshold Models of Collective Behavior'. *American Journal of Sociology*, Vol. 83, No. 6 (1978), pp. 1420-1443; Burbidge, D., *The Shadow of Kenyan Democracy: Widespread Expectations of Widespread Corruption* (Burlington: Ashgate, 2015).

salaries, while the rest of the population tend to earn more similar salaries. When one calculates the arithmetic mean, it is the sum of all the salaries divided by the number of people. The mode, instead, is calculated by counting the number of people in each salary range, and identifying that which is most usual. Normally, because there tend to be more people with similar salary levels at the lower end of the salary scale, the modal salary will be less than the mean salary. The difference generates the shocking headline that most people earn less than the average wage (“average” in terms of “mean”), and the same logic creates the even more shocking headline that most people have more than the average number of limbs.

These various methods of calculation provide different ways of trying to understand what is “normal” or “most common”, and they become very interesting when looked at sociologically. Specifically, what counts as the mode or arithmetic mean can become an important comparative metric for assessing how one is doing as an individual. For example, knowing that you are paid *below average* or *above average* may make you feel differently about the respect other people have for the work that you do. If the modal salary is lower than the mean salary, society is in the dilemma that the most common salary level is also below average (i.e. below the arithmetic mean).

The dilemma extends from comparative salary assessments to all kinds of ways in which human beings compare themselves with each other and generate expectations about what is acceptable behaviour. In this sense, the terms “normative” and “normal” align quite closely. Much of what we consider decent behaviour is a product of its accepted normality. The test case is when one enters a new culture and has to re-learn one’s sense of normal, and avoid assuming negative intent when seeing people acting in a way one’s previous culture would have abnormal (a-normal).

Because, as indicated, the networked age is bringing about changes to the size and communicative apparatus of one's community, it also has major influence on the way humans go about their interpersonal comparisons. To repeat, as my community grows, I may not grow; as my community becomes more interconnected, I may not become as interconnected as everyone else. Even though I may not change much in my own routines, my sense of the average person in my community is under radical alteration, which brings about significant changes to me in terms of comparative assessments of myself.

A helpful resource for thinking through the meaning of this can be found in the thought of sociologist Mark Granovetter. In a 1973 article, 'The Strength of Weak Ties',⁹⁹ Granovetter pointed out that the people most likely to find information about future job opportunities are not those who have good, close friends but those who have many loose friends. It tends to be that among one's closest friends, many of them also know each other independently of oneself, and therefore they are privy to the same sort of information which gets relayed between them, without much new coming in. However, one's loose friends tend to be part of other networks involving very different sources of information. Knowing what is coming down-the-pipeline is therefore about staying in touch with lots of people occasionally, rather than relying on close or deep relationships with only a few.

Granovetter's insight applies much more broadly than information about new jobs (which readers who already have decent jobs may be keen to hear). The overall theory is that one's information and understanding is heavily influenced by the sort of network one is a part of, with some people enjoying dense, close ties within a limited informational network and others enjoying wide, loose ties within an expanded informational network. It has been found that the world is

⁹⁹ Granovetter, M., 'The Strength of Weak Ties'. *American Journal of Sociology*, Vol. 78, No. 6 (1973), pp. 1360-1380.

connected by four degrees of separation.¹⁰⁰ What that means is that everyone in the world can be reached by asking a friend-of-a-friend-of-a-friend-of-a-friend. Of course, one has to ask the right friend each time, so the connections do not feel this close as a matter of course. Insofar as it is correct, it implies that you are, at any one time, only a handful of people away from being friends with the Supreme Leader of North Korea. Tell your friend to disarm.

In terms of who your friends-of-friends are, it is possible to compute visualisations of the social network on Facebook, as in the below example from Caleb Jones.¹⁰¹ The image shows the friendship links between one's friends, giving a sense of the density of links between different groups, and who are especially effective at "bridging" one's social network to new sets of people.

¹⁰⁰ Jessica Leber notes the previous consensus on six degrees of separation but finds it to now be closer to four. Leber, J., 'You Are Connected To Everyone On Earth By Just 4 Degrees Now'. *Fast Company* (28 Oct 2013).

<https://www.fastcompany.com/3020687/you-are-connected-to-everyone-on-earth-by-just-4-degrees-now>.

¹⁰¹ Jones, C., 'How to Visualize Your Facebook Friend Network'. *All Things Graphed* (28 Aug 2014).

<http://allthingsgraphed.com/2014/08/28/facebook-friends-network/>.

Mutual friend visualisation

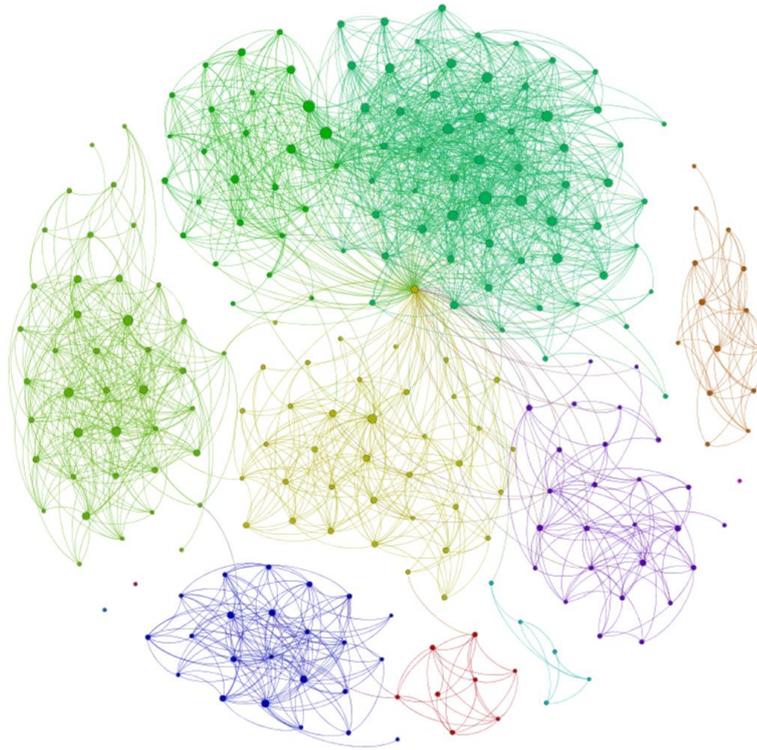


Figure 3.1 An example of mutual friend visualisation, emphasising the friends that bridge between social groups

Researchers at Facebook used data on interpersonal networks to empirically calculate whether a previously assumed six degrees of separation holds among its users. Are you six friends away from knowing the entire 2.38 billion users of Facebook? The researchers found, surprisingly, that in fact Facebook users are only spread by an arithmetic mean of 3.57 degrees of separation.¹⁰² That is, for the

¹⁰² Bhagat, S., Burke, M., Diuk, C., Filiz, I. O., Edunov, S., 'Three and a half degrees of separation'. *Facebook Research* (4 Feb 2016).

average Facebook user fewer than three friends-of-friends can put them in touch with anyone else on the platform.

Estimated average degrees of separation between all people on Facebook

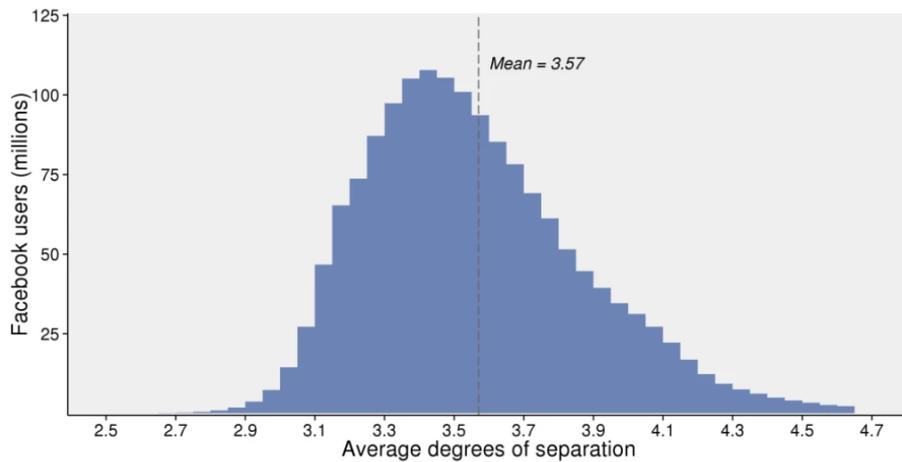


Figure 3.2 Estimated average degrees of separation between all people on Facebook, showing how many friends will link one with any other part of the network

Now, it may be thought that the fewer than expected degrees of separation is because the sorts of people who use Facebook are more similar to each other and share the same sorts of friends already than the world population tends to in general. However, the researchers compared these 2016 results with an earlier study conducted in 2011 and found that in fact the degrees of separation had decreased over this time period from 3.74 to 3.57 despite the rapid increase in numbers of Facebook users that had occurred. Interestingly, therefore, the

<https://research.fb.com/blog/2016/02/three-and-a-half-degrees-of-separation/>.

comparison suggests that as the user base of Facebook has increased, the degrees of separation between users has decreased. All this suggests that it is not that Facebook is consolidating links between a pre-existing group but that it is, in fact, tracking the more general increased connectivity among the world's population. Paradoxically, the more people there are, the easier it is that everyone knows each other.

Most analyses of the networked age's connectivity tend to stop there. The focus they make is on the density of our interpersonal links, using data available ready-made by social media platforms. The analysis suffers, however, from having no real sense of people's connections prior to the networked age, meaning there is little qualitative assessment of *how well* people's interpersonal connections are being developed now, compared to before. This is a difficult comparison to make as it involves value-based assessment of the quality of interactions and categorisation of the types of information shared online versus offline. Notwithstanding this, it can be helpful to at least evaluate the sociological dynamics at play in these new networks and the quality of interactions, with focus on the relationship between the lone citizen and the changing size and nature of his or her wider community.

Perhaps the most important sociological feature of the networked age's developments lies in the "friendship paradox". This is the theory—again grounded in mathematics—that we each, on average, have fewer friends than our friends have. What does this mean? It means that if we were to ask our friends how many friends they have, we would find that they on average tend to have more than we do, and that this phenomenon applies to most of us. How is that possible? The explanation lies in the way friendship groups are, in fact, biased sub-samples of all the people out there in the world, in that they over-represent the well-connected. Popular people are over-represented in friendship groups and unpopular people are under-represented. A

sample friendship group—such as one’s own—will therefore tend to contain more of the popular people than the unpopular people. That means that, on aggregate, popular people will be over-represented in friendship groups (often counted many more times than unpopular people are). On average, therefore, one’s friends will tend to have more friends than oneself.

In the networked age, the increased connectivity between people accelerates this phenomenon, such that the friendship paradox of one’s friends tending to have more friends than oneself increases to unprecedented levels. It can be no other way, because increasing the links between people will almost invariably multiply the connectivity of those already well connected by a factor higher than those who are less well connected. In the networked age, one’s friends have even more friends compared to oneself than before, widening the gulf. A stop to this dynamic can come when everyone is friends.¹⁰³ Approaching that end point, the most well-connected will of course see radical reduction in their comparative greater connectivity, restoring equality. But it is not clear if humans will ever reach that point due to the limits on human cognitive capacities¹⁰⁴ and the enduring desire for ever-deepening, nurturing relationships which prioritise exclusivity and the giving up of one’s time and attention to particular people.¹⁰⁵ So far, digital technologies grant us improved capacity to expand outwards, but do precious little to make us more

¹⁰³ The other way of stopping the dynamic is to ensure that absolutely everyone has the same number of friends as everyone else.

¹⁰⁴ Robin Dunbar argues that the human brain gives a natural limit of about 150 stable friends. Dunbar, R. I. M., ‘Coevolution of neocortical size, group size and language in humans’. *Behavioral and Brain Sciences*, Vol. 16, No. 4 (1993), pp. 681-735.

¹⁰⁵ A perennial feature of human nature, it seems. See reflections on the differences between humans and ants in Christakis, N. A., *Blueprint: The Evolutionary Origins of a Good Society* (New York: Little, Brown Spark, 2019); and Free Thoughts Podcast, ‘The Evolution of Cooperation (with Nicholas A. Christakis)’ (14 Feb 2020). <https://www.libertarianism.org/podcasts/free-thoughts/evolution-cooperation-nicholas-christakis>.

nurturing or more exclusive. They help us stay in touch, not be deeper friends. Our friendship reach enjoys digital tentacles spread across the world, but who can say they have deeper friendships than their ancestors had?

In terms of empirical studies, Nathan Hodas and colleagues confirm the friendship paradox among Twitter users, identifying that the core dynamic holds true for more than 98 percent of Twitter users.¹⁰⁶ That is, the vast majority of Twitter users follow people who have more followers than them. The authors also find the phenomenon extends to the posting of viral content, such that those one follows are much more likely to post viral content than oneself: on average, 86 percent of Twitter users post less viral content than those they follow.¹⁰⁷ Johan Bollen and colleagues push analysis even further by asking whether there is a “happiness paradox” of one’s friends being, on average, more happy than oneself.¹⁰⁸ Measuring happiness levels among a network of 39,110 Twitter users, the researchers find that a happiness paradox is indeed evident, meaning that most people in one’s Twitter network¹⁰⁹ tend to be happier than oneself. This phenomenon reinforces, perhaps, the strange way in which social media tends to report positives while traditional media tends to report negatives. There is little relational dynamic at play in traditional media so what matters is purely interest in what is said. In social media one tends to follow both what is of interest and also who is attractive for forming social relationships with, in turn over-representing happy people. And

¹⁰⁶ Hodas, N. O, Kooti, F. & Lerman, K., ‘Friendship Paradox Redux: Your Friends Are More Interesting Than You’. *Association for the Advancement of Artificial Intelligence* (2013), pp. 1-9.

¹⁰⁷ Ibid, p. 4.

¹⁰⁸ Bollen, J., Gonçalves, B., van de Leemput, I. & Ruan, G., ‘The happiness paradox: your friends are happier than you’. *EPJ Data Science*, Vol. 6, No. 4 (2017), pp. 1-10.

¹⁰⁹ Twitter does not have “friends” like Facebook does, so the authors discuss the happiness of Twitter users in terms of the networks of users that are engaged in reciprocal following (“friends” in the sense that person A follows person B and person B also follows person A).

happy people tend to say happy things, making social media cover happier topics on average than traditional media used to.

As is probably already clear, the expansion of the friendship paradox to all types of digital interconnection in the networked age carries with it a great risk: a shrinking of the comparative happiness and sense of relevance for most people as the density of their networks increases. Such concerns link with wider concerns over increased loneliness—just when it seemed we were getting closer to each other than ever before—and increased social anxiety by means of many rapid comparisons with those who seem to be doing better than oneself. As the prior analyses show, there are less than two percent of Twitter users who follow people who have, on average, fewer followers than they have. The paradox becomes a trap when one thinks that the way to combat one's comparative disconnect is to get more online friends, to follow more popular people, or to share more viral content. In fact these draw one ever deeper into the friendship paradox, increasing one's disappointments. From the individual's point of view, the apparent progress of the networked age in generating greater connectivity in fact means that *my own contribution* and *my own sense of self* reduce, a reduction accelerated in line with the accelerated erosion of interpersonal barriers.

It has been additionally observed that the networked age is encouraging a “Fear Of Missing Out” (FOMO), something which should not be separated from wider analysis of the friendship paradox and what it means to be in a world where those one knows are, on average, happier, more successful and better connected than oneself. FOMO is the concern that one will not get the opportunities that others will if one is removed from a particular network or subscription. It lies behind many explanations ordinary people give as to why they stay members of social media platforms even though they

also feel it consumes a lot of their time.¹¹⁰ They fear missing out on connections with long-lost friends, important news updates and work opportunities. Removing oneself from a network creates the painful possibility of not knowing how wonderful the thing that one missed out on would have been.

FOMO can be combined with the friendship paradox, such that an impression that one will indeed miss out on a lot is heightened by the sense that there are a great number of opportunities others online seem to be connecting with. In fact, like many of the factors discussed above, one's friends tend to over-represent the opportunities available because they are, on average, better connected to opportunities than oneself. Not properly appreciating this dynamic, most people in the networked age feel that these dynamics are peculiar to them, perhaps because they joined social media relatively late in the day or because they live less attractive or desirable lives than the majority of people online. Our assessment of our self-worth gets wrapped-up in the friendship paradox and FOMO as our comparisons with others gives us poor feedback on how we really are as compared to the average person online.

Seth Stephens-Davidowitz takes the controversial step of sociologically analysing what people tend to use the internet for privately, offering a sense of the non-social activity online. On the topic of how this use compares with social media use, he writes:

Facebook is digital brag-to-my-friends-about-how-good-my-life-is serum. In Facebook world, the average adult seems to be happily married, vacationing in the Caribbean, and perusing the *Atlantic*. In the real world, a lot of people are angry, on supermarket checkout lines, peeking at the *National Enquirer*, ignoring the phone calls from their spouse, whom they haven't

¹¹⁰ Newport, 2016, pp. 184-8.

slept with in years. In Facebook world, family life seems perfect. In the real world, family life is messy. It can occasionally be so messy that a small number of people even regret having children. In Facebook world, it seems every young adult is at a cool party Saturday night. In the real world, most are home alone, binge-watching shows on Netflix. In Facebook world, a girlfriend posts twenty-six happy pictures from her getaway with her boyfriend. In the real world, immediately after posting this, she Googles “my boyfriend won’t have sex with me.”¹¹¹

While neither private use of the internet nor social media use can give an overall picture of who one is, it is certainly the case that social media use in particular tends to involve a great deal of comparison with others online as posts are curated for wider consumption. At play in such comparisons are a number of difficult mathematical paradoxes challenging the sense of self of the person-in-community. What can explain the overall effect these are having and the kind of civic response that will reverse their dynamics?

c. Time & attention as scarce resources

At the core of all of these tensions is a commodification of one’s attention span such that the lone citizen faces increased pressure over time for the opportunity cost of focusing on one thing and not another. When this combines with greater doubt over one’s place in an ever-shifting community, the opportunity cost of the use of one’s attention has direct effect on the bigger questions of *who I am* and *what I am for*.

¹¹¹ Stephens-Davidowitz, S., *Everybody Lies: What the Internet Can Tell Us About Who We Really Are* (London: Bloomsbury, 2017), pp. 152-3 (emphasis in original).

Traditionally, politics and economics has always been framed in terms of managing the *scarcity of resources*. In the networked age it would be better to say the need lies in managing *scarcity of attention*. The commodification of personal data—what Shoshana Zuboff calls behavioural surplus¹¹²—places value in any metrics that can reliably be associated with what people do and where their interests lie. For online activity, this is dictated by what we give attention to, which therefore forms the basis to our online profile and the behavioural futures market. At the same time, the lust after users' attention creates an online environment oriented towards ever-shorter attention spans when the quest becomes winning away someone's attention from other things, rather than keeping their attention on a single thing. As Henry Kissinger writes:

Inundated via social media with the opinions of multitudes, users are diverted from introspection; in truth many technophiles use the internet to avoid the solitude they dread. All of these pressures weaken the fortitude required to develop and sustain convictions that can be implemented only by traveling a lonely road, which is the essence of creativity. [...] Political leaders, overwhelmed by niche pressures, are deprived of time to think or reflect on context, contracting the space available for them to develop vision.¹¹³

Although there is insufficient space here to enter into debates on the effects of technology for cognitive capacity and work habits, it is important to draw out the way in which the economic scarcity of attention lies behind many of the interpersonal dilemmas present in online community formation. Our ability to maintain our attention on what we value seems to decline just as it seems to be increasing among

¹¹² Zuboff, 2019, pp. 74-82.

¹¹³ Kissinger, H. A., 'How the Enlightenment Ends'. *The Atlantic* (Jun 2018). <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>.

those we follow and are friends with, because their *display of attentiveness* over-represents attentiveness towards what is happening in the online world. The greater pressure being placed on the opportunity cost of our uses of attention prioritises how we manage our own time, creating a gap in genuine empathy for others:

The moral need for more human empathy is not a novelty of the digital age. But the need is *growing*, not just because some new habits of the information society might deplete or inhibit empathy, but because the increasingly networked and interdependent nature of the human family entails that we shall find ourselves exposed to ever more circumstances that seem to call for it. [...] Determining who we ought to feel empathy for, when, and to what degree, is a moral problem that is getting ever harder, not easier, to get right.¹¹⁴

It is difficult to draw any conclusions on where the empathy need is greatest. It is certainly the case, however, that a shortfall in empathy is likely when platforms encourage self-reflection of one's lack of success compared to others, rather than self-reflection of one's ability to assuage others' lack of success. To get a better sense of the role of platforms in shaping our sense of social need, it is necessary to explore how platforms themselves entice particular ways of being and acting in our networked age.

d. New categories of community, platform & institution in a networked age

Behind all of these changes to *who we are* and *what we do as only one citizen*, there are macro changes to the communicative platforms that

¹¹⁴ Vallor, 2016, p. 138 (emphasis in original).

mediate between us. It is important to resist the temptation to “determinism” that these shifts elicit: the human race does not simply change its habits and values based on the media that help organise the way we relate to one another. Nevertheless, there is clearly an influential role played by the structure and framing of platforms of mediation, making feasible new ‘imagined communities’¹¹⁵ and ‘affective publics’.¹¹⁶

Our social norms are deeply influenced by the economic frameworks we involve ourselves in. That was a point made long ago by Montesquieu, who famously wrote:

Commerce cures destructive prejudices, and it is an almost general rule that everywhere there are gentle mores, there is commerce and that everywhere there is commerce, there are gentle mores.

Therefore, one should not be surprised if our mores are less fierce than they were formerly. Commerce has spread knowledge of the mores of all nations everywhere; they have been compared to each other, and good things have resulted from this.¹¹⁷

Later in this report we explore what the changes of the networked age mean for the civic virtue of listening to one another, but it is necessary here to ask the broader question of whether these changes will be in line with what Montesquieu observed for 18th century commerce: will the networked age involve a diffusing of the knowledge of manners of

¹¹⁵ Anderson, B., *Imagined Communities: Reflections on the Origin and Spread of Nationalism* (London: Verso, 2016 [1983]), Revised Ed.

¹¹⁶ Papacharissi, 2015.

¹¹⁷ Montesquieu, *The Spirit of the Laws* (Cambridge: Cambridge University Press, 1989 [1748]), p. 338; Online Library of Liberty, ‘Montesquieu thought that commerce improves manners and cures “the most destructive prejudices” (1748)’ (2019). <https://oll.libertyfund.org/quotes/85>.

all nations, allowing choice of the best? Will the shift from firms and companies to platforms and online communities refine our human relations? Will such transition produce the type of network able to nurture its good points and weed out its bad points over time? That self-correcting hypothesis lay at the heart of neoclassical optimism in the free market¹¹⁸; and early interest in the internet's political potential shamelessly transposed the very same optimism about the natural human equilibrium arrived at through freedom onto cyberspace.¹¹⁹ But instead of John Perry Barlow publishing 'A Declaration of the Independence of Cyberspace', it would have been more accurate to publish 'A Declaration of the Interdependence of Cyberspace'.¹²⁰ The internet and the development of artificial intelligence have proved so intimately bound to the reciprocal, relational and reconstituting progress of the human race and human governance that its development has nothing of independence about it. And precisely because of its reciprocal, relational and reconstituting bind with humanity, the being and choices of the lone citizen remain of persistent relevance in defining its nature and predicting its trajectory.

To get even a start, therefore, on these questions of how and in which causal direction individuals and their institutional layout affect one another, it is necessary to get clear on what is really taking place in the formation of new categories of community, platform and institution in the networked age. Danah Boyd provides a helpful suggestion:

¹¹⁸ Smith, A., *The Wealth of Nations* (New York: The Modern Library, 2000 [1776]); Hayek, F. A., *The Road to Serfdom* (London: Routledge, 1962 [1944]).

¹¹⁹ Barlow, J. P., 'A Declaration of the Independence of Cyberspace' (Davos, 8 Feb 1996). <https://www.eff.org/cyberspace-independence>.

¹²⁰ Castro, D., 'A Declaration of the Interdependence of Cyberspace'. *Computerworld* (8 Feb 2013). <https://www.computerworld.com/article/2494710/a-declaration-of-the-interdependence-of-cyberspace.html>.

Networked publics are publics that are restructured by networked technologies. As such, they are simultaneously (1) the space constructed through networked technologies and (2) the imagined collective that emerges as a result of the intersection of people, technology, and practice.¹²¹

As the level and nature of intersections change, so too does our imagined collectivity.

How can we better understand edits to our points of intersection among people, technology and practice? A key measurement to track is differences in levels of intersection exclusivity—the exclusivity of society’s intersections of peoples, technologies and practices. Microeconomics theory has, over the years, helped establish ways of determining what is excludable and what non-excludable. Microeconomics commonly divides types of goods four ways based on (i) levels of excludability and (ii) the degree of rivalry of consumption. “Excludability” refers in microeconomics to the extent to which other people can be stopped from consuming the good. The air we breathe (a public good) and the apples on a tree in a public place (common pool resource) are, in this sense, non-excludable. “Rivalry” refers to the extent to which consumption of a good leads to diminishment of its supply. Apples on a tree are, in this sense, rivalrous, while the air we breathe is not. Fish in the ocean have changed from being non-rivalrous (there are plenty of fish in the sea) to rivalrous (fishing quotas). An easy example of a rivalrous good is oil (a private good) and an easy example of a non-rivalrous good is the music of a concert (a club good)—only those admitted to the venue can hear the music, but one person’s hearing does not diminish the amount of music left for others in the room to hear.

¹²¹ Boyd, D., ‘Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications’. Ch 2 of Papacharissi, Z. (ed.), *A Networked Self: Identity, Community, and Culture on Social Network Sites* (New York: Routledge, 2011), p. 39 (emphasis in original).

Types of economic goods

	EXCLUDABLE	NON-EXCLUDABLE
RIVALROUS	private goods	Common pool resources
NON-RIVALROUS	Club goods	Public goods

Figure 3.3 Matrix showing types of economic goods by level of rivalry and excludability

Commentators who believe the internet opens the world and its information up to everybody are, essentially, fixated on the way in which it seems to be non-excludable and non-rivalrous, and therefore a public good. There is, of course, some need for a device that gets one online, plus some bandwidth, but after that people with glaringly different income levels are able to use and enjoy what is online in similar amounts, in line with the principle of net neutrality. And when you take from the internet, there is no reduction in the amount of internet available for everybody else. These factors make the internet look and feel like something non-excludable and non-rivalrous. Indeed, the internet sometimes even seems to be one of those strange economic goods whereby the more you take, the more there is. User-generated content provides much of the “resource” of the internet, which only increases the more the internet is consumed. In this sense,

economically-speaking, the internet is a lot like education, and also a lot like a pathway through the jungle. The more people are educated, the easier it is to get educated (education does not diminish through its accumulation nor its sharing). Likewise, the more people who use a pathway through the jungle, the easier it is to find and use the pathway in future. In this way, the internet too seems to be a good that increases through consumption (rather opposite to the idea of “private goods” which has underpinned neoclassical economic theory for the past 250 years).

While there is a great deal of truth in saying that consuming the internet is like treading the jungle path (an Amazon jungle path?¹²²), much criticism has since been levied against assumptions that the internet means a free world. In *The Internet Trap*, Matthew Hindman argues that large corporations such as Google and Facebook take the profits from the attention economy leaving no room for others, implying a sort of rivalry in the accumulation and use of personal data.¹²³ The idea that the internet is a welcome place for any type of start-up or new idea—“grad students in a garage”—no longer holds; the monopolistic positions of the tech giants seem to outweigh any vestiges of regard for net neutrality and make it next to impossible for

¹²² Founder of Amazon, Jeff Bezos, originally called the company Cadabra Inc, and then wanted to change to relentless.com (which will still redirect to amazon.com when typed into a browser). The reason for settling on ‘Amazon’ was that it would help make the accompany appear first alphabetically and because, just as the Amazon river is the largest in the world, so too did Bezos want to make the largest bookstore in the world. Laliberte, M., ‘You’ll Never Guess What Amazon Was Almost Called—and Why They Had to Change the Name’. *Reader’s Digest* (2019). <https://www.rd.com/culture/amazons-name-origin/>; Byers, A., *Jeff Bezos: The Founder of amazon.com* (New York: The Rosen Publishing Group, 2007), p. 47.

¹²³ Hindman, M., *The Internet Trap: How the Digital Economy Builds Monopolies and Undermines Democracy* (Princeton: Princeton University Press, 2018).

newcomers to generate a robust digital audience for their products and services without operating in line with the tech giants' overall goals.

It is important, at this point, to note an extreme reversal of principles of non-excludability and non-rivalry through the creation of blockchain technologies. While it is not possible here to do justice to the extent to which blockchain technologies hold the potential to reverse the internet's non-diminishment through use, it is nevertheless a useful example to highlight in terms of showing the diversity of edits to our points of intersection among people, technology and practice, and therefore the diversity of edits to our networked publics coming down the line. Blockchain technologies provide distributed, open ledgers able to record data changes such that those changes can be known publicly and are historicised permanently. Because blockchain writes down the past history of data transfers in a way that is shared between so many actors that it cannot effectively be altered or manipulated, the data's history is able to act as a permanent public record. Just as the title deed made land even more of an excludable good than it was previously, so too can blockchain technologies *create scarcity online when there was no scarcity before*. Applied to transactions, the most widespread initial use has been through cryptocurrencies, but there is little to stop the use of blockchain technologies to create enduring scarcities and contracts for all kinds of virtual realities.

Take one important form of organised scarcity that is in fact necessary for representative democracy to flourish: democratic societies use legislatures guided by the simple rule that one person speaks at a time. One could imagine a way of imitating this online by creating blockchain scarcity for one's speaking time online (in the form of minutes spoken becoming a finite, coin-like commodity). Speaking time could then be given from one individual to another. A technological creation such as this would more fully release the

potential of digital technologies for democratic deliberation by restoring equivalent respect for speaking and listening.

The example of blockchain is an extreme case of inversion of the internet's presumed non-excludability and non-rivalry of consumption, showing just how far humanity can push the networked age back towards neoclassical logics of supply and demand. Nevertheless, there are key changes at play in the way humans relate online that are significantly reshaping our points of intersection and therefore the nature of our networked publics.

Seeking to identify the new institutional layout arising through the networked age, Anne Helmond writes of the 'platformization of the web', by which she means 'the rise of the platform as the dominant infrastructural and economic model of the social web and the consequences of the expansion of social media platforms into other spaces online.'¹²⁴ This is part of the ascendancy of Web 2.0, whereby websites have often come to host participatory user activity and third party app development as central to their business model. Netscape founder Marc Andreessen described the "platform" in the following way:

Definitionally, a "platform" is a system that can be reprogrammed and therefore customized by outside developers—users—and in that way, adapted to countless needs and niches that the platform's original developers could not have possibly contemplated, much less had time to accommodate.¹²⁵

¹²⁴ Helmond, A., 'The Platformization of the Web: Making Web Data Platform Ready'. *Social Media + Society*, Vol. 1, No. 2 (2015), pp. 1-11, p. 5.

¹²⁵ *Ibid*, p. 3.

In this realm of participatory, ‘multi-sided markets’,¹²⁶ the question of who is ultimately in charge rises to the fore, drawing with it deeper questions about citizenship and legitimate governance. Robert Gorwa offers three options: ‘platforms govern, platforms are governed, and platform companies are companies.’¹²⁷ By the first, ‘platforms govern’, he means the way in which the original stakeholders involved in setting-up the platform retain discretion and influence over what is shared on the platform, with the end goal of managing and curating user experience.¹²⁸ The second, ‘platforms are governed’, signifies how various stakeholders can get involved in regulating and transforming platform companies from the outside, through law and regulation, or else through social and political pressure. Platforms can also be governed internally in ways that go against the platform’s interests in monopolising its economic position. Mark Zuckerberg, for example, has made plans for a community guidelines committee to adjudicate over disputed content on Facebook (which some dub a Facebook Supreme Court¹²⁹). One of the reasons is that ‘it will provide assurance that these decisions are made in the best interests of our community and not for commercial reasons’.¹³⁰ The third option, ‘platform companies are companies’ takes an altogether different approach of accepting that rapid growth creates its own governance problems simply by virtue of the way the size and success of companies tests anew jurisdictions and regulatory frameworks. Their depth of resources gives them an edge in lobbying, which secures their status quo dominance ‘[d]espite their constant invocation of the rhetoric of

¹²⁶ Ibid, p. 2.

¹²⁷ Gorwa, R., ‘What is platform governance?’ *Information, Communication & Society*, Vol. 22, No. 6 (2019), pp. 854-871, p. 857.

¹²⁸ See also Bucher, 2018.

¹²⁹ Emmanouilidou, L., ‘Facebook wants to create a “Supreme Court” for content moderation. Will it work?’ *PRI* (5 Sep 2019).

<https://www.pri.org/stories/2019-09-05/facebook-wants-create-supreme-court-content-moderation-will-it-work>.

¹³⁰ Ibid.

disruption and innovation'.¹³¹ Julie Cohen concurs with much of this approach. She describes:

platforms more often seem to be principally concerned with establishing their own regulatory independence. Speaking at a recent network security conference, Microsoft's president crystallized that ambition, sketching a future in which platform firms function as "a trusted and neutral digital Switzerland." Several months later, chastising the NSA after a powerful hacking exploit that it had developed was stolen and then used by cybercriminals, he characterized "nation-state action and organized criminal action" as "the two most serious forms of cybersecurity threats in the world today." Statements like these, which position platforms as conscientious, neutral stewards of the global digital infrastructure, set a lofty tone that elevates the more self-interested process of strategic positioning operating continually in the background.¹³²

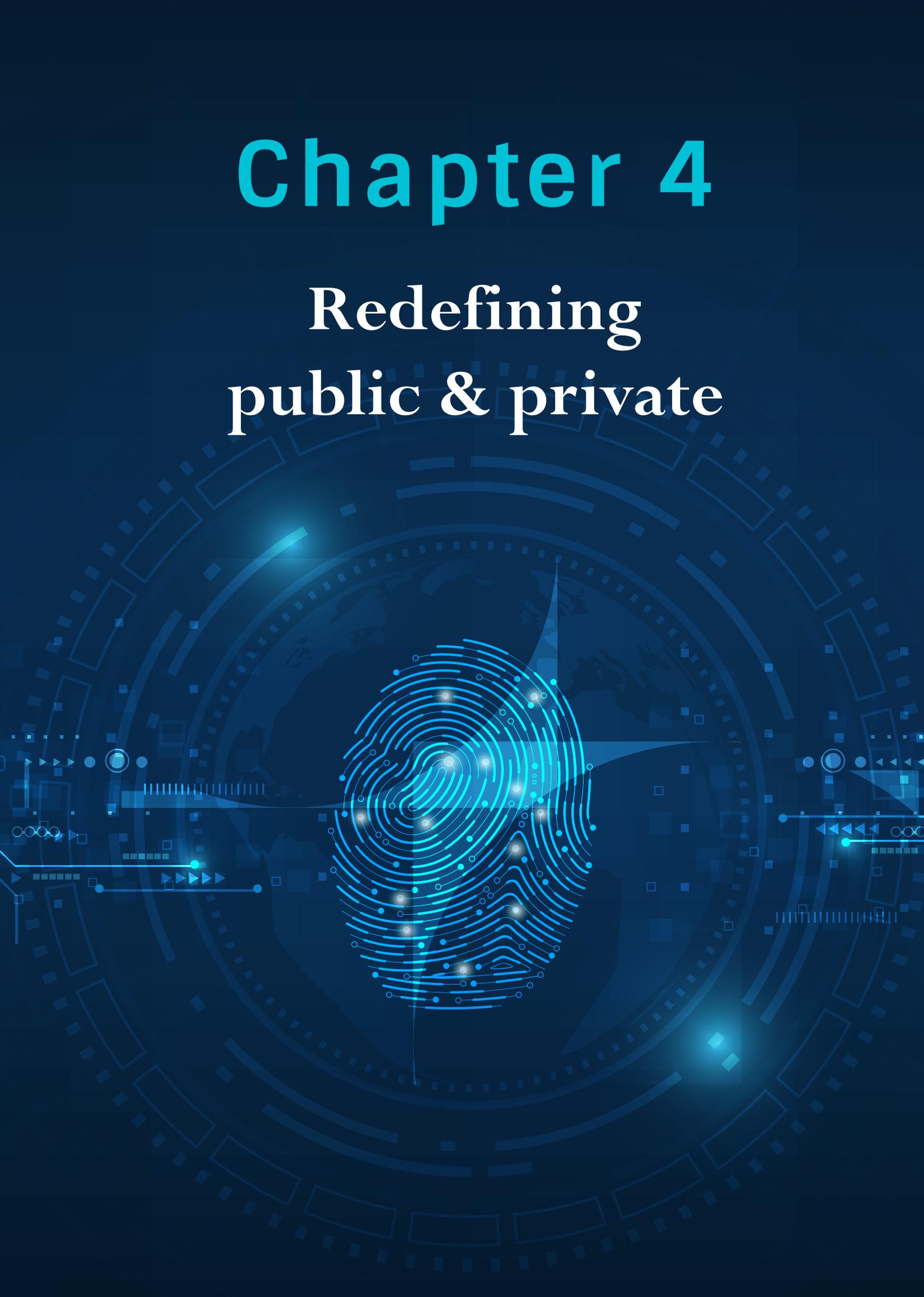
These various options of how to treat the governance elements of the new platforms, companies, communities, networks and institutions of the networked age show both that much has changed and that existing regulatory strategies have scope for application today. What is being emphasised here is the need to avoid assuming that a perfect regulatory framework can be arrived at without open engagement with the civic ideals that will shape the networked age through the connection between macro institutional developments and the needs and aspirations of the only one citizen.

¹³¹ Gorwa, 2019, p. 860.

¹³² Cohen, 2017, p. 201.

Chapter 4

Redefining public & private



Redefining Public & Private



a. Why do we have a private realm?

A problem with privacy is that, theoretically-speaking, it has no settled value. That is to say, the value of privacy is in the way it is valued by the agent, and it seems to have no value outside of that. Some people have privacy but do not value it, and others have no privacy and would not value it if they did. Privacy sometimes has unknowable value, in that we do not know what else we would lose if we lost it: we do not know all the things we cherish only insofar as they are kept private. I value my privacy because I have something that is important to keep to myself. You value your privacy because you value others thinking you might have something important to keep to yourself. Whether we have a pair of kings in our poker hand or not, we would rather keep the cards face down.

Privacy is, therefore, one of the most untrackable human goods, and yet, surprisingly, one of the most trackable legal goods. The legal right to privacy is dated to a paper written in 1890 by Samuel Warren and Louis Brandeis, writing in rebellion against the way in which the parties of Warren's wife were being mercilessly reported on in the newspapers.¹³³ One of the reasons privacy is relatively easy to track in law, while relatively difficult to track as a philosophical ideal, is that the law thrives in saying what is *not* good, *forbidden*, and *out of bounds* (and struggles with the opposite: determining what is best, balanced, or to be pursued proportionally¹³⁴). The law's natural preference for saying what we should *not* do suits the protection of privacy well because a right to privacy amounts to saying what should *not* be forced

¹³³ Geuss, R., *Public Goods, Private Goods* (Princeton: Princeton University Press, 2001), p. 105.

¹³⁴ Urbina, F. J., *A Critique of Proportionality and Balancing* (Cambridge: Cambridge University Press, 2017).

into the open, rather than giving indication as to what should be. In this way, the right to privacy can be argued as not really a right at all, but an *immunity*¹³⁵—immunity from third party interventions without consent (as with “private” property¹³⁶).

The immunity emphasis is clear in the most well-known manifestations of our right to privacy:

UNIVERSAL DECLARATION OF HUMAN RIGHTS, ARTICLE 12

No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

EUROPEAN CONVENTION ON HUMAN RIGHTS, ARTICLE 8;

UK HUMAN RIGHTS ACT 1998, ARTICLE 8

Right of respect for private and family life

1. Everyone has the right of respect for his private and family life, his home and his correspondence.
2. There shall be no interference by a public authority with the exercise of his right except such as is in accordance with the law and is necessary in a democratic society [...].

Not being ‘subjected to arbitrary interference’ and having a ‘right of respect’ with ‘no interference by a public authority’ constitute,

¹³⁵ For elucidation of Wesley Newcomb Hohfeld’s distinctions in this respect, see Finnis, J., *Natural Law & Natural Rights* (Oxford: Oxford University Press, 2011), p. 199; for complementary application to digital governance issues, see also Cohen, J. E., ‘Law for the Platform Economy’. *U.C. Davis Law Review*, Vol. 51 (2017), pp. 133-204, p. 153.

¹³⁶ For discussion of privacy assigned as a property right, see Posner, R. A., ‘The Right of Privacy’. *Georgia Law Review*, Vol. 12, No. 3 (1978), pp. 393-422, p. 399.

therefore, the details of privacy. But all this begs the question as to what its deeper purpose is.

Why have privacy? Answering, the advocacy organisation *Liberty* argues:

The right to a private life is based on the principles of human dignity and is linked to many other rights such as equal treatment and free expression. A society that does not respect personal privacy is one where dignity, autonomy and trust are dangerously undermined.¹³⁷

This is an appeal to privacy as fundamental for the realisation of other rights, in that it preserves appropriate space between the state and society, with some separation required for the formation of a social contract between the two, just as one requires an identification of distinct parties before any type of contract can be formed.

Privacy has taken on a *methodological role* in protecting and promoting representative democracy, even though the codification of privacy in law came late in the story of global democratization. For Alexis de Tocqueville, for example, focus was made on private groups and associations, and how they contribute to a lively civil society and the thriving of democratic norms, but he wrote little about the role of individual-level privacy for forming robust democracies.¹³⁸

Part of the difficulty is that if one were to determine what particular good is preserved through privacy it would ruin the distinctiveness of a “private realm”; society and the state could simply take to the task of

¹³⁷ Liberty, ‘Privacy’ (2019).

<https://www.libertyhumanrights.org.uk/human-rights/privacy>.

¹³⁸ Tocqueville, A., *Democracy in America* (London: Penguin Books, 2003 [1835]).

directly promoting that particular good, whatever it is.¹³⁹ An easy example can be seen in consumer credit ratings. One can make the argument, quite plausibly, that having individual finances (as opposed to state or community control of all finances) helps individuals take personal responsibility for their finances, and therefore helps the overall ability of companies and organisations to issue appropriate loans and financial services. The *why* to the privacy in this case is therefore quite clear—it is to make citizens responsible on an individual basis for what they choose to spend and what they choose to save. With such a clear articulation of the aim, however, it becomes easy to short-circuit the process and jump straight to directly promoting that particular good,¹⁴⁰ in this case issuing loans accurately according to the likelihood that the receiver is able to pay back over time. We therefore allow data to be collected that can give a credit score to each citizen, and believe this to ultimately be in the interests of citizens because without a credit score lenders will be less willing to issue loans. In the process, the privacy over one's financial data is eroded, but that is felt acceptable because the original purpose for the privacy was individual responsibility over one's financial affairs, something that is being accelerated through the sharing of data.

It is therefore in the interests of advocates of privacy that not too much of the value of privacy is explained, lest such short-circuiting be advocated. Defending privacy without explaining all of its value is a tactic supported by the fact that even if we did know everything about the consequences of having or not having privacy, we cannot be sure of privacy's self-determined use among agents going forward,

¹³⁹ This is the approach taken by Richard Posner, who argues that privacy can be seen as an 'intermediate' good in the sense that 'people are assumed not to desire or value privacy or prying in themselves but to use these goods as inputs into the production of income or some other broad measure of utility or welfare.' Posner, 1978, p. 394.

¹⁴⁰ Indeed, Posner's argument that we should value privacy as an intermediate good for obtaining other economic goods ends up in advocating that a right to privacy needs to be strengthened much more for corporations than for citizens. See Posner, 1978, pp. 404-6.

meaning it is impossible to fully calculate the future consequences of dismissing a right to privacy. We add to this the fact that even if we could establish the value of privacy for one person, its self-defined valuation means we cannot extrapolate this information for valuing others' privacy. Privacy is a hodgepodge of underdetermined goods, much like citizenship itself; it is a kind of determined protection of the indeterminate. As Julie Cohen comments, '[p]rivacy shelters dynamic, emergent subjectivity from the efforts of commercial and government actors to render individuals and communities fixed, transparent, and predictable. It protects the situated practices of boundary management through which the capacity for self-determination develops.'¹⁴¹

The linkage of democracy-privacy-citizenship is strong in our networked age, forming a chain of resistance against the idea that optimisation of our common goods is best achieved in a singular, public domain. Advocates of the democracy-privacy-citizenship linkage should not worry that it has become increasingly difficult to articulate its benefits in a non-partisan manner, because that inarticulacy preserves its depth as an alternative tradition, both older and newer than the high-modern rationalism that is the basis to theories of optimal decision-making through data aggregation. We fail to agree what the value of privacy is, and that is part of its value as a bulwark against having to decide.

The inarticulacy of the benefits to privacy rest on deeper incommensurability between values, for which we cannot use a single metric to determine policy optimisation.¹⁴² It is like the difficulty of saying why I prefer to make pizza dough myself, rather than having

¹⁴¹ Cohen, J. E., 'What Privacy Is For'. *Harvard Law Review*, Vol. 126 (2013), pp. 1904-1933, p. 1905.

¹⁴² Finnis, J., 'Commensuration and Public Reason'. Ch 12 of Chang, R. (ed.), *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1997).

someone else do it. There are many reasons I can give, and for each of them you can say why it would be better that someone else does it (they will do it better, it will give me a break, they enjoy it more, they will be quicker at getting our pizzas ready). I am not convinced by your counter-arguments, however, because none of them touch on my real reason for doing it, which is something about the way I enjoy self-sufficiency. That reason holds regardless of the quality of the output (how good my pizza tastes), which makes the justification of the method separate to justification of the output.

Privacy has taken on a *methodological role* in protecting and promoting democracy, touching on the essence of citizenship. It is not simply a bulwark against the rise in influence of AI, multinationals and government but a defining of the uniqueness of our human contribution.

b. What is our public realm?

To see the current role privacy is making, it can help to contrast it with our sense of what is genuinely “public”. In his book *Public Goods, Private Goods*, Raymond Geuss outlines how there were at least two distinct understandings of the public in the classical world. The first is the public as ‘a place to which everyone has free access, and thus where everything that happens can be observed by anyone’.¹⁴³ The second is the public as ‘the realm of things that concern or affect everyone’.¹⁴⁴ It is important to realise the difference between the public as *what everyone can access and see* and the public as *what everyone is affected by*, because it lies at the heart of much confusion over what should count as public in our networked age. The networked age maximises what we can access and see, but it is a mistake to think that it likewise

¹⁴³ Geuss, 2001, pp. 52-3.

¹⁴⁴ Ibid.

maximises what everyone is affected by, a mistake often made when pundits assume interconnectedness means interdependence.

When we say that everyone affected by a decision should have a say in it, is it because the decision affects a place to which everyone has access, or because the decision affects each of us as individuals? At base, we are conflicted as to whether the greater connectivity provided by the networked age increases the size of the public realm or, in fact, crowds out the public realm with a new type of individualism. Are we in a public marketplace, visible to everyone, or a black market? Is this a new interdependence or an interconnectedness shunning the common good?

The networked age brings the distinction between a public of common access and a public of common fate into central view, because the internet is on the one hand a place of private individuals and private companies involving themselves only irregularly with what is in the common interest and, on the other hand, a place of growing interdependence as we develop our products and services as a community of users.

To some extent, the openness afforded by the networked age is about openness of access, but with a push-back of attempted exclusivity in use. Everyone can login to Gmail, but no-one should be allowed, we think, to login to *my* Gmail—not even Google! Google’s generation of data for personalised advertising through word recognition of the emails that users were writing led to widespread outcry.¹⁴⁵ Ethical difficulties such as these transgress the lines between public and private. Perhaps even more importantly, they transgress lines between *types of public* by involving, on the one side, corporations that believe

¹⁴⁵ Zuboff, S., *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power* (London: Profile Books, 2019), pp. 47, 141.

in a public realm defined by free access—with email services provided free of charge—and, on the other side, users who believe that the public is what concerns everyone collectively—which their emails do not, while corporations’ ethical standards do.

There is a certain moment when common access turns into common fate, and the political instinct wakes up. We need to provide a clearer basis for determining the role and territory of that political instinct, in order to give a better sense of how civic ideals can help guide our networked age. Central to this is redefinition of public and private through more critical reflection of the relationship between interconnectivity and interdependence.

The legal scholar Dwight Newman explores the basis to what counts as groups and communities and recommends distinguishing between two types: “sets” and “collectives”. He explains:

A *set* is a collection of persons that one would identify as a different set if the included persons changed. A *collectivity* is a collection of persons such that one would still identify it as the same collectivity were some or all of the included persons to change (provided some other conditions were met) and such that the included persons properly identify themselves non-trivially as members of this collectivity.¹⁴⁶

Manchester United fans are a collectivity in the sense that individual fans are born, die, and sometimes abandon the team, and yet the collectivity of “Manchester United fans” endures. When the players “thank the fans”, they don’t specify who they are talking about and don’t make effort to exclude from thanks those fans not yet born and

¹⁴⁶ Newman, D., *Community and Collective Rights: A Theoretical Framework for Rights Held by Groups* (Oxford: Hart Publishing, 2011), p. 4 (emphasis in original).

those who have died. Liverpool FC's anthem, "You'll Never Walk Alone" likewise bypasses the idea that maybe there was a first Liverpool fan who did walk alone before others joined him or her. These groups or movements are collectivities. In contrast, the set of people with tickets to watch Manchester United play this Saturday is not an enduring collectivity. The individuals who can attend the stadium changes based on who has the tickets, so the group of "ticket-holders" is no bigger than the sum of its parts. The fan base, on the other hand, reaches in its self-definition into the past and into the future, and tolerates varied degrees of intensity among its membership while still including those members equally in outward-facing descriptions.

We need to work out whether our networked age makes us more of a collectivity. Is the networked age one of a set of app users or a collectivity of netizens?

A difficulty in answering this question lies in the way our networked community is not yet fully emergent, making us at times appear more like a set of specific users of individual services and web platforms and, at other times, more like a genuine collectivity or citizenry. We enter the networked age on *transactional and contractual terms* with digital technologies that appear to make our lives easier, bit by bit; we stay on *more permanent terms* with digital technologies that share aspects of our being and our identity, and make numerous promises about future efficiency gains and data consolidations. The latter trend forms the basis to membership of a collectivity, while the former, set membership.

If everything in the networked age were a question of set membership, it would certainly be appropriate to apply a consent framework to all our ethical dilemmas, and let individuals determine their preferences individually between each other and between themselves and corporations. But the movement from digital technologies as ad hoc

tools to digital technologies as part of our behaviour, our sense of self and our sense of community frustrates a consent framework because it detaches citizens from set membership and starts to build collectivities. Consent requires (a) individuation between users, (b) separation of users and service providers, and (c) knowledge about the changes of before and after. These conditions rarely hold in the networked age. Just as the growth of a family can't be navigated solely by contractual consent (except—as the exception that proves the rule—with entry into marriage), so too does the networked age's growth in interdependence and connectivity move us away from, rather than closer to, anything that a consent framework can offer for solving our ethical and civic dilemmas. A consent view of privacy was provided clearly by US Supreme Court Justice William O. Douglas in 1967:

Privacy involves the choice of the individual to disclose or to reveal what he believes, what he thinks, what he possesses. [...] [T]he individual should have the freedom to select for himself the time and circumstances when he will share his secrets with others and decide the extent of that sharing.¹⁴⁷

Shoshana Zuboff comments on Douglas' view by writing simply that '[s]urveillance capitalism lays claim to these decision rights.'¹⁴⁸ While the point is well made that it is much harder now than ever before to isolate and affirm individual autonomy over what one shares, it is not as simple as saying that corporations have robbed citizens of their rights. The shape of our communities has also changed, along with our level of interdependence. Zuboff's account confuses the intentionality of corporate actors with general changes in communicative practices. An impression of intentionality is racked up in her account when it serves to build a sense of moral censure, while an impression of

¹⁴⁷ Goldfarb, R., *In Confidence: When to Protect Secrecy and When to Require Disclosure* (New Haven: Yale University Press, 2009), p. 165.

¹⁴⁸ Zuboff, 2019, p. 90.

inevitability racked up when she seeks to demonstrate to the reader the scale of changes taking place.¹⁴⁹

Taking forward a consent-based approach, there is a great deal of interest in requiring consent for cookies and the storing of private data under the European Union’s General Data Protection Regulation (GDPR), which came into force in May 2018. Zuboff explains:

The regulations introduce several key new substantive and procedural features, including a requirement to notify people when personal data is breached, a high threshold for the definition of “consent” that puts limits on a company’s reliance on this tactic to approve personal data use, a prohibition on making personal information public by default, a requirement to use privacy by design when building systems, a right to erasure of data, and expanded protections against decision making authored by automated systems that imposes “consequential” effects on a person’s life.¹⁵⁰

The dilemma here is that a consent framework is being employed alongside greater recognition of its frailties and, hitherto, its abuse by corporations when they obtain consent for wide-reaching activities through terms and conditions that are essentially unknowable and

¹⁴⁹ She states, for example, that ‘[k]ey to our conversation is this fact: surveillance capitalism was invented by a specific group of human beings in a specific time and place. It is not an inherent result of digital technology, nor is it a necessary expression of information capitalism.’ On the other hand, she describes these developments as “laws of motion” that drive surveillance competition’, explaining her use of the phrase as follows: ‘I intentionally repurpose familiar language because I want to stress certain continuities in the function of an element or process. This is the case with “laws of motion,” borrowed from Newton’s laws of inertia, force, and equal and opposite reactions.’ Zuboff, 2019, pp. 85, 66. Either these are inevitable changes or chosen; the problem is common to all Marxist criticisms of economic change.

¹⁵⁰ Zuboff, 2019, p. 481.

allow a use of data in a way that a user might not be happy with if they understood its true extent. While the net effect of the regulations can be positive in encouraging companies to more clearly and directly explain how they will use private data, ultimately the consent framework will not work for the same reasons as are leading to its current refinement. GDPR works on point (c), increasing the knowledge among users about the likely changes before and after consent is given. It further ensures companies stick to what their original plans were for the use of the data, and forces them to notify users when their data is hacked. These changes help stabilise that before and after, therefore, but do not treat points (a) and (b), the difficulties of properly individuating users (e.g. proven identity and stable legal personhood) and the difficulty of maintaining separation between users and service providers (e.g. user-generated content and online community-building; see Section 3.d).

The GDPR regulations make corporations sensitive about the data they are collecting from users, and increase the chances that there will be a future request about that use of data that the corporation must then invest significantly in responding to, with possible reputational damage depending on the outcome. That stick-based approach is not the same as providing interpersonal ethics and civic ideals that can define the nature and ethical boundaries of our new shared spaces. In response to the GDPR regulations, programmer Daniel Kladnik developed the app ‘I don’t care about cookies’, which is an extension that can be added to one’s browser to make the computer accept all cookies.¹⁵¹ The app can be downloaded for free and means one’s browser automatically accepts all cookie requests, allowing the user to browse the internet freely without the never-ending pop-ups asking for consent.

¹⁵¹ Chrome Web Store, ‘I don’t care about cookies’ (29 Jun 2019).
<https://chrome.google.com/webstore/detail/i-dont-care-about-cookies/fihnjjcciajhdojfnbdddfaoknhlnja>.

Even if we were to assume that we can solve the full information dilemma when building a consent framework, the type of consent we give would still fail to capture all that we do online and who it is that is doing it (point (b), the separation of users and service providers). The GDPR regulations provide little suggestion about how to extend a consent framework to the relational use of social media. For social media platforms, it becomes impossible to properly obtain the consent of readers prior to what is posted, and companies do not make prior approval of it either. Instead, we are agreeing to community guidelines that the corporation will use as the framework for reacting to what is done on their platform. Even if one can make the legal case that I consent to Facebook's use of my personal data through agreeing to their terms and conditions, I cannot effectively consent to the offensive posts coming my way in years to come from other users. I just hope that the community guidelines lead to the kind of community I want to abide in, and look forward to the corporation reacting with sufficient clout and speed to encourage those community norms among users over time. In this sense, it is right that Facebook self-describes as a community—in essence, a collectivity—and not a set of users. We are not consenting to a certain type of service provision but agreeing to community guidelines reactively enforced over time. For the live-stream shooting of Muslims in Christchurch, New Zealand, in March 2019, Facebook users cannot be said to have consented to viewing that type of content. Instead, Facebook declared that 'anyone who breaks certain rules in broadcasting content on Facebook Live will be temporarily barred from using the service, with the possibility of a 30-day ban on a first offense.'¹⁵² That was a tightening from its previous guidelines that dictated a temporary ban would only come in response to multiple offenses.

¹⁵² Metz, C. & Satariano, A., 'Facebook Restricts Live Streaming After New Zealand Shooting'. *New York Times* (14 May 2019).
<https://www.nytimes.com/2019/05/14/technology/facebook-live-violent-content.html>.

Again, trying to recover a framework of set membership for navigating the ethical dilemmas of the networked age, there is renewed interest in age restrictions as a way of limiting negative uses of the internet by minors. The idea here is that age can provide a set membership that divides the internet into two spaces: one for those under 18 and one for those over 18, with age verification allowing one's "upgrade" from the former to the latter. If such proposals worked, however, there is no reason why there can't also be content restricted to those below 18, such as chat rooms where everyone is confident that everyone else is a minor, or sites where minors can share videos and photos without fear of access by adults seeking to groom them or the viewing of content by paedophiles. Some service providers and companies may also prefer to work in the "only under 18" world.

The UK government sought to implement an over-18 restriction by requiring that websites conduct age verification if pornography constitutes more than a third of the website's content. In the original proposals, age verification was more than simply answering an online question. The plans required uploading a photo of one's government-issued ID cards or else going to a local shop to pay to have one's ID verified, which would then provide the citizen with a 'porn pass' that could be used for adult online access.¹⁵³ Due to the near impossibility of implementing these designs, the plans were shelved.¹⁵⁴ Apart from the dilemma that legal jurisdictions do not match the boundaries of the online world, and therefore make single-country implementation almost impossible, age verification can also be criticised for avoiding broader ethical questions that the networked age throws up. There are many reasons why pornography may be wrong which are not treated by an age barrier in its consumption—those filmed or photographed

¹⁵³ Waterson, J., 'UK online pornography age block triggers privacy fears'. *The Guardian* (16 Mar 2019).

<https://www.theguardian.com/culture/2019/mar/16/uk-online-porn-age-verification-launch>.

¹⁵⁴ BBC News, 'UK's controversial "porn blocker" plan dropped' (16 Oct 2019). <https://www.bbc.co.uk/news/technology-50073102>.

may have been coerced or abused, the pornography may involve minors, the pornography may be “revenge porn” (where someone posts pictures or videos of a former partner publicly without their consent out of revenge), etc. Some take the more straightforward view that sex is beautiful *because intimate*, and so any non-private enjoyment of sex runs contrary to the beauty of human love. It is clear that if one were to hold any one of these views, age verification falls short of ensuring ethical behaviour. Both consent frameworks and age verification frameworks presuppose set membership, a type of group of decreasing relevance for understanding our online life.

The point is that we entered into the networked age on *transactional and contractual terms* with digital technologies that appeared to make our lives easier, and we stayed on *more permanent terms* as these technologies increasingly shared in our being, identity and habituated behaviour. They increasingly influence our habits and our habitat, changing our collective space from a market of transactions to a home of relationships and membership. We consented to the market but we didn’t consent to the home, and can’t anyway. As Zuboff aptly puts it, it seems ‘decision rights vanish before one even knows that there is a decision to make’.¹⁵⁵

We are each travelling this road from *transactional* to *permanent* at different speeds, with some far behind and some far ahead, though the general direction clearly one-way. This report, in setting out *an agenda for rebuilding our civic ideals for the networked age*, joins a wider discourse and literature pointing towards our emergence as a collectivity in the midst of technological advancements. What we advise is that our coming of age not be a *growth in dependency* but an *emergence of a citizenry*, guided not just by regulations but civic ideals that give effect to long-term principles for our new habits and habitat.

¹⁵⁵ Zuboff, 2019, p. 94.

Central to this is redefinition of the distinction between private and public realms, and greater certainty on the unique contribution of privacy to citizenship. No matter how networked we become, our ideals are sourced from outside the network, from the idea of ourselves when all technological interventions are said and done.¹⁵⁶

c. End-to-end encryption

In January 2019 it was revealed that Mark Zuckerberg, Facebook CEO, was pushing to integrate the messaging services of three of its core social media platforms—Facebook, Instagram and WhatsApp.¹⁵⁷ The idea was two-fold: first, to allow users to centralise their communications across platforms; and second, to make privacy settings more consistent by making end-to-end encryption a default setting in messaging services for each of the three (covering communications made in Facebook Messenger, not across the Facebook app). As Andy Greenberg explains, end-to-end encryption means that:

messages are encrypted in a way that allows only the unique recipient of a message to decrypt it, and not anyone in between. In other words, only the endpoint computers hold the

¹⁵⁶ As Zuboff writes, ‘We cannot evaluate the current trajectory of information civilization without a clear appreciation that technology is not and never can be a thing in itself, isolated from economics and society.’ Zuboff, 2019, p. 15.

¹⁵⁷ Isaac, M., ‘Zuckerberg Plans to Integrate WhatsApp, Instagram and Facebook Messenger’. *New York Times* (25 Jan 2019). <https://www.nytimes.com/2019/01/25/technology/facebook-instagram-whatsapp-messenger.html>.

cryptographic keys, and the company's server acts as an illiterate messenger, passing along messages that it can't itself decipher.¹⁵⁸

End-to-end encryption means that messenger service providers themselves do not read the message, and send it to the receiver in a way that no-one can hack its contents in transit. The message needs the receiver's key to be opened.¹⁵⁹

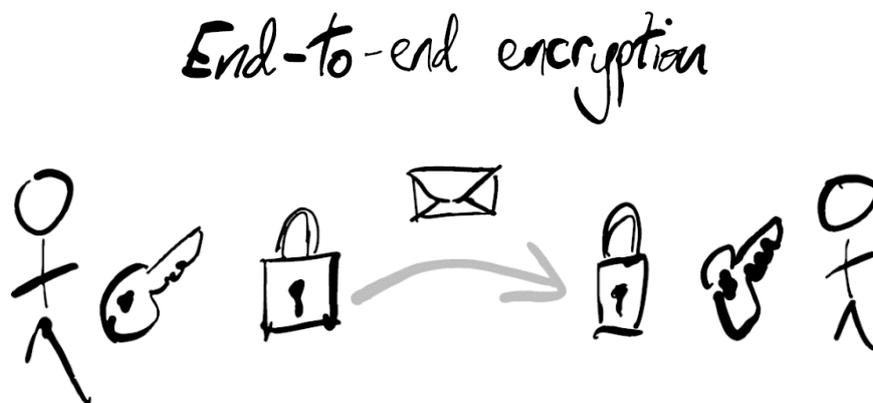


Figure 4.1 Simplification of end-to-end encryption

For Facebook, the proposed change would mean that users are able to move their communications easily between Facebook Messenger, Instagram and WhatsApp. In addition, all three platforms would commit to a default of end-to-end encryption, making messages private so that no-one, not even Facebook, can read what is written. The double change would help assuage fears among WhatsApp users that the other Facebook apps will erode WhatsApp's existing privacy provisions over the course of integration.

¹⁵⁸ Greenberg, A., 'Hacker Lexicon: What is End-to-end Encryption?' *Wired* (25 Nov 2014). <https://www.wired.com/2014/11/hacker-lexicon-end-to-end-encryption/>.

¹⁵⁹ The development of sending and receiving keys alongside messages is explained in Singh, S., *The Code Book: The Secret History of Codes & Code-breaking* (London: Fourth Estate, 1999), pp. 298-9.

It is clear that the move to end-to-end encryption can happen on all three platforms without any centralising of the messaging services, so the concern for privacy cannot be the prime motive for the proposed integration. Cross-platform messaging integration has numerous cost advantages for Facebook, and also makes it harder for governments to contemplate breaking the company up in future to encourage competition in the industry.¹⁶⁰ The added privacy element through requiring end-to-end encryption helps offset some of the likely data consolidation that will result, an especially important area of scrutiny because of Facebook's ability to form profiles of its users to help target ads. End-to-end encryption was a policy of WhatsApp from March 2016, covering individual chats, group chats, images, videos, voice messages, file attachments, calls and live location updates;¹⁶¹ it has been an important reason for WhatsApp's growing popularity over the years. WhatsApp cofounder Jan Koum grew up in communist Ukraine and had as special priorities that the app would promote privacy and freedom of speech (as well as be simple to use, with no adverts).¹⁶²

The shift to end-to-end encryption is part of a broader sea change in valuing privacy in the wake of the Cambridge Analytica scandal and other such surprises over the extent to which our online communications are captured and analysed by big data companies for ulterior motives. That is the "push" factor for greater privacy. There is, however, a "pull" factor too, which brings us closer to this chapter's mission of redefining public and private to help guide our civic ideals in our networked age.

¹⁶⁰ BBC News, 'Facebook to integrate WhatsApp, Instagram and Messenger' (25 Jan 2019).

<https://www.bbc.co.uk/news/technology-47001460>.

¹⁶¹ WhatsApp, 'WhatsApp Encryption Overview: Technical white paper' (19 Dec 2017).

<https://www.whatsapp.com/security/WhatsApp-Security-Whitepaper.pdf>.

¹⁶² Techworld, 'Best secure mobile messaging apps' (4 Jan 2019).

<https://www.techworld.com/security/best-secure-mobile-messaging-apps-3629914/>.

The “pull” factor is the *methodological role* that privacy is playing in shaping the architecture of the internet and our new forms of democracy and citizenship. Simply put, the *online-yet-private domain* is becoming increasingly important for defining the nature and extent of our citizenship in a world where what is done online has direct effect on our democratic life. The methodological role privacy plays draws in corporate giants, who seek to anticipate these changes and find continued room and use for their products and services. An end-to-end encrypted internet offers a different market for personalised internet use that combines with democratic ideals in a way that puts nonencrypted services out to pasture.

Facebook’s move to centralise its messaging services across platforms generated enormous resentment within the company, with critics observing that the independence promised to Instagram and WhatsApp at the time of their acquisitions was being taken away.¹⁶³ But what about the welfare of users? Is their data better protected under a centralised system that uses end-to-end encryption, or under siloed services, only some of which encrypt appropriately? This is in the first place a faulty question, because as mentioned there is no reason why siloed services cannot employ end-to-end encryption to equal effect. Nevertheless, as a practical matter the question is important because combined app centralisation and privacy tightening is the likely next step, in keeping with the ongoing attraction of app service consolidation so frequently suggested to improve efficiency and the marketization of data. Zuboff describes the payoff in terms of companies’ cyclical pursuit of ‘behavioral value’, described in Figure 4.2 below.¹⁶⁴

¹⁶³ Isaac, 2019.

¹⁶⁴ Zuboff, 2019, p. 70.

Zuboff's behavioural value cycle

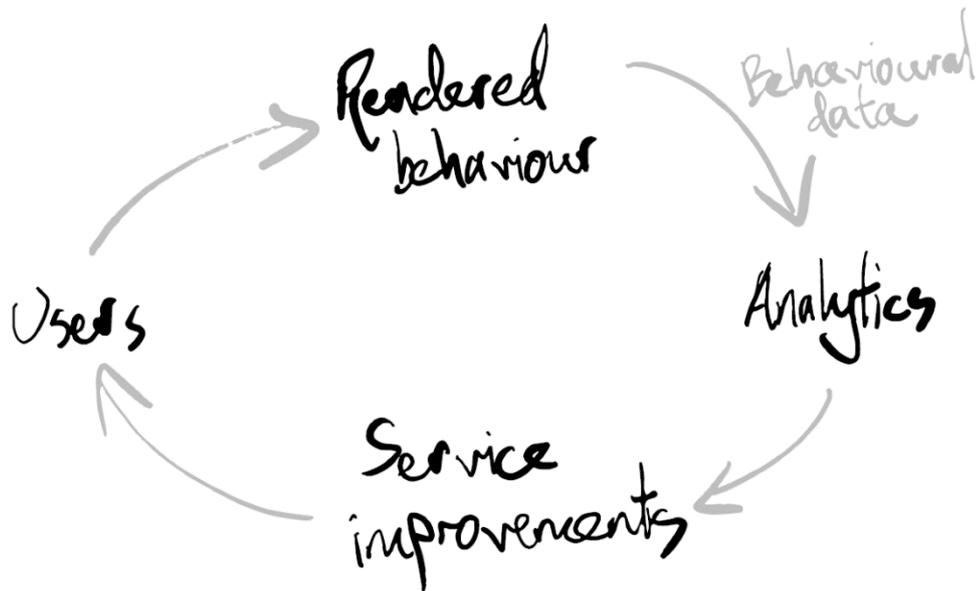


Figure 4.2 Zuboff's behavioural value cycle showing the use of personal data by corporations

Does a move towards end-to-end encryption stop the cycle of 'behavioral value reinvestment' in its tracks by making companies unable to capture users' behavioural data? There are two important reasons why this is highly unlikely. The first is that even if a platform cannot read users' messages to each other, it can still use the metadata about the use of its message services to build profiles of users. As Alan Woodward explains, 'By abstracting out and looking at who's talking to who, for how long, and when ... you can build up a very statistical picture of people very quickly. [...] In many ways, it is the context of what you say in those messages that is more important than the

messages themselves.’¹⁶⁵ The quest for understanding “data about data” formed the initial inspiration for monetary valuation of our online behaviour through the ‘Carol Brady moment’—when the search query ‘Carol Brady’s maiden name’ became the most frequent Google search every time the latest episode of *Who wants to be a millionaire?* was aired.¹⁶⁶ Monetisation is not about collecting individual pieces of private data but correlating behavioural data to form a profile of the habits, interests and dispositions of potential customers. On this front, algorithms trained through machine learning can find and isolate correlations with any type of behavioural data in order to generate pictures of user tendencies and likely user profiles, even if such data comes from the metadata of how a platform is being used, rather than specifically what is said.

The second reason end-to-end encryption is unlikely to stop the cycle of ‘behavioral value reinvestment’ is that end-to-end encryption protects the message but not the two ends. People and their computers can be hacked as usual, regardless of the security of the messages when in transit, and then the messages read as if the hacker is the original user.¹⁶⁷ If the hacker impersonates the user and continues to use the messaging service they initiate a ‘man-in-the-middle attack’ whereby they continue to send back and forth messages between the two original users and thereby avoid detection while still getting to read all messages.¹⁶⁸ Even when we are talking about

¹⁶⁵ Murphy, H., ‘How Facebook could target ads in age of encryption’. *Financial Times* (27 Mar 2019).

<https://www.ft.com/content/0181666a-4ad6-11e9-bbc9-6917dce3dc62>.

¹⁶⁶ Zuboff, 2019, p. 75.

¹⁶⁷ Srivastava, M., ‘WhatsApp voice calls used to inject Israeli spyware on phones’. *Financial Times* (13 May 2019).

<https://www.ft.com/content/4da1117e-756c-11e9-be7d-6d846537acab>;

Bershidsky, L., ‘End-to-End Encryption Isn’t as Safe as You Think’. *Bloomberg* (14 May 2019).

<https://www.bloomberg.com/opinion/articles/2019-05-14/whatsapp-hack-shows-end-to-end-encryption-is-pointless>.

¹⁶⁸ Greenberg, 2014.

corporations that do not illicitly interfere with user accounts, a consolidation of profiles across platforms can link the metadata about one's encrypted messages with actual profiles and personal data that is available to the corporations. This is where Facebook's proposed centralisation of its messaging services is extremely problematic. With 2.38 billion monthly active users of the Facebook app as of April 2019 (more than 30 percent of the world's population),¹⁶⁹ a centralisation of messaging services will mean that the metadata about one's WhatsApp use will, in a majority of cases, be able to be directly linked to one's Facebook profile, which does not employ end-to-end encryption. That produces a toxic and highly intrusive combination of encrypted and nonencrypted data gathering.

Of important concern for civil society is one advantage companies enjoy when using end-to-end encryption: it reduces corporate liability for what goes on online. A selling point of end-to-end encryption's ability to protect privacy is that the platform itself is not able to access the substance of what is being communicated (the 'illiterate messenger'). From an ethical and legal point of view, this encourages the view that the corporation is not to blame for anything sent or received. It gives the corporation a level of immunity from accusations that its service provision shares culpability for enabling acts of wrongdoing that happened through use of the platform. Such immunity is, of course, extremely attractive to tech firms struggling with the idea that they can be at fault for a great deal of user activity, even when such activity goes against their community guidelines. As discussed earlier, community guidelines are often enforced reactively, which makes them ill-equipped for generating the sort of legal or moral immunity that would allow corporations confidently to let their platforms be user-led.

¹⁶⁹ Hutchinson, A., 'Facebook Reaches 2.38 Billion Users, Beats Revenue Estimates in Latest Update'. *Social Media Today* (24 Apr 2019). <https://www.socialmediatoday.com/news/facebook-reaches-238-billion-users-beats-revenue-estimates-in-latest-upda/553403/>.

Philosophically-speaking, gaining immunity by limiting one's powers to intervene in advance of the wrongdoing is, unfortunately, on rather shaky grounds. The logic is akin to a man handcuffing himself to the banister before his wife arrives, so that he can claim he can't help with the washing-up. Will she grant him immunity due to his incapacity? Unlikely. For versions eight and up of Apple's iOS operating system, the encryption is described as 'so deep that Apple could no longer comply with government warrants asking for customer information to be extracted from devices.'¹⁷⁰ The dilemma goes at least as far back as the December 2015 mass shooting in San Bernardino, California, in which 14 people lost their lives in a terrorist attack. Authorities obtained one of the shooter's iPhones but the NSA were unable to crack the encryption. The FBI then successfully applied for a court order to mandate Apple to create and provide software to get past the phone's security features. Apple resisted and appealed the court order until the FBI eventually managed to break through the phone's security through third party support. Director of the FBI, James Comey, commented at a news conference:

What concerns me about this is companies marketing something expressly to allow people to hold themselves beyond the law. [...] The notion that someone would market a closet that could never be opened—even if it involves a case involving a child kidnapper and a court order—to me does not make any sense.¹⁷¹

For its part, Apple argued that they were not in the business of collecting users' personal information but about developing efficient devices that people trusted. It would be unprecedented for courts to

¹⁷⁰ Sanger, D. E. & Chen, B. X., 'Signaling Post-Snowden Era, New iPhone Locks Out N.S.A.' *New York Times* (26 Sep 2014).
<https://www.nytimes.com/2014/09/27/technology/iphone-locks-out-the-nsa-signaling-a-post-snowden-era-.html>.

¹⁷¹ *Ibid.*

require companies to build new software to break into citizens' accounts.

And so the race for immunity proceeds apace. As Alina Selyukh explains, 'On the newer devices, Apple says it's just not technically feasible for the company to unlock passcodes or otherwise extract data, warrant or no warrant.'¹⁷² The dilemma this creates for law enforcement agencies prompted Amber Rudd, UK Home Secretary, to call for a requirement that tech companies 'build "back doors" into their encrypted content for security services to use when they require access in the fight against terrorism.'¹⁷³ For tech companies, the general concern is that this simultaneously creates scope for the abuse of users' data and privacy by state actors. Apart from possible uses and abuses of any "back door" by those states where the tech companies are headquartered, there is great concern that compliance with national laws in less democratic jurisdictions may mean the provisions are used to violate basic human rights. Comparing 2009 and 2018 popular protests in Iran, Alp Toker, founder of the civil society group NetBlocks.org, argued that the greater use of smartphones actually made communications more regulated and less open. With many citizens using older versions of iPhones and not downloading all security updates to their devices, they are more vulnerable to state-sponsored hacking and surveillance.¹⁷⁴ When one adds the fact that smartphones are now in the habit of transmitting live location data,

¹⁷² Selyukh, A., 'A Year After San Bernardino And Apple-FBI, Where Are We On Encryption?' *National Public Radio* (3 Dec 2016). <https://www.npr.org/sections/alltechconsidered/2016/12/03/504130977/a-year-after-san-bernardino-and-apple-fbi-where-are-we-on-encryption>.

¹⁷³ Techworld, 2019. For discussion in the UK on the appropriate policies and laws surrounding data sharing among public bodies, see Law Commission, 'Data Sharing between Public Bodies: A Scoping Report', No. 351 (10 Jul 2014).

¹⁷⁴ Quinn, M., 'One Difference Between 2009 vs 2018 Iran Protests? 48 Million Smartphones'. *Voice of America* (3 Jan 2018). <https://www.voanews.com/usa/one-difference-between-2009-vs-2018-iran-protests-48-million-smartphones>.

the vulnerability of protesters to state tracking has increased in tandem with their increased capacities for rapid collective action.

The central importance of end-to-end encryption for the methodological role that privacy plays in shaping contemporary democracy was brought into full view in the 2019 Indian elections, explained below. In essence, privacy now plays a role in defining what areas of civic activity lie fully outside the state, and is therefore necessary for defining the nature of our social contract. End-to-end encryption generates added distance between citizens and the state, and so strengthens civil society's hand in negotiating the social contract. At the same time, this shift to privatised individualism in our communicative and political relations fails to make civil society more united. The trend is likely only to expand in importance as end-to-end encryption becomes more of a default provision among internet messaging services.

In 2018, a research team in the BBC set out to understand the reasons why fake news gets spread by ordinary people, and to that end conducted in-depth studies in India, Kenya and Nigeria. To get at what fake news was being shared in the form of encrypted messaging services, the researchers adopted an ethnographic approach of visiting participants in their homes, having participants guide them through their use of social media and their online social network, and obtaining access to participants' phones for a seven-day period. Overall, the researchers found that the success of fake news depended a great deal on its type of emotional appeal, observing that facts were far less important than whether stories encouraged 'the emotional desire to bolster national identity.'¹⁷⁵ The lead researcher into the India study, Santanu Chakrabarti, comments that the popularity of end-to-end

¹⁷⁵ Chakrabarti, S., 'Nationalism is a driving force behind fake news in India, research shows'. *BBC News* (12 Nov 2018a).
<https://www.bbc.com/news/world-46146877>.

encrypted messaging means fake news is spread precisely in order to help verify it:

One of the more counterintuitive behaviours we discovered was that of people sharing messages of dubious provenance or facticity within the networks, because they want someone in the network to verify the information contained in the messages. Usually, though, these messages are not shared with an explicit request to verify; but it is shared in the *expectation* that someone in the network itself would get back to the sender disproving the contents of the message if it were untrue. Usually, for every individual there is someone in the network who is trusted—either on specific issues, or because they are generally considered to be “learned” or “educated”—to point out if something is untrue. [...] As a result, fake news messages are spreading, even if the intention is very much there to check verification.¹⁷⁶

Alongside this, there is a sense in which the news, if true, is extremely important and urgent to share. As Pallavi Mishra, manager of a fact checking contractor to Facebook, commented, for users ‘being the “first” to share things in their circles gave them a rush’.¹⁷⁷ If the message is false, it seems to the user that the costs are low, so it makes sense to spread the word, just in case. In fact, the costs can be high, for example in terms of spreading negative perceptions of others, with fact checking achieving little in terms of revising the overall impression.

¹⁷⁶ Chakrabarti, S., ‘Duty, Identity, Credibility: “Fake News” and the ordinary citizen in India’, with Stengel, L. & Solanki, S. *BBC News* (2018b). <http://downloads.bbc.co.uk/mediacentre/duty-identity-credibility.pdf>, p. 44 (emphasis in original).

¹⁷⁷ Rai, S., ‘Alarming Lessons From Facebook’s Push to Stop Fake News in India’. *Bloomberg* (20 May 2019). <https://www.bloomberg.com/news/articles/2019-05-20/alarming-lessons-from-facebook-s-push-to-stop-fake-news-in-india>.

Fake news march in India



Figure 4.3 Fake news image of a march in India supposedly in support of a rapist

The above is, purportedly, a march in Mandsaur, Madhya Pradesh, in June 2018, in defence of an accused rapist. The accused, a Muslim, was alleged to have raped an eight year-old, and these Muslims protesting for his release are introduced with the caption, ‘The Quran sanctions rape of young girls if they are of a different religion’.¹⁷⁸ The post was widely shared on WhatsApp, linking to an article explaining the same. A closer look, however, reveals the gross dishonesty at play. The news website hosting the article was registered the same month as the original story broke, and a reverse image search found that the original photo was of a march in support of the victim, against the accused.¹⁷⁹ The exact opposite to what was being portrayed. For the

¹⁷⁸ Ibid, p. 53.

¹⁷⁹ Sidharth, A., ‘Viral: Fake news of march by Muslims demanding release of Mandsaur rape accused’. *Alt News* (2 Jul 2018).

placards the protesters are carrying, the text was photoshopped. The original had, ‘Won’t tolerate attacks on daughters, stop this brutality’, which was changed to ‘Release Irfan’, the accused.¹⁸⁰

Ironically, in 2012 the US State Department had admonished the Indian government for blocking websites and social media platforms that the Indian government believed fostered intercommunal tensions.¹⁸¹ Back then, the US was idealising the Arab Spring, with the internet and social media seen as resources of resistance against authoritarian regimes in the Middle East. That initial narrative of the internet as the great emancipator collapsed, of course, with the shift to fake news and computational propaganda making the internet now appear to some as democracy’s worst nightmare (see Section 5.b). The Indian government’s stance has, instead, been more consistent: the fear is that inter-religious tensions are often heightened to the point of violent conflict, and so an internet that helps polarise discourse, especially between Hindus and Muslims, will make social harmony more difficult over the long-term. One does not need to incite violence to bring about violence—it can be enough that one increases the expectation that others will be violent towards one’s community, and that leads to pre-emptive strikes.¹⁸²

As part of a growing commitment to nationalism, India’s Bharatiya Janata Party (BJP) argues that ‘Hindus should have first claim over India and that India should be a culturally Hindu nation, rather than a

<https://www.altnews.in/viral-fake-news-of-march-by-muslims-demanding-release-of-mandsaur-rape-accused/>.

¹⁸⁰ Ibid.

¹⁸¹ Garton Ash, T., *Free Speech: Ten Principles for a Connected World* (London: Atlantic Books, 2016), p. 32.

¹⁸² Burbidge, D., ‘The Security Dilemma in North Kivu, the Democratic Republic of the Congo’. *Conflict Trends*, Vol. 3 (2009), pp. 42-49.

secular state governed by a diverse range of voices.’¹⁸³ That message was taken forward in the 2019 election campaign by the BJP, but in contrast to the 2016 US elections or 2016 UK referendum on the EU, the key online platform for campaigning was not public posts on social media but end-to-end encrypted messages. It is possible to systematically restrict fake news on public platforms, though extremely laborious and regularly ineffective in practice. Saritha Rai explains the process in a typical day for those contracted by Facebook in India:

Hundreds of texts, audio and video clips are flagged daily to Vishvas’ [a news company’s] fact-checkers by Facebook’s automated tool, or through dedicated email or WhatsApp helplines. At 9:30 each morning, the team chooses two dozen stories to debunk from among hundreds mostly related to the elections, political parties or prominent leaders. Verifications begin with reaching out to the subject or their PR representative. A variety of tools like image recognition and sun-shadow calculators are employed. By the day’s end, about 20 pieces of fake content are debunked in detailed articles on Vishvas’ website. Facebook drastically reduces their distribution and they are fed to its machine-learning algorithms.¹⁸⁴

Difficult though this is, it is a possible strategy for dealing with fake news. The problem is that little of the method can apply to regulating the content of end-to-end encrypted messages. Rohit Chopra argues that the decentralised nature of social media sites such as WhatsApp better facilitate hateful messages than more centralised or public networks because they make accountability and responsibility harder to establish, achieving separation yet coordination between official and

¹⁸³ Chopra, R., ‘In India, WhatsApp is a weapon of antisocial hatred’. *The Conversation* (23 Apr 2019a). <https://theconversation.com/in-india-whatsapp-is-a-weapon-of-antisocial-hatred-115673>.

¹⁸⁴ Rai, 2019.

nonofficial groups.¹⁸⁵ Privacy gives autonomy to civil society actors, an autonomy protected even from the community guidelines of app developers.

WhatsApp tried to respond to the use of end-to-end encryption for spreading fake news and hate speech in India by restricting some of the app's functions for its 230 million Indian users.¹⁸⁶ WhatsApp groups were already restricted to 256 members each, and so the company further placed a limit in India on message forwarding, such that any given message could only be forwarded five times to other users (compared to a limit of 20 times in other countries).¹⁸⁷ It is worth remembering that originally a forwarded message did not even get described as one by WhatsApp. But as part of deterring the use of WhatsApp for copying and pasting content, the company since makes clear which messages have been forwarded and which have been written personally, in order to preserve the platform's *raison d'être* of personalised messaging. Overall, these efforts are blunt tools, however, as they restrict the sharing of truth to the exact same degree as the sharing of lies. And, in any case, there are numerous methods for overcoming the restrictions, such as apps that imitate WhatsApp but remove the restrictions, add-on software that automates the dissemination of messages through many user profiles, and the fact that one can always get political workers and volunteers to use their accounts for the party's aims.¹⁸⁸

¹⁸⁵ Chopra, R., *The Virtual Hindu Rashtra: Saffron Nationalism and New Media* (Noida: HarperCollins India, 2019b).

¹⁸⁶ Varma, T., 'The WhatsApp election: The fake news challenge to democracy'. *European Council on Foreign Relations* (30 Apr 2019). https://www.ecfr.eu/article/commentary_the_whatsapp_election_the_fake_news_challenge_to_democracy.

¹⁸⁷ Kastrenakes, J., 'WhatsApp limits message forwarding in fight against misinformation'. *The Verge* (21 Jan 2019). <https://www.theverge.com/2019/1/21/18191455/whatsapp-forwarding-limit-five-messages-misinformation-battle>.

¹⁸⁸ Vengattil, M., Kalra, A. & Phartiyal, S., 'In India election, a \$14 software tool helps overcome WhatsApp controls'. *Reuters* (15 May 2019). <https://www.reuters.com/article/india-election-socialmedia->

All this proves that there is nothing per se guaranteed by calls for greater privacy without a sense of what the wider purpose of privacy is. Ultimately, if there is no connection between privacy and one's search for truth, the private domain becomes a place of equal protection for falsehood and honesty, without any clear civic contribution. That is not to say that the state's view on what is true should inform on the right to privacy. But it is to say that the methodological role privacy plays in demarcating the distinction between citizen and state means it must itself be oriented towards discovery, deliberation and authenticity if it is to play a robust role as an ideal of citizenship in our networked age. We need to redefine the private realm to connect with that deeper role—a redefinition that can then be systematically taken forward for methodologically reshaping contemporary democracy.

d. Privacy as conscience & space for personal moral development

Privacy has many uses and protects many things, and yet over the years it has been increasingly threatened. Its unique contribution nevertheless persists in the minds of many commentators of democracy. James Scott remarks:

...private spheres have been made the object of official intervention. Much of the work of Michel Foucault was an attempt to map these incursions into health, sexuality, mental illness, vagrancy, or sanitation and the strategies behind them. Nevertheless, the idea of a private realm has served to limit the ambitions of many high modernists, through either their own

whatsapp/in-india-election-a-14-software-tool-helps-overcome-whatsapp-controls-idUSKCN1SL0PZ.

political values or their healthy respect for the political storm that such incursions would provoke.¹⁸⁹

That limiting of ‘the ambitions of many high modernists’ is the enduring role that privacy seems able to play: inciting anger, fear or consternation when interventions into our private lives go too far. Practically speaking, though, that limiting, reactionary role is mitigated when advances into the private domain are made subtly and incrementally. The networked age has come hand-in-hand with a surveillance that is hard to detect and market-based,¹⁹⁰ and that makes it difficult for privacy to play a limiting role without reference to its wider positive purpose.

What is that positive purpose? Raymond Geuss points to wide-ranging examples of the way privacy has been valued over the ages and concludes that such heterogeneity means one cannot say the concept has any universal meaning or value.¹⁹¹ Our term “public” is derived rather straightforwardly from the Latin “publicus”, but there is nothing like the same connection for “privacy” (the closest being “privatus”, meaning non-office holder,¹⁹² or else “secretum”). What is private has often simply meant what is shameful when done publicly, which makes it relatively easy for Geuss to demonstrate how its use has been under constant dispute and renegotiation. He argues that what counts as public or private is really a matter of what society thinks of sufficiently common concern for public decision-making, determined in part by our current needs and challenges. Geuss summarises:

¹⁸⁹ Scott, J. C., *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* (New Haven: Yale University Press, 1998), p. 101.

¹⁹⁰ Zuboff, 2019.

¹⁹¹ Geuss, 2001.

¹⁹² *Ibid*, p. 71.

The distinction between public and private, then, is a relative one in a number of distinct senses: first, it depends on the level of social knowledge of the possible consequences of acts. Second, it depends on what we allow or disallow to count as “consequences” of action. Third, it also depends on the value judgments of the members of the society and thus on their views and decisions about what consequences “need” to be controlled. Finally, it depends on an initial decision about who is considered to be “directly concerned.”¹⁹³

These areas of indeterminacy notwithstanding, Geuss makes interesting reference to St Augustine’s autobiographical work *Confessions*, written between AD 397 and 400, as particularly insightful for giving a positive view of the role privacy can play in someone’s personal development. Dissatisfied with Manicheism—an ancient religion teaching that there were two opposite powers of light and darkness in constant struggle—and feeling social pressures on all sides, Augustine decided to withdraw to an isolated villa in Cassiciacum, in modern-day Italy, with a few close friends for intense philosophical and religious discussion. Augustine’s account reveals the extent to which this period became for him a quest for self-knowledge. The effort, overall, is to build-up the interior dialogue between one’s soul and God, to better know what is the right thing to do with one’s life. Augustine’s example profoundly shapes Western Europe’s conviction over the distinctive value of privacy for moral self-development, as compared, say, with sub-Saharan African thought on moral development through realising the fundamental interdependence of human relations.¹⁹⁴ Geuss explains Augustine’s particular position:

¹⁹³ Ibid, p. 85.

¹⁹⁴ As John Murungi explains, ‘African jurisprudence takes human beings in their social setting. Contrary to the claims of Western modernity, such a setting is not a social construction. Human beings are not social beings because they socialize with one another. They socialize with one another because they are social beings. The claim

The “interior” phenomena Augustine is analysing are ones we have come to describe with the terminology of *privacy*. [...] [W]e would say that our states of mind, both cognitive and desirous, are “private”—we have access to them that no other human being has. [...] My mental states are supposed to be epistemically *inherently* private (and they are then construed as ontologically private). [...] [I]n Augustine’s case, “the private” is in its origins a domain to which he wishes to withdraw when social demands become too oppressive, the villa at Cassiciacum, and eventually an ontologically privileged place of withdrawal within his own mind.¹⁹⁵

As discussed in Section 3.c, the networked age makes attention a scarce resource—both in terms of our attention span and in terms of the usefulness of our attention for shaping consumption trends. What Augustine shows us is an example of taking ownership of one’s capacity for attention through retreat into the private sphere (built through a kind of presentism: memory of what has passed, awareness of the present, expectation of the future). Machine learning, algorithmic decision-making and surveillance capitalism all specialise in rapidity of response, but the private realm gives self-direction. Julie Cohen labels this as having space for ‘emergent subjectivity’ and writes that it ‘can evolve in ways that produce a robust sense of

that human beings are social beings is not to be taken as denying individuality. That is, it is not to be taken as denying individual rights or individual autonomy. Individual rights or individual autonomy in a social vacuum are theoretical constructions that are removed from human reality. In Africa, legality is not a social construction. It is a natural social setting in which the African has his or her being.’
Murungi, J., ‘The Question of an African Jurisprudence: Some Hermeneutic Reflections’. Ch 43 of Wiredu, K. (ed.), *A Companion to African Philosophy* (Malden: Blackwell Publishing, 2004), p. 523. See also Burbidge, D., ‘Connecting African Jurisprudence to Universal Jurisprudence through a Shared Understanding of Contract’. Ch 5 of Onazi, O. (ed.), *African Legal Theory and Contemporary Problems: Critical Essays* (Dordrecht: Springer, 2014).

¹⁹⁵ Geuss, 2001, pp. 71-2 (emphasis in original).

agency, supportive and resilient networks of relational ties, and critical independence of mind'.¹⁹⁶ That is, for her, the deep value behind privacy and something threatened in the networked age. Her thesis on the effects of its erosion is the following:

If [...] the capacity for critical subjectivity shrinks in conditions of diminished privacy, what happens to the capacity for democratic self-government? Conditions of diminished privacy shrink the latter capacity as well, because they impair the practice of citizenship. But a liberal democratic society cannot sustain itself without citizens who possess the capacity for democratic self-government. A society that permits the unchecked ascendancy of surveillance infrastructures cannot hope to remain a liberal democracy.¹⁹⁷

The good of privacy lies in its ability to help self-initiate the search for truth, spoiling any supposed neutrality between false and honest information. Privacy supports discovery, deliberation and authenticity in an age of triviality and scattered moral extension. Defined on these terms it forms a basic civic ideal for our networked age.

¹⁹⁶ Cohen, 2013, p. 1910.

¹⁹⁷ Ibid, p. 1912.

Chapter 5

Algorithmic vs democratic decision-making

0.78



Algorithmic vs Democratic Decision-Making



a. Defining algorithmic decision-making

In the book *The Master Algorithm*, Pedro Domingos describes an algorithm in simplified form as ‘a sequence of instructions telling a computer what to do.’¹⁹⁸ Anyone who uses an internet search engine is harnessing algorithmic decision-making; the software uses algorithms to decide what hits to display, and in what order to display them. The first search engines, up to the end of the twentieth century, matched basic keywords to try to find what one was looking for. Generally, the search produced either nothing (perhaps because the keyword was misspelt) or vast numbers of hits, leaving one to scour the whole list in the hope of finding what was wanted. We can call that the first phase of searching. Algorithms then began to be developed which gave different scores to different sites, and ranked them in order of probability that they were what one was looking for. This was based on the frequency of links to those sites, so that different users would be given the same results for the same search. We can call that a second phase, in which the big names were Yahoo and Google. But then search engines were personalised. They would collect data about a given user, what kinds of sites they visited, which results of previous searches they had clicked on, and so on. Google did this better than Yahoo, and increasingly became the search engine of choice. So far so good: Google was successful because time and again it put top of the list what users were actually looking for. We can call the personalisation of search engines a third phase. The search engine was using what it learned about the person inputting the search terms to enhance the experience.

¹⁹⁸ Brogan, J., ‘What’s the Deal With Algorithms? Your 101 guide to the computer codes that are shaping the ways we live’. *Slate* (2 Feb 2016). <https://slate.com/technology/2016/02/whats-the-deal-with-algorithms.html>.

Early in the 21st century Google realised they could use what they had learned about the user for another purpose.¹⁹⁹ We can call this a fourth phase. If they understood an individual's preferences, then they could sell that knowledge to advertisers, who could target their advertising more accurately. Actually, the data that Google had collected about an individual was much too valuable a resource to sell on; it would be more profitable to sell services through the Google search engine. The services would be measured, and charged, by the click-through rate, which is the frequency with which users click on items that are presented. This is not quite as robust a measure as knowing what that individual subsequently purchased, but it provides a better indication than is readily available from newspaper advertising or billboards. It thus restored the feedback which was provided by coupons, originally described by Claude Hopkins in 1923,²⁰⁰ and which was lost with the advent of television advertising. The aim of advertising is persuasion;²⁰¹ the aim of targeted advertising is targeted persuasion.

One might think that we have now exhausted the opportunities available for developing online advertising through algorithms: the machine learns the user's preferences, and thereby maximises that individual's click-through rate. But we are now moving to a fifth phase. Stuart Russell, who wrote one of the definitive textbooks on AI,²⁰² put it like this, 'The solution is simply to present items that the user likes to click on, right? Wrong. The solution is to change the user's preferences so that they become more predictable.'²⁰³ This principle is not restricted to advertising. It can be applied to voting preferences too. How can a machine alter the preferences of an

¹⁹⁹ Zuboff, S., *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power* (London: Profile Books, 2019), p. 74.

²⁰⁰ Hopkins, C. C., *Scientific Advertising* (London: 1968 [1923]).

²⁰¹ Packard, V., *The Hidden Persuaders* (London: Longmas, Green & Co., 1957).

²⁰² Russell, S. J. & Norvig, P., *Artificial Intelligence: A Modern Approach* (Harlow: Pearson, 2016), 3rd Ed.

²⁰³ Russell, S. J., *Human Compatible: Artificial Intelligence and the Problem of Control* (United Kingdom: Allen Lane, 2019), p. 8.

intelligent, thoughtful citizen? Very easily. Even sticking to the truth, the machine can select the stories and the statistics to push a citizen in the direction it wants their thought to go, as well as target them to go out and vote or not.²⁰⁴ Newspapers have been doing it for generations, targeted on the preferences that led the reader to choose that newspaper. The difference is the degree of personalisation that algorithmic targeting is achieving.

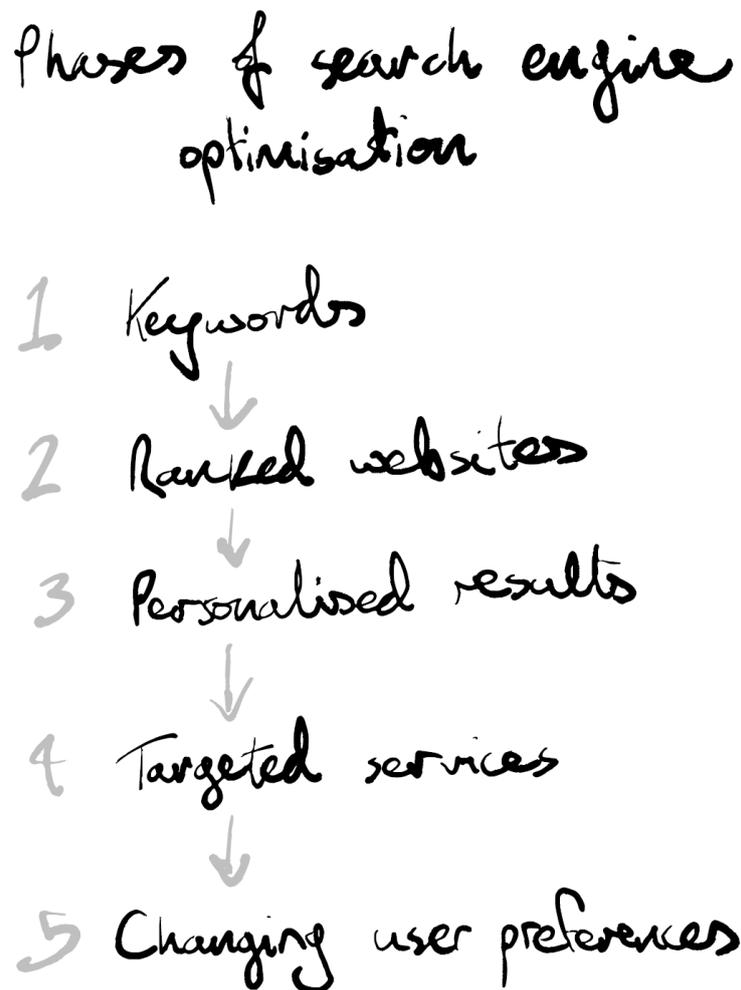


Figure 5.1 Five phases of change for search optimisation

²⁰⁴ Madsen, J. K., *The Psychology of Micro-Targeted Election Campaigns* (Basingstoke: Palgrave Macmillan, 2019).

Need it be all bad? Not if responsible citizens can develop what Solomon prayed for, the skill to listen with discernment.²⁰⁵ Harold Macmillan used to tell a story about the Oxford Professor of Moral Philosophy, J. A. Smith, when asked about the purpose of education. His reply was set in the idiom and culture of 1914. ‘Gentlemen’, he argued, ‘You are now about to embark upon a course of studies Together, they form a noble adventure. But I would like to remind you of an important point. Some of you, when you go down from the University, will go into the Church, or to the Bar, or to the House of Commons, or to the Home Civil Service, or the Indian or Colonial service, or into various professions. Some may go into the Army, some into industry and commerce, some may become country gentlemen. A few—I hope a very few—will become teachers or dons. Let me make this clear to you. Except for those in the last category nothing that you will learn in the course of your studies will be of the slightest possible use to you in after life—save only this—that if you work hard and intelligently you should be able to detect when a man is talking rot, and that, in my view, is the main, if not the sole, purpose of education.’²⁰⁶ In the networked age, citizens need to develop the critical judgment to know when what the machine is feeding them is rot.

Algorithmic processes are about instructions, identification of variables, and execution. Algorithmic decision-making was defined in Section 1.c as *optimised responses following pattern identification*. The challenge throughout this chapter, therefore, is to interpret the similarities and differences between that form of decision-making and democratic decision-making, a growing question of concern for the future of citizenship in democratic societies. Julie Cohen expertly explains the tension at play:

²⁰⁵ 1 Kings 3:9.

²⁰⁶ John Alexander Smith, Speech to Oxford University students, 1914.

Privacy scholars and philosophers of technology have begun to question whether information processing practices that subject individuals to predictive and effectively preemptive judgments impair due process guarantees. In the era of Big Data, the most individualized judgments are not necessarily the most dignifying. Due process in the era of comprehensive, preemptive computation may entail limits on fine-grained personalization in a range of public administrative processes. While it might seem tempting, for example, to calibrate disability benefits based on the precise level of need, or to engage in real-time monitoring of Medicaid recipients' food purchases to supervise nutritional choices, a liberal democratic society cannot simply deploy surveillance technologies to close the gap unfilled and unfillable by perfect technologies of justice.²⁰⁷

We have here a tension between, on the one hand, what is felt to be optimal decision-making according to principles of machine efficiency, and on the other, what is felt important for citizens to decide in line with principles of justice and democracy. Simply reflecting on outcomes of the choices and decisions being made does not end the debate,²⁰⁸ because there seems something perennially relevant about citizen participation in the decisions that most affect their lives. This chapter argues that democratic decision-making is, in fact, better than even the most efficient algorithmic decision-making

²⁰⁷ Cohen, J. E., 'What Privacy Is For'. *Harvard Law Review*, Vol. 126 (2013), pp. 1904-1933, p. 1931.

²⁰⁸ As James Scott remarks, with respect to government planning centred only on means-ends calculations: 'There may, of course, be no alternative to planning, especially when the urgency of a single goal, such as winning a war, seems to require the subordination of every other goal. The immanent logic of such an exercise, however, implies a degree of certainty about the future, about means-ends calculations, and about the meaning of human welfare that is truly heroic. That such plans have often had to be adjusted or abandoned is an indication of just how heroic are the assumptions behind them.' Scott, J. C., *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* (New Haven: Yale University Press, 1998), p. 95.

whenever and wherever the ultimate goal of the decision-making process is interpersonal unity.

b. Defining democratic decision-making

It seems the rise of the networked age is not healing social divisions, but increasing them. This appearance is ironic, because the networked age promises unprecedented togetherness. We are being tied ever more closely, and in that binding we seem to be increasingly falling apart. Why is this so?

The networked age does not only change our social ties but also our *practice of decision-making*. Society and politics are inextricably linked, such that when the shape of society changes, so too do the demands placed on public decision-making. Just as there has at times been naïve optimism about the benefits of greater social connectivity, so has there been naivety about the way in which the internet is able to help decision-making simply by making information more available. The initial enthusiasm is understandable, however, as it connected with some of the oldest justifications for democratic decision-making made by the Marquis de Condorcet. In the 18th century he famously advanced the “jury theorem” that if each person is more likely than not to make a correct decision, the probability that a group votes on the correct decision together increases as the number of people increases. With this justification of the democratic method, it seems natural that the availability of more information and more people that the networked age affords will only increase the likelihood of good decision-making.

Instead, we are confronted with a dystopia whereby the networked age seems to be facilitating a rise in disinformation, fake news, emotivism and post-truth, the latter occurring when ‘objective facts

are less influential in shaping public opinion than appeals to emotion and personal belief'.²⁰⁹ As Shannon Vallor puts it, 'the relevant facts about many topics (quantum physics, global climate, evolution, high-frequency trading) are not equally accessible from all perspectives, so if adding more perspectives in these cases merely dilutes the views of those few who have reliable knowledge, the result will not be truth but confusion.'²¹⁰

How do citizens decide how to vote when given the opportunity? Leaving aside for the moment the use of referenda to allow citizens to express a preference on a specific issue, consider how they decide how to vote for who will govern them. Voter A may have a deep concern for public services, and she may seek to identify the party (if the choices are between well-defined parties) which will best promote education, health, law and order, etc. She may also have a concern for the poor, and favour redistributive taxation. Voter B may have different priorities, and seek simply to vote for the party which in his judgment will leave him best off. Other factors may come into play, such as the perceived trustworthiness of an individual candidate, or their ability to evoke empathy from fellow citizens. Not all of these preferences are mutually exclusive, though some of them may conflict.

How good are humans at navigating these choices, most of which involve multiple objectives? Does democracy deliver the government that people want, or at least that they deserve? To varying extents, yes, and few would advocate the abolition of democracy. Could machines help us to do better? Methods are being developed for multi-objective optimisation. Suppose one is developing the next generation

²⁰⁹ 'post-truth' in *English Oxford Living Dictionaries* (Oxford University Press, 2019). <https://en.oxforddictionaries.com/definition/post-truth>.

²¹⁰ Vallor, S., *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (New York: Oxford University Press, 2016), p. 179.

of semiconductor chips. They are required to be as cheap as possible, as reliable as possible, as small as possible, and to dissipate as little heat per unit area as possible, etc. Such situations of competing objectives are common in life. Suppose a machine was better than at analysing which candidate, if elected, would be more likely to deliver one's optimal combination of preferences. Might there be something to be said for benefitting from that guidance?

Our foundational views of democratic engagement rested, in part, on Condorcet's wisdom, whereby the greater exchange of views was to make our democratic engagement more refined and accurate, a view echoed by John Stuart Mill.²¹¹ Revelations of hate-filled misinformation destabilise our foundations and our older understanding of the value of democracy: the idea that by getting together as equals we have a better chance of solving the problems we hold in common. Similarly, if machine learning is better at multi-objective optimisation, why should it not hold sway for political affairs too? Jessica Baldwin-Philippi comments:

Deliberative democrats have viewed the wealth of information to which the Internet is home as a way to inform citizens so that they are more capable of informed discussion, debate, and opinion formation, while others argue that an abundance of highly specialized content can fragment people and insulate them alongside fellow-minded citizens and opinions. Or worse, the technology itself, via algorithms, does the job of filtering for citizens. Yet another area of disagreement among scholars is whether the ability to form online communities both productively returns us to our communitarian roots and isolates us from one another and our civic duties.²¹²

²¹¹ Mill, J. S., *On Liberty* (New York: Bedford, 2008 [1859]).

²¹² Baldwin-Philippi, J., *Using Technology, Building Democracy: Digital Campaigning and the Construction of Citizenship* (Oxford: Oxford University Press, 2015), pp. 10-11.

How do we push past this collapse in confidence in deliberative democracy? One option is to simply come to terms with the death of democratic society and admit that authoritarian rule is, ultimately, necessary when making difficult, complicated or long-term decisions. This could be because there are some questions that are simply too convoluted for humans to amass all the relevant information on, or because individual articulations of views will contain too much self-interest to point towards a good overall decision.

That latter point was admitted in Condorcet's "paradox", a model showing that voting according to one's preferences can spiral into cycles of collective irresolvability. Condorcet's concern was further developed in the 20th century by Kenneth Arrow, who demonstrated the "impossibility theory", that tactical voting can always upset what would be the optimal outcome of a vote according to true preferences, so long as there are at least three options and at least three voters choosing between them.

The frustration that there are some decisions too complicated for humans to master either in terms of amassing all the relevant information or coordinating private interests leads many to suppose that AI would better handle them. From a political theory point of view, it follows the same logic as justifications for authoritarianism, only that rather than saying the authoritarian or dictator has *greater* capacity, the futurists hint that AI has *greatest* capacity. On these terms, there is some critical overlap between justifications for one-party rule and the singularity thesis (the theory that at a certain point technological developments will become irreversible and lead to computerised super-intelligence). Both one-party rule and the singularity thesis take the institution of greatest capacity to be the most likely to deliver good decisions, differing only in terms of what institution that will likely be. Shoshana Zuboff sees the two as coming together, describing Chinese ambitions for the use of digital

technologies (for example, as seen in its social credit system “Sesame Credit”), in the following manner:

The state will assume the role of the behaviorist god, owning the shadow text and determining the schedule of reinforcements and the behavioral routines that it will shape. Freedom will be forfeit to knowledge, but it will be the state’s knowledge that it exercises, not for the sake of revenue but for the sake of its own perpetuation.²¹³

In fear of such a rise in techno-political authoritarianism, some advocate retreating from technological progress altogether. Others insist that democracy is our defence against it because of democracy’s separation of powers and overt avoidance of centralised leadership, a move to saying democracy is the least bad way of ruling because it stops abusive domination of minorities by the majority.²¹⁴ Such responses are weak in the sense that they are not able to establish any inherent value in civic decision-making that identifies it as better than data-driven algorithmic decision-making. They value democracy for what it was understood as able to achieve before the networked age came, and fail to give a persuasive account as to why the people at the frontier of the networked age’s developments—either in Silicon Valley or in the highest offices of authoritarian regimes—should, as individuals, themselves prefer to embrace democratic forms of engagement when trying to optimise decision-making.

The problem with the Condorcet approach is that its simplistic account of human nature—the assumption that we are, on average, each more likely than not to make a correct decision—puts the explanatory burden on the relationship between people and the

²¹³ Zuboff, 2019, p. 394.

²¹⁴ Dworkin, R., *Taking Rights Seriously* (London: Bloomsbury, 2013 [1977]).

information they have, rather than the nature of the decision. It makes the twin assumptions that an objectively correct decision exists out there and that impartial information about the relevant issue is the main means to arriving at that correct decision. Sometimes these assumptions do hold, like when guessing the number of sweets in a jar, but often they do not. In terms of the expressly political—that normative hierarchy we place on our pursuit of public goods—they do not hold. Take, for example, a decision to limit immigration, which of course involves the facts of numbers of people and their likely economic effects, but also involves asking what kind of society we want to become. The view that political questions are reducible to establishing facts and causal mechanisms is a view filled with blind spots, because it omits that normative dimension. While political questions do involve assessments of facts and causes, they also involve a self-reflective social coordination that is agency-driven and imbued, whether explicitly or implicitly, with value judgments.

For the most part, we tend to leave the task of ascertaining facts to experts, and that is a persistent norm in democracies, making Condorcet’s jury theorem misrepresentative of the nature of democratic civic engagement. Indeed, the growth of fake news is precisely parasitical on our trust in experts (which fake news tries to exploit) rather than any trust we have or had in collectivised fact-finding and aggregative voting. Fake news does not pretend to be the wisdom of the crowds (the average view), but secret, shocking revelations *that only a few are privy to*. The battle between fake news and true news is about determining the trustworthy authority to provide specialist knowledge and insight, a necessary but insufficient tool for democratic engagement, and not the main reason we choose to do things democratically.

We democratically engage when we suffer deep disagreement over the values that give the very definition of what counts as “optimal”. It is not, therefore, a process of optimisation but a process of navigating

distinct optimisations. Things get political when it is hard to stack up the values at play in a given choice, that is, when there is understandable meaning and purpose behind each of the options, and those meanings and purposes cannot be reduced to a single metric of evaluation (the values are incommensurable). Take, for example, the following dilemma:

Deloitte acknowledges that according to its own survey data, most consumers reject telematics on the basis of privacy concerns and mistrust companies that want to monitor their behavior. This reluctance can be overcome, the consultants advise, by offering cost savings “significant enough” that people are willing “to make the [privacy] trade-off,” in spite of “lingering concerns....” If price inducements don’t work, insurers are counseled to present behavioral monitoring as “fun,” “interactive,” “competitive,” and “gratifying,” rewarding drivers for improvements on their past record and “relative to the broader policy holder pool.” In this approach, known as “gamification,” drivers can be engaged to participate in “performance based contests” and “incentive based challenges.”²¹⁵

We have here some sort of dilemma between, on the one hand, optimisation of the service being provided and, on the other, the privacy of the customer. Deloitte is, unfortunately, trying to emasculate the latter concern by either raising the importance of the former or, else, hiding the issues at stake through gamification. The approach—increasingly prevalent in the networked age—of framing possible secondary effects (in this case, making the product more interactive) as instead a primary reason why it is being done (which is, rather, so that the company gains behavioural data for its product development) has been an element of marketization that is a frequent

²¹⁵ Zuboff, 2019, p. 216.

subject of concern among political theorists and not a phenomenon limited to the networked age. The general idea is that liberal economics leads to justifications of what would be comfortable for the customer, rather than moral for the citizen. An incremental blurring of the perceived relationship between individual actions and their effects on others diminishes civic virtue.²¹⁶

Incommensurable values lie all around us, and yet the networked age is trying to optimise something. What is it? Clickability? Attention? Consumption? Surveillance? Social connectivity? Information? Social welfare? Profit? All of these things at different times and at the same time. But the goal of any particular software depends ultimately on its original design. It is, of course, possible that a program ends up optimising something it was not originally designed to by the human authors, but it is still optimising something, albeit accidentally. An example may help.²¹⁷ Let's say engineers seek to use machine learning to help identify diabetes from diabetic retinopathy images. The machine is trained to work out what kinds of images likely indicate diabetes is present. If the original data it is used to train on is biased then the optimisation procedure will optimise something the designers did not intend. If none of the images from younger patients have diabetes, for instance, and all of the images from older patients do, the machine learning algorithm may end up estimating age rather than the likelihood of diabetes. This is a mistake, but the fact of it does not prove that the machine has a life of its own or a mind of its own. It is still optimising in line with how it was designed, only that the original design did not act as intended with the training data available.

²¹⁶ Milbank, J. & Pabst, A., *The Politics of Virtue: Post-Liberalism and the Human Future* (London: Rowman & Littlefield, 2016).

²¹⁷ Thanks to Lionel Tarassenko for suggesting this example.

There is always a human conversation to be had as to why an optimisation is being organised and what kinds of correlations and estimations count as helpful.²¹⁸ That brings wider society back in.

c. Algorithms & justice

Justice is sometimes used to illustrate the difference between scientific decisions and moral decisions.²¹⁹ Whether or not the accused committed the crime can, in a broad sense, be considered a scientific question. What the just penalty should be is not. A person stands in court, accused of a crime. The first job of the court is to decide whether or not the accused committed the crime. Whatever investigative methods may be applied in practice, this is in part a question of reliable and valid evidence. As methods of forensic science become more advanced and more dependable, they have much to offer courts in amassing evidence. If one day machines are trained in evaluating probabilities that individual pieces of evidence stand up to scrutiny, then in cases that depend on complex evidence of varying degrees of confidence, there is every reason to expect that machines will help in identifying whether the accused committed the crime.

It is not as if humans are completely reliable or consistent in reaching decisions that affect the freedom of another human. A widely cited case suggested that parole decisions are affected by whether the judge

²¹⁸ In philosophical terms, this is the forming of a normative hierarchy of incommensurable values, requiring identification of emergent goods (as opposed to aggregating utility). ‘Emergence’ as used in Smith, C., *What is a Person? Rethinking Humanity, Social Life, and the Moral Good from the Person Up* (Chicago: The University of Chicago Press, 2010).

²¹⁹ Steane, A., *Faithful to Science: The Role of Science in Religion* (Oxford: Oxford University Press, 2014).

makes them before or after lunch,²²⁰ though the study does not seem entirely to stand up to scrutiny.²²¹ As early as 1928 the Illinois State Board of Parole published a study comparing predictions of 3,000 parole outcomes by unweighted statistical analysis of 21 factors with subjective judgments by three prison psychiatrists. The psychiatrists were slightly better than the statistical analysis in predicting desistance, but significantly less accurate in predicting recidivism. However careful humans seek to be in discharging their responsibilities, any human endeavour involves uncertainty and variability. Humans are not free of inconsistency, nor are they free from bias (which is why we all need hidden bias training). What about machines? They have the potential to be more consistent than humans, since, as far as we know, their information processing is not affected by the timing of lunch. Bias is also an issue here, however, since machine learning relies heavily on the data used to train the machine.²²² The basic approach is to identify characteristics of convicts likely to determine recidivism, and then to train the machine on previous recidivism rates in order that it learns how to predict future recidivism. But such correlations might be hard to distinguish from prejudice, especially as the more successful the machines are the more selective further data will become (only those granted parole can reoffend while on parole), in a way that has the potential to reinforce bias. It is a general principle of a legal judgment that it should be open to challenge or appeal, but that is hard in the case of machines, partly because of the difficulty of knowing how they have made use of the training data, and partly because in many cases the code is proprietary,

²²⁰ Danziger, S., Levav, J. & Avnaim-Pesso, L., 'Extraneous factors in judicial decisions'. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 108, No. 17 (2011), pp. 6889–6892.

²²¹ Weinshall-Margel, K. & Shapard, J., 'Overlooked factors in the analysis of parole decisions'. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 108, No. 42 (2011), p. E833.

²²² Safiya Umoja, N., *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018).

and not available for scrutiny.²²³ Similar difficulties may arise for other decisions in which human freedom is restricted, such as taking children away from their parents, and here too if the data can be assembled in a useable form, the machines may be able to assist in improved assessment of the inevitable uncertainties.

Decisions about parole are in some ways intermediate between decisions about guilty or not guilty and decisions about sentencing. If the decision about guilt or innocence is in principle (however hard this might be in practice) scientific on the basis of available evidence, the decision about a just sentence is not, since no amount of scientific enquiry can establish what is just. A utilitarian attempt might be, “What sentence will yield the greatest aggregate happiness summed over everyone who will be affected?”, but if the effect on others would be small, then the societal cost of shorter sentences might be outweighed by the increase in happiness of the guilty person, which might not be recognised as just. An alternative consequentialist attempt might be, “What sentence will minimise the probability of reoffending?”, to which the answer might be execution or a life sentence, since then the probability of reoffending would be zero. Whatever collateral benefits restricting a person’s freedom may deliver, justice comes first, and justice involves moral decisions that need other modes of enquiry than science can deliver.

Does this mean that passing judgment in a court of law is something that only humans can do? Perhaps the question should be rephrased as whether humans will invariably do better than machines. In the first decade of the 21st century it became widely recognised that there were inconsistencies between sentencing for similar offences, in what the UK press called a postcode lottery. To address this need, in 2010 the

²²³ See the emerging debate in Wachter, S., Mittelstadt, B. & Floridi, L., ‘Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation’. *International Data Privacy Law*, Vol. 7, No. 2 (2017), pp. 76-99.

Sentencing Council for England and Wales was set up to promote greater transparency and consistency in sentencing. This set out factors which judges must consider in passing sentence,²²⁴ and ranges of tariffs for different kinds of crimes. In a sense, it was a step in the direction of algorithmic sentencing. It is perfectly conceivable that a machine might build on this, together with a strong reliance on case law, to make more consistent judgments than a human is capable of.

Inverse reinforcement learning, and new techniques called imitation learning, may be suitable for this purpose. Conventional reinforcement learning works by giving the machine a goal, and then encouraging or discouraging it to keep going in the same manner according to whether it achieves or fails to achieve the goal. For example, software for a self-driving car may have the goal “drive me to my destination safely.” If a human observer in the car monitoring the performance deems it necessary to intervene, then the machine is penalised for incorrect behaviour and encouraged when it achieves the goal without intervention. In that way the machine learns to drive more safely. Imitation learning enables the machine instead to learn how a human achieves the goal. An expert human driver controls the car, and the machine is tasked with predicting what the driver will do next. The machine is penalised when it gets the predictions wrong and rewarded when it gets them right. When a satisfactory level of accuracy is achieved in a sufficient range of circumstances, the machine is deemed to be able to drive to the standard of the human. Of course, it can still make bad decisions, especially in circumstances that differ too much from anything it has encountered in training, but so can a human driver. Notwithstanding difficult questions of human

²²⁴ Factors include some, such as remorse, which are difficult to evaluate. See Maslen, H., *Remorse, Penal Theory and Sentencing* (London: Hart Publishing, 2015).

perception of risk, the machine may eventually drive more consistently and safely than a human.²²⁵

Could this principle be applied to court judgments? Just as a machine-driven car is only as good as the sensory data available, so the machine would need to be able to acquire the kind of information required by the criteria laid down by the Sentencing Council, together with (and this is easy for machine learning) access to all the case law and acts of Parliament. The machine could then be trained by imitation learning until it reaches an acceptably accurate degree of prediction. Now comes the advantage. Humans have only limited opportunities to learn from others, including their peers. The machine, having first been trained on historic cases, can learn from every case going. The experience base can therefore be as large as the number of cases available. Machine analysis of cases might not be perfect (what would that mean?), but could achieve a greater degree of consistency than human judges. But in the backdrop there remains open debate on the exact difference between scientific enquiry and moral reasoning, and the extent to which the justice system relies on the latter especially for determining guilt and appropriate sentencing. There is something about legal equality and rule of law that seems to rest on humans judging other humans so as to preserve society's understanding of moral culpability. One person says to another, "the Court finds you guilty", with all aware that we share as humans this capacity for wrongdoing, with mutual free will at the centre of why one should not just be guilty but feel guilty when committing a wrong.

²²⁵ For discussion of the moral decisions involved in programming driverless cars, see Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J. & Rahwan, I., 'The Moral Machine experiment'. *Nature*, Vol. 563 (2018), pp. 59-64; Frank, D., Chrysochou, P., Mitkidis, P. & Ariely, D., 'Human decision-making biases in the moral dilemmas of autonomous vehicles'. *Scientific Reports*, Vol. 9, No. 13080 (2019), pp. 1-19.

d. The distinct role of public reason in democratic deliberation

One of the challenges of growing up is coming to terms with one's irrationality. We may like to think of ourselves as rational beings who carefully weigh the evidence and arguments, and then come to a logical conclusion supported by the best of human reasoning. If only it were so! Perhaps sometimes it is, in academic scholarship and in engineering and medicine, and maybe in business and commerce. But in life choices in general, and democratic deliberation in particular?

For life decisions, there seems to be growing evidence that it is often the other way around. Humans are capable of committing to a particular position or course of action, and then using their powers of reasoning to justify that commitment.²²⁶ It could be that they have reached that conclusion through the exercise of what we are pleased to call rationality, however subject that may be to all sorts of overt and latent influences. Suppose you are a hunter-gatherer living in the savannahs, wanting to kill an antelope for food. What evolutionary advantage would language give you? It might enable you to communicate the ballistics of spear throwing to your companions, explaining to them that for maximum range you should throw the projectile at an elevation of 45 degrees. But it might carry even more survival advantage if, having decided yourself that you want to go antelope hunting for which you need assistance, your linguistic fluency helps you to persuade your companions to go along with your plan. Fast forward to the present day, where politicians often seem capable of arguing very persuasively for particular policies for reasons which might have little to do with how they actually chose those policies in the first place. Can a more rational and consistent process of collective decision-making not be found through the use of AI, algorithms and digital technologies?

²²⁶ Haidt, J., *The Righteous Mind: Why Good People Are Divided by Politics and Religion* (London: Penguin, 2013).

If machines were to contribute to democratic deliberation, what sorts of values would need to be built into them? Interestingly, some go so far as to argue that it is not a requirement that the machine shares or feels such values, only that it uses the values or set of values in its decision-making processes.²²⁷ Complicating, however, assumptions that rational solutions can be found for achieving social coordination is the fact that there is a growing move away from the conventional description of *Homo economicus* within economics and the social sciences generally.²²⁸ Not everything that matters can be measured in monetary terms,²²⁹ and this makes it difficult to identify social policies that are comparatively more optimal without reference to the norms, cultures and endogenous behaviours of the people concerned. It is possible and in fact quite regular that corporations set targets for their CEOs other than simply maximising shareholder value.²³⁰ In many circumstances, identity plays a decisive role in motivation and moral choices.²³¹ It is even possible to conceive of free market economics based on robust ethical principles.²³² There is growing evidence that the relatively recent evolution of humans has involved a co-evolution

²²⁷ Russell, 2019.

²²⁸ Collier, P., 'Greed is dead: The recognition that we need to rely on each other rather than ourselves'. *Times Literary Supplement* (6 Dec 2019). <https://www.the-tls.co.uk/articles/greed-is-dead/>.

²²⁹ Sandel, M. J., *What Money Can't Buy: The Moral Limits of Markets*, (London: Penguin, 2012).

²³⁰ Mayer, C., *Prosperity: Better Business Makes the Greater Good* (Oxford: Oxford University Press, 2018).

²³¹ Akerlof, G. A. & Kranton, R. E., *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being* (Princeton: Princeton University Press, 2010).

²³² Bruni, L. & Sugden, R., 'Fraternity: Why the Market Need Not Be a Morally Free Zone'. *Economics and Philosophy*, Vol. 24, No. 1 (2008), pp. 35-64; Bruni, L. & Sugden, R., 'Reclaiming Virtue Ethics for Economics'. *Journal of Economic Perspectives*, Vol. 27, No. 4 (2013), pp. 141-164; Burbidge, D., 'Space for virtue in the economics of Kenneth J. Arrow, Amartya Sen and Elinor Ostrom'. *Journal of Economic Methodology*, Vol. 23, No. 4 (2016), pp. 396-412; Collier, P., *The Future of Capitalism: Facing the New Anxieties* (London: Allen Lane, 2018).

of prosocial skills and the genetic wherewithal to manifest them;²³³ humans need to cooperate and they have evolved to cooperate.²³⁴ Thus, building the value of cooperation into the score function of machine decision-making seems to be a requirement for fully appreciating the way humans are and the way societies are.

It seems that even when people want to make good choices from the highest of motives, there is an inseparable emotional element in their decision-making process.²³⁵ Contrary to what every parent of a teenager is tempted to think—if only my son/daughter were not so emotional, he/she would live much better—the evidence suggests that where an individual’s rational cognitive processes become disconnected from their emotions, they make worse life choices. It seems that humans have evolved so that the emotional dimension is essential to moral decision-making.²³⁶ Although work on affective computing is advancing apace,²³⁷ it is too early to say how, if ever, machines will acquire the kind of emotional capacity which seems to be so crucial in good moral decision-making. Easy moral decisions could presumably be made by humans or machines, with the same outcome either way. For hard moral decisions, of the kind that require democratic deliberation, it may be that either/or debates about human/machine involvement should be replaced by careful consideration as to how the best of human moral capacity, including

²³³ Henrich, J., *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (Princeton: Princeton University Press, 2017).

²³⁴ Christakis, N. A., *Blueprint: The Evolutionary Origins of a Good Society* (New York: Little, Brown Spark, 2019).

²³⁵ Damásio, A., *Descartes’ Error: Emotion, Reason, and the Human Brain* (New York: G. P. Putnam, 1994).

²³⁶ Jeeves, M. A. & Ludwig, T. E., *Psychological Science and Christian Faith: Insights and Enrichments from Constructive Dialogue* (West Conshohocken: Templeton Press, 2018).

²³⁷ See, for example, Picard, R. W., *Affective Computing* (Cambridge, MA: MIT Press, 1997).

its emotional component, can harness the best of machine learning in the face of numerous uncertainties.

Achieving this kind of balance is not an easy question by any stretch of the imagination. Perhaps unfortunately, it is not possible to optimise only one thing or a commensuration of distinct values when seeking to pursue human flourishing, due to the irreducibility of human values to a single, commensurate unit of analysis. This means that algorithmic decision-making on its own will likely fail to capture distinctly human valuations. And, to the extent that it removes humans from the decision-making process, it defers human valuation, leading to mistrust in the collective priorities being formed. As Zuboff remarks:

Surveillance capitalism offers solutions to individuals in the form of social connection, access to information, time-saving convenience, and, too often, the illusion of support. [...] More significantly, it offers solutions to institutions in the form of omniscience, control, and certainty. The idea here is not to heal instability—the corrosion of social trust and its broken bonds of reciprocity, dangerous extremes of inequality, regimes of exclusion—but rather to exploit the vulnerabilities produced by these conditions.

It is useful to note that despite the much-touted social advantages of always-on connection, social trust in the US declined precipitously during the same period that surveillance capitalism flourished.²³⁸

However, such criticism amounts only to a negative argument: saying that algorithmic optimisation is not always best in the way currently being exercised. It is important to also offer something more positive, in terms of the distinct value of democracy and the persistent relevance of moral reasoning for democratic and judicial decision-

²³⁸ Zuboff, 2019, p. 383.

making. Something more positive would better answer why those at the frontier of these technological developments—whether in Silicon Valley or in the highest offices of authoritarian regimes—should, as individuals, themselves prefer democratic forms of engagement in the quest to optimise decision-making.

The social trust that is everywhere eroding forces us to reflect on the basis to our social unity. Trust itself is, ultimately, about intentional unity.²³⁹ This means we have to start treating unity as a good per se—something that helps reveal the vocation of democratic engagement for our networked age, and the unique value of citizenship for decision-making.

Our civic unity is in the domain of our citizenship, not our material wealth. The togetherness of our “comm-unity” is not accidental: it is not that we just happen to be in the same place at the same time (something that the networked age changes). “Comm-unity” can be purposefully chosen as a good in itself. If it is not, the networked age will merely become a centralisation of data nodes: a centralisation that then drives us apart by optimising unimportant aspects of our human flourishing, such as our desire that our mood be validated, or our wish to continue fighting with someone even though they are no longer physically in the same room. These techniques of our networked age mostly reduce the capacity for human moral reasoning by substituting relationships with rapidity of response, leading to mistrust in the collective priorities being formed.

²³⁹ Burbidge, D., ‘Urban trust in Kenya and Tanzania: Cooperation in the provision of public goods’. *Canadian Journal of African Studies*, Vol. 47, No. 3 (2013), pp. 465-482, p. 468.



'I'd like to join the parish council.

I would be willing to push the nuclear button'

Figure 5.2 Local democracy²⁴⁰

Analysis of decision-making and public reason has got lost in discussing what is optimal according to correlations and likely outcomes, which AI and big data have the potential to be much better at determining, when citizenship is at least in part about decision-making for the purpose of unity itself. Citizenship is about bringing people together because togetherness is a good thing. We are social in nature, and evolutionarily specialise in forming long-term, nurturing relationships.²⁴¹ Being together is *a part of being*, making our willingness to learn from each other, respect each other, make each other laugh and love each other integral to our species' success. If we lose our collective unity we lose our distinctiveness in being human.

²⁴⁰ The Telegraph, 'Best of Matt' (28 Dec 2018).
<https://www.telegraph.co.uk/news/2016/03/31/best-of-matt/matt-cartoon15>.

²⁴¹ Christakis, 2019.

In order to foster this distinctiveness, we do some things for the sake of unity itself, which means the way of deciding things is more important than what is decided. We should not wait for imagined future economic gain or technological progress in the hope that that will bring about unity; unity is brought about consciously and purposefully by communities that value it for its own sake, shown in their democratic decision-making.

How is it that unity can be considered a good thing in itself? A simple way of seeing this is in the “battle of the sexes” dilemma outlined in game theory. The decision that needs to be made is what to do in the evening, with the two options being visiting a restaurant or going to the cinema. The husband prefers the former while the wife prefers the latter. The payoff structure of the game theory design typically describes that, even though each person has a preferred activity, the pair ultimately prefer to do something together.

battle of the sexes payoff matrix

		HUSBAND	
		Cinema	Restaurant
WIFE	Cinema	3, 2	1, 1
	Restaurant	0, 0	2, 3

Figure 5.3 Game theory payoff matrix for the “battle of the sexes”

Doing something together is more important than what activity, in particular, is done. The game thus has two Nash equilibria, meaning there is no optimal solution that can be rationally pursued at the individual level (it could be either the cinema or the restaurant). Instead, finding a solution becomes about developing understanding and communication between the parties, even if it just takes the form of agreeing to flip a coin together.

The battle of the sexes game represents more of life than we may realise, and helps us see that decision-making strategies cannot always be optimised through making information about preferences more available, or by increasing the number of people in the group.

On one level the battle of the sexes explains the point, but on another level it falls short, because it is still trying to provide a single metric for evaluating incommensurable goods. It is helpful for showing that the optimisation exercise is limited by its own logic for treating distinctly human valuations, but it is unhelpful for really explaining the real workings of unity, which can only be experienced.

Pushed to its extremes, the case being made is that it is sometimes better to make bad decisions together than good decisions apart. That is often the case in marriage, where it is important that both husband and wife have the same sort of ideas of what they would do if they won the lottery, even though those ideas may be awful.

In a good marriage you will pursue those awful ideas together.²⁴² To the extent that our networked age is removing the space for intentional unity, it is leading to mistrust in the collective priorities being formed.

²⁴² As Chris Stewart poetically recounts when autobiographically explaining his first day moving into a dilapidated house in the south of Spain with his wife: “Well, here we are. This is home. Here we lay our bones.” We laughed and walked arm in arm up to the terrace where we sat dangling our legs over the drop below while the sun slipped down behind the hill.

What we needed was a cup of tea. [...] Nothing that we had brought with us up to the house was suitable for that purpose and I refused adamantly to unload and go back across the river to where we had left the trailer before I had drained my first cup.

We eventually found a bent aluminium pot. The sort of pot you boil up handkerchiefs in. It looked as if a mule had trodden on it. Then we built a fire of twigs, filled the pot with water from the pomegranate-dribbling hose, and suspended it over the flames with some bits of rusty wire. When the water began to smoke—not steam, oddly enough, but smoke—we removed it from the heat and put in some sort of tea-bag we’d located. Then we covered it with a flat stone to mast.

“Cups, cups, cups ... what shall we do for cups?” But of course! There were some empty tuna-fish tins lying around here and there. I took a couple and went to scrub them in the water-butt. “Have six minutes elapsed yet?” They had, and we poured the loathsome grey liquid into the tuna tins.

“You didn’t wash the cups very well,” said Ana accusingly.

“I did the best I could—they’re alright.”

A scum of fish-oil was floating on top of the tea. We sat back and sighed, gazing at the lovely view of rivers and mountains below us, while we sipped what must surely have been the most detestable beverage ever to pass the lips of man.

Nonetheless, we have kept as family treasures the paraphernalia of that first brew and on November 26 each year we celebrate El Valero Day by seeking to surpass in vileness that first momentous cup of tea.’ Stewart, C., *Driving Over Lemons: An Optimist in Andalucía* (London: Sort of Books, 2011), pp. 99-100.



Conclusion

Citizenship 
in a Networked Age

Conclusion



IT IS OFTEN FELT THAT OUR SOCIAL DIVISIONS ARE ECONOMIC, AND THAT IF WE WERE BETTER AT SHARING OUR GROWTH EQUALLY, WE WOULD HEAL THESE DIVISIONS.²⁴³ That tells one side of the story. The other side is our civic unity, which is in the domain of our citizenship. The togetherness of our community is not accidental. It's not just that we happen to be in the same place—something that the networked age breaks up—and it's not just that some of our lifestyle choices lead to strong communities while others do not. Community has to be purposefully chosen as a good in itself. Analysis of decision-making and public reason has got lost in discussing what is optimal according to correlations and likely outcomes, which AI, algorithms and big data have the potential to be much better than humans at determining. But citizenship is about decision-making in part for the purpose of community itself. It is about bringing people together because that togetherness is a good thing. If AI, algorithms and big data can better determine optimal material outcomes, that frees us up to decide things for the sake of social unity. Over time, we should focus less on optimal material outcomes and more on optimal togetherness. Of course, many decisions involve both determining material outcomes and keeping us together as a community. The argument here is simply that it is wrong to see decision-making solely in terms of material outcomes, and that we should keep at the forefront of our minds humans' unique ability to engage in moral reasoning. Exploring this, as was done in Chapter 1, affords us breathing space for engineering a new take on citizenship for our networked age. We cannot wait for imagined future economic gain to bring about unity; unity is brought about consciously and purposefully by communities that value it for its own sake. As Chapters 2 and 3 pointed out, while politics has often

²⁴³ Piketty, T., *Capital in the Twenty-First Century* (Cambridge, MA: Harvard University Press, 2014).

been discussed in terms of managing *the scarcity of economic resources*, it would be better in the networked age to say it is about managing *the scarcity of time*, because achieving unity above all takes time in nurturing interpersonal relations. Chapter 5 demonstrated that a decision where everyone has had their say is qualitatively different to a decision made instantaneously, *even when the same thing is decided*. When we choose to let everyone have their say we are “spending” our scarce resource of time, because we believe participatory engagement should be prioritised in this instance. As we give people our time, we learn more about *why* they think the way they do, and we see areas where giving up something minor to us is of major benefit to them. Citizenship in a networked age is about this management of the scarcity of time for the sake of unity. As Danielle Allen writes:

The precondition of democratic decision making is unity. If a political system that relies heavily on majority rule cannot keep minorities affixed to it through loyalty, then every fresh, durable minority faction that comes into being will bring with it the threat of breakup. A first secession will provide grounds for a second, and on and on; the polity will face a threat to its very existence.²⁴⁴

THE RECOMMENDATIONS OF THIS REPORT ARE THEREFORE EACH CONNECTED TO THE NEED FOR UNITY AS A SOCIETAL GOAL. Ultimately, we are entering a changing public decision-making environment that increasingly relies on forms of optimal response through data aggregation, pattern identification and the use of algorithms and machine learning to draw out the most salient features of data flows. In the midst of these decision-making upheavals, we need to pay attention to what they mean for our growth as a self-aware

²⁴⁴ Allen, D. S., ‘The Road From Serfdom: How Americans can become citizens again’. *The Atlantic* (Dec 2019).
<https://www.theatlantic.com/magazine/archive/2019/12/danielle-allen-american-citizens-serfdom/600778/>.

moral community. The report therefore makes the following recommendations to rebuild our civic ideals in this networked age:

Recommendation 1

Identify and protect human
uniqueness for moral
decision-making

Chapter 1 demonstrated that no matter the advances in big data and big computation, there will always be a space for goal-setting that is uniquely human. The feared displacement of human decision-making through algorithms, machine learning and digital technologies should not be considered a threat to democracy because humans remain unique in the way they are able to order pursuit of the common good through their moral reasoning.



Recommendation 2

Nurture the complementary skills of humans and machines for collective decision-making

Human life is circumscribed by mortality in a way that machine life is not, as Chapter 1 further explained. This gives humans a distinct, self-aware capacity for moral decision-making which no machine can replace. The challenge is to develop methods of moral decision-making which benefit from machine optimisation of technical aspects in service of the human judgment of the moral whole. For pursuit of human flourishing in this networked age, the choice is not between citizens and machines; rather it is about supporting fruitful areas of human-machine collective decision-making.



Recommendation 3

Engage in consensus-building
about civic ideals for a
networked age

Civic engagement is needed because there is no one goal worth pursuing on which everyone is agreed. The effect of new technologies is generally not to raise new moral and social issues which did not previously exist, but rather to give a new focus and poignancy to norms of behaviour which long predate modern communications, a point offered in Chapter 2. The economic success of big data industries is bound up with the ownership and privileged use of that data. Regulation will need to be based, however, not on the relationship between corporations and consumers but on the relationship between democratic institutions and citizens. In order to use methods of machine learning in support of collective decision-making, there needs to be sufficient agreement about the values that these optimisations are oriented around. This provides a new context and motivation for developing our principles as a society and defining what human flourishing means.



Recommendation 4 Teach listening as a civic virtue

Chapter 3 demonstrates that the civic virtue most appropriate for the networked age is listening well. While education has had a major focus from Ancient Greece onwards on the art of rhetoric and what it means to be a good speaker, we are now entering an age where primacy must instead be placed on being a good listener. That means greater focus on humility, introversion, trust and social capital, and reduced focus on agency, leadership and self-determination. Such a shift in focus would of course be counter-cultural: maturity is commonly associated with taking ownership of one's own affairs. The networked age, however, requires deep levels of sharing and connective collaboration, and therefore deep levels of trust. The internet provides so much information at one's disposal that our "what?" questions are already receiving great assistance; education instead needs to help focus on the "why?" question, which requires listening for reasons and motivations. Getting a good grasp of the "why" helps generate social trust and solidarity because different viewpoints remain traceable to the common ground that we are all seeking happiness. The benefits lie both in terms of economic advancement and political togetherness, accessed through developing better listening. The networked age encourages economic specialism and offers more channels of communication between those specialisms. The onus, therefore, is on the art of listening in being able to join together distinct pursuits so as to establish and strategize their meaning-based contributions. On the political front, the increased ability of everyone to state their views and publicise them to everybody else means there is no longer any scarcity of good speakers but, instead, a scarcity of good listeners. Citizenship through good listening needs to be taught as a civic virtue.



Recommendation 5

Maintain distance between thought and speech

Privacy provides special space for citizens to reflect on what matters most to them, as outlined in Chapter 4, and works best with some distance being maintained between thought and speech. Privacy is widely recognised as a fundamental human right, based on principles of human dignity, and is linked to other rights such as equal treatment and free expression. The networked age gives unprecedented opportunity and even encouragement for individuals to voluntarily forgo much of their privacy. In turn, the distinction between the private realm and the public realm becomes increasingly blurred. End-to-end encryption gives some protection to individual privacy, but the metadata can still be harvested. Machine learning can help in the fight to filter out fake news, but ultimately the reader must be aware of the need to form their views independently and thoughtfully, often in private first. In developing civic ideals, digital privacy provides space, therefore, for society's moral development. In particular, it protects against the demand for instant response. In private, individuals and groups are able to self-initiate their search for truth, meaning, and purpose.



Recommendation 6

Promote the value of privacy
for personal moral development

Privacy is essential for each person's ethical development in our networked age. Its value is being underappreciated because it is framed in terms of immunity from others' interventions and incursions, without reference to its civic contribution. Efficiency-gains in sharing data so often seem to outweigh individualistically-defined benefits of immunity from others, relegating privacy to a secondary concern. As Chapter 4 further explains, privacy is necessary, however, for each person's moral development—that is its civic contribution. Privacy needs to be redefined in terms of personal moral development. There is a universal civic need to develop views initially in closed settings with robust forgetfulness for bad or incorrect things said.



Recommendation 7

Revalue democracy in terms
of the ability to bring about
social unity and trust

Chapter 5 argued that democratic decision-making is all about agreeing on a way forward in the context of diverse and often conflicting interests. These decisions involve both technical and moral components. Machine learning surpasses human abilities in an increasing range of technical decisions. Democratic decision-making, however, involves building unity among people as part of the process of deciding what is just and moral. A crucial ingredient in democratic decision-making, which machines will never be able to replace, is the value of the discourse which goes into the decision-making process. At its best, under wise and inclusive leadership, democracy facilitates shared commitment to the agreed course of action. Good democratic decision-making needs to be based on sound technical and moral reasoning, with shared ownership of the outcome.



Bibliography

- Acemoglu, D. & Robinson, J. A., *Why Nations Fail: The Origins of Power, Prosperity and Poverty* (London: Profile Books, 2012).
- Akerlof, G. A. & Kranton, R. E., *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being* (Princeton: Princeton University Press, 2010).
- Allen, D. S., *Talking to Strangers: Anxieties of Citizenship since Brown v. Board of Education* (Chicago: The University of Chicago Press, 2004).
- Allen, D. S., 'The Road From Serfdom: How Americans can become citizens again'. *The Atlantic* (Dec 2019).
<https://www.theatlantic.com/magazine/archive/2019/12/danielle-allen-american-citizens-serfdom/600778/>.
- Anderson, B., *Imagined Communities: Reflections on the Origin and Spread of Nationalism* (London: Verso, 2016 [1983]), Revised Ed.
- Anderson, C., 'The End of Theory: The Data Deluge Makes the Scientific Method Obsolete'. *Wired* (23 Jun 2008).
<https://www.wired.com/2008/06/pb-theory/>.
- Anderson, C. W., *Apostles of Certainty: Data Journalism and the Politics of Doubt* (Oxford: Oxford University Press, 2018).
- Aristotle, *The Politics* (London: Penguin, 1992).
- Aristotle, *The Nicomachean Ethics* (London: Penguin, 2004).
- Askonas, J., 'How Tech Utopia Fostered Tyranny'. *The New Atlantis* (Winter, 2019). <https://www.thenewatlantis.com/publications/how-tech-utopia-fostered-tyranny>.
- Augustine, *City of God* (London: Oxford University Press, 1963 [426]).
- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J. & Rahwan, I., 'The Moral Machine experiment'. *Nature*, Vol. 563 (2018), pp. 59-64.
- Baldwin-Philippi, J., *Using Technology, Building Democracy* (Oxford: Oxford University Press, 2015).
- Barlow, J. P., 'A Declaration of the Independence of Cyberspace' (Davos, 8 Feb 1996). <https://www.eff.org/cyberspace-independence>.
- BBC, 'Ethics guide: Philosophers justifying slavery' (2014).
http://www.bbc.co.uk/ethics/slavery/ethics/philosophers_1.shtml.
- BBC News, 'Facebook to integrate WhatsApp, Instagram and Messenger' (25 Jan 2019). <https://www.bbc.co.uk/news/technology-47001460>.

- BBC News, 'UK's controversial "porn blocker" plan dropped' (16 Oct 2019).
<https://www.bbc.co.uk/news/technology-50073102>.
- Bello, J., 'The making and consequences of everyday political discussion networks'. DPhil thesis, University of Oxford (2011).
- Bentham, J., *An Introduction to the Principles of Morals and Legislation* (London: W. Pickering, 1823).
- Bershidsky, L., 'End-to-End Encryption Isn't as Safe as You Think'. *Bloomberg* (14 May 2019).
<https://www.bloomberg.com/opinion/articles/2019-05-14/whatsapp-hack-shows-end-to-end-encryption-is-pointless>.
- Beyer, J. L., *Expect Us: Online Communities and Political Mobilization* (Oxford: Oxford University Press, 2014).
- Bhagat, S., Burke, M., Diuk, C., Filiz, I. O., Edunov, S., 'Three and a half degrees of separation'. *Facebook Research* (4 Feb 2016).
<https://research.fb.com/blog/2016/02/three-and-a-half-degrees-of-separation/>.
- Bollen, J., Gonçalves, B., van de Leemput, I. & Ruan, G., 'The happiness paradox: your friends are happier than you'. *EPJ Data Science*, Vol. 6, No. 4 (2017), pp. 1-10.
- Bostrom, N., *Superintelligence: paths, dangers, strategies* (Oxford: Oxford University Press, 2014).
- Boyd, D., 'Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications'. Ch 2 of Papacharissi, Z. (ed.), *A Networked Self: Identity, Community, and Culture on Social Network Sites* (New York: Routledge, 2011).
- Brogan, J., 'What's the Deal With Algorithms? Your 101 guide to the computer codes that are shaping the ways we live'. *Slate* (2 Feb 2016).
<https://slate.com/technology/2016/02/whats-the-deal-with-algorithms.html>.
- Bruni, L. & Sugden, R., 'Fraternity: Why the Market Need Not Be a Morally Free Zone'. *Economics and Philosophy*, Vol. 24, No. 1 (2008), pp. 35-64.
- Bruni, L. & Sugden, R., 'Reclaiming Virtue Ethics for Economics'. *Journal of Economic Perspectives*, Vol. 27, No. 4 (2013), pp. 141-164.
- Bucher, T., *If... Then: Algorithmic Power and Politics* (Oxford: Oxford University Press, 2018).
- Burbidge, D., 'The Security Dilemma in North Kivu, the Democratic Republic of the Congo'. *Conflict Trends*, Vol. 3 (2009), pp. 42-49.
- Burbidge, D., 'Urban trust in Kenya and Tanzania: Cooperation in the provision of public goods'. *Canadian Journal of African Studies*, Vol. 47, No. 3 (2013), pp. 465-482.
- Burbidge, D., 'Connecting African Jurisprudence to Universal Jurisprudence through a Shared Understanding of Contract'. Ch 5 of Onazi, O. (ed.),

- African Legal Theory and Contemporary Problems: Critical Essays* (Dordrecht: Springer, 2014).
- Burbidge, D., *The Shadow of Kenyan Democracy: Widespread Expectations of Widespread Corruption* (Burlington: Ashgate, 2015).
- Burbidge, D., 'Space for virtue in the economics of Kenneth J. Arrow, Amartya Sen and Elinor Ostrom'. *Journal of Economic Methodology*, Vol. 23, No. 4 (2016), pp. 396-412.
- Burbidge, D., 'The Inherently Political Nature of Subsidiarity'. *American Journal of Jurisprudence*, Vol. 62, No. 2 (2017), p. 143-164.
- Byers, A., *Jeff Bezos: The Founder of amazon.com* (New York: The Rosen Publishing Group, 2007).
- Cambridge Dictionary (Cambridge University Press, 2019).
<https://dictionary.cambridge.org/dictionary/english/intelligence>.
- Castro, D., 'A Declaration of the Interdependence of Cyberspace'. *Computerworld* (8 Feb 2013).
<https://www.computerworld.com/article/2494710/a-declaration-of-the-interdependence-of-cyberspace.html>.
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M. & Floridi, L., 'Artificial Intelligence and the "Good Society": the US, EU, and UK approach'. *Science and Engineering Ethics*, Vol. 24, No. 2 (2018), pp. 505-528.
- Chakrabarti, S., 'Nationalism is a driving force behind fake news in India, research shows'. *BBC News* (12 Nov 2018a).
<https://www.bbc.com/news/world-46146877>.
- Chakrabarti, S., 'Duty, Identity, Credibility: "Fake News" and the ordinary citizen in India', with Stengel, L. & Solanki, S. *BBC News* (2018b).
<http://downloads.bbc.co.uk/mediacentre/duty-identity-credibility.pdf>, p. 44 (emphasis in original).
- Chopra, R., 'In India, WhatsApp is a weapon of antisocial hatred'. *The Conversation* (23 Apr 2019a). <https://theconversation.com/in-india-whatsapp-is-a-weapon-of-antisocial-hatred-115673>.
- Chopra, R., *The Virtual Hindu Rashtra: Saffron Nationalism and New Media* (Noida: HarperCollins India, 2019b).
- Christakis, N. A., *Blueprint: The Evolutionary Origins of a Good Society* (New York: Little, Brown Spark, 2019).
- Christian, B., *The Most Human Human: What Artificial Intelligence Teaches Us About Being Alive* (London: Penguin Books, 2011).
- Chrome Web Store, 'I don't care about cookies' (29 Jun 2019).
<https://chrome.google.com/webstore/detail/i-dont-care-about-cookies/fihnjjcciajhdojfnbdddfoaknhlnja>.
- Cohen, J. E., 'What Privacy Is For'. *Harvard Law Review*, Vol. 126 (2013), pp. 1904-1933.

- Cohen, J. E., 'Law for the Platform Economy'. *U.C. Davis Law Review*, Vol. 51 (2017), pp. 133-204.
- Collier, P., *The Future of Capitalism: Facing the New Anxieties* (London: Allen Lane, 2018).
- Collier, P., 'Greed is dead: The recognition that we need to rely on each other rather than ourselves'. *Times Literary Supplement* (6 Dec 2019). <https://www.the-tls.co.uk/articles/greed-is-dead/>.
- Dafoe, A., 'AI Governance: A Research Agenda'. Future of Humanity Institute, University of Oxford (Aug 2018).
- Damásio, A., *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: G. P. Putnam, 1994).
- Danziger, S., Levav, J. & Avnaim-Pesso, L., 'Extraneous factors in judicial decisions'. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 108, No. 17 (2011), pp. 6889–6892.
- Deneen, P. J., *Why Liberalism Failed* (New Haven: Yale University Press, 2018).
- Drexler, K. E., 'Reframing Superintelligence: Comprehensive AI Services as General Intelligence'. Technical Report, No. 2019-1, Future of Humanity Institute, University of Oxford (2019).
- Dubois, E. & Blank, G., 'The echo chamber is overstated: the moderating effect of political interest and diverse media'. *Information, Communication & Society*, Vol. 21, No. 5 (2018), pp. 729-745.
- Duff, A., 'The Problem of Rule in MacIntyre's Politics and *Ethics in the Conflicts of Modernity*'. *Politics & Poetics*, Vol. 4 (2018), pp. 1-21.
- Dunbar, R. I. M., 'Coevolution of neocortical size, group size and language in humans'. *Behavioral and Brain Sciences*, Vol. 16, No. 4 (1993), pp. 681-735.
- Dworkin, R., *Taking Rights Seriously* (London: Bloomsbury, 2013 [1977]).
- Elshtain, J. B., *Sovereignty: God, State, and Self* (New York: Basic Books, 2008).
- Emmanouilidou, L., 'Facebook wants to create a "Supreme Court" for content moderation. Will it work?' *PRI* (5 Sep 2019). <https://www.pri.org/stories/2019-09-05/facebook-wants-create-supreme-court-content-moderation-will-it-work>.
- English Oxford Living Dictionaries* (Oxford University Press, 2019). <https://en.oxforddictionaries.com/definition/post-truth>.
- Ensminger, J., *Making a Market: The institutional transformation of an African society* (Cambridge: Cambridge University Press, 1996).
- Ensminger, J., 'Market Integration and Fairness: Evidence from Ultimatum, Dictator, and Public Goods Experiments in East Africa'. Ch 12 of Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E. & Gintis, H. (eds.), *Foundations of Human Sociality: Economic Experiments and*

- Ethnographic Evidence from Fifteen Small-Scale Societies* (Oxford: Oxford University Press, 2004),
- Finnis, J., ‘Commensuration and Public Reason’. Ch 12 of Chang, R. (ed.), *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1997).
- Finnis, J., *Natural Law & Natural Rights* (Oxford: Oxford University Press, 2011).
- Frank, D., Chrysochou, P., Mitkidis, P. & Ariely, D., ‘Human decision-making biases in the moral dilemmas of autonomous vehicles’. *Scientific Reports*, Vol. 9, No. 13080 (2019), pp. 1-19.
- Free Thoughts Podcast, ‘The Evolution of Cooperation (with Nicholas A. Christakis)’ (14 Feb 2020).
<https://www.libertarianism.org/podcasts/free-thoughts/evolution-cooperation-nicholas-christakis>.
- Garton Ash, T., *Free World: Why a Crisis of the West Reveals the Opportunity of Our Time* (London: Penguin Books, 2005).
- Garton Ash, T., *Free Speech: Ten Principles for a Connected World* (London: Atlantic Books, 2016).
- Gatherer, D., ‘Modelling versus Realisation: Rival Philosophies of Computational Theory in Systems Biology’. Ch 3 of Reiss, M. J., Watts, F. & Wiseman, H. (eds.), *Rethinking Biology: Public Understandings* (New Jersey: World Scientific, 2019).
- Gaventa, J., *Power and Powerlessness: Quiescence and Rebellion in an Appalachian Valley* (Urbana: University of Illinois Press, 1980).
- Geuss, R., *Public Goods, Private Goods* (Princeton: Princeton University Press, 2001).
- Ghahramani, Z., ‘Probabilistic machine learning and artificial intelligence’. *Nature*, Vol. 521 (2015), pp. 452-459.
- Gianakon, S. E., ‘Citizenship and the Holy in Late Antiquity’. Humboldt Universität, Fulbright essay (May 2018).
- Goldfarb, R., *In Confidence: When to Protect Secrecy and When to Require Disclosure* (New Haven: Yale University Press, 2009).
- Goodhart, D., *The Road to Somewhere: The New Tribes Shaping British Politics* (London: Penguin, 2017).
- Gorwa, R., ‘What is platform governance?’ *Information, Communication & Society*, Vol. 22, No. 6 (2019), pp. 854-871.
- Graham, M., ‘Big data and the end of theory?’ *The Guardian* (9 Mar 2012).
<https://www.theguardian.com/news/datablog/2012/mar/09/big-data-theory>.
- Granovetter, M., ‘Threshold Models of Collective Behavior’. *American Journal of Sociology*, Vol. 83, No. 6 (1978), pp. 1420-1443.

- Granovetter, M., 'The Strength of Weak Ties'. *American Journal of Sociology*, Vol. 78, No. 6 (1973), pp. 1360-1380.
- Greenberg, A., 'Hacker Lexicon: What is End-to-end Encryption?' *Wired* (25 Nov 2014). <https://www.wired.com/2014/11/hacker-lexicon-end-to-end-encryption/>.
- Haidt, J., *The Righteous Mind: Why Good People Are Divided by Politics and Religion* (London: Penguin, 2013).
- Harari, Y. N., *Homo Deus: A Brief History of Tomorrow* (London: Vintage, 2015).
- Hawley, S. H., 'Theopolis Monk: Envisioning a Future of A.I. Public Service'. Ch 14 of Lee, N. (ed.), *The Transhumanism Handbook* (Cham: Springer, 2019).
- Hayek, F. A., *The Road to Serfdom* (London: Routledge, 1962 [1944]).
- Helmond, A., 'The Platformization of the Web: Making Web Data Platform Ready'. *Social Media + Society*, Vol. 1, No. 2 (2015), pp. 1-11.
- Henrich, J., *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (Princeton: Princeton University Press, 2017).
- Hindman, M., *The Internet Trap: How the Digital Economy Builds Monopolies and Undermines Democracy* (Princeton: Princeton University Press, 2018).
- Hodas, N. O., Kooti, F. & Lerman, K., 'Friendship Paradox Redux: Your Friends Are More Interesting Than You'. *Association for the Advancement of Artificial Intelligence* (2013), pp. 1-9.
- Hopkins, C. C., *Scientific Advertising* (London: 1968 [1923]).
- Huffington, A., 'GPS for the Soul: A Killer App for Better Living'. *HuffPost* (16 Apr 2012). https://www.huffpost.com/entry/gps-for-the-soul_b_1427290?guccounter=1.
- Hutchinson, A., 'Facebook Reaches 2.38 Billion Users, Beats Revenue Estimates in Latest Update'. *Social Media Today* (24 Apr 2019). <https://www.socialmediatoday.com/news/facebook-reaches-238-billion-users-beats-revenue-estimates-in-latest-upda/553403/>.
- Isaac, M., 'Zuckerberg Plans to Integrate WhatsApp, Instagram and Facebook Messenger'. *New York Times* (25 Jan 2019). <https://www.nytimes.com/2019/01/25/technology/facebook-instagram-whatsapp-messenger.html>.
- Jeeves, M. A. & Ludwig, T. E., *Psychological Science and Christian Faith: Insights and Enrichments from Constructive Dialogue* (West Conshohocken: Templeton Press, 2018).
- Jones, C., 'How to Visualize Your Facebook Friend Network'. *All Things Graphed* (28 Aug 2014). <http://allthingsgraphed.com/2014/08/28/facebook-friends-network/>.

- Jordan, M., 'Artificial Intelligence – The Revolution Hasn't Happened Yet'. *Medium* (19 Apr 2018). <https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7>.
- Karpf, D., *Analytical Activism: Digital Listening and the New Political Strategy* (Oxford: Oxford University Press, 2017).
- Kastrenakes, J., 'WhatsApp limits message forwarding in fight against misinformation'. *The Verge* (21 Jan 2019). <https://www.theverge.com/2019/1/21/18191455/whatsapp-forwarding-limit-five-messages-misinformation-battle>.
- Kelion, L., 'Why Amazon knows so much about you'. *BBC News* (Feb 2020). <https://www.bbc.co.uk/news/extra/CLQYZENMBI/amazon-data>.
- Kissinger, H. A., 'How the Enlightenment Ends'. *The Atlantic* (Jun 2018). <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>.
- Kissinger, H. A., Schmidt, E. & Huttenlocher, D., 'The Metamorphosis'. *The Atlantic* (Aug 2019). <https://www.theatlantic.com/magazine/archive/2019/08/henry-kissinger-the-metamorphosis-ai/592771/> (emphasis in original).
- Laliberte, M., 'You'll Never Guess What Amazon Was Almost Called—and Why They Had to Change the Name'. *Reader's Digest* (2019). <https://www.rd.com/culture/amazons-name-origin/>.
- Law Commission, 'Data Sharing between Public Bodies: A Scoping Report', No. 351 (10 Jul 2014).
- Leber, J., 'You Are Connected To Everyone On Earth By Just 4 Degrees Now'. *Fast Company* (28 Oct 2013). <https://www.fastcompany.com/3020687/you-are-connected-to-everyone-on-earth-by-just-4-degrees-now>.
- Leverhulme Centre for the Future of Intelligence, 'Trust and Transparency' (2019a). <http://www.lcfi.ac.uk/projects/ai-trust-and-society/trust-and-transparency/>.
- Leverhulme Centre for the Future of Intelligence, 'Politics and Democracy' (2019b). <http://lcfi.ac.uk/projects/ai-trust-and-society/politics-and-democracy/>.
- Leyser, O. & Wiseman, H., 'Integrative Biology: Parts, Wholes, Levels and Systems'. Ch 2 of Reiss, M. J., Watts, F. & Wiseman, H. (eds.), *Rethinking Biology: Public Understandings* (New Jersey: World Scientific, 2019).
- Liberty, 'Privacy' (2019). <https://www.libertyhumanrights.org.uk/human-rights/privacy>.
- MacIntyre, A., *A Short History of Ethics: A history of moral philosophy from the Homeric Age to the twentieth century* (London: Routledge, 1966).

- MacIntyre, A., *Ethics in the Conflicts of Modernity: An Essay on Desire, Practical Reasoning, and Narrative* (Cambridge: Cambridge University Press, 2016).
- Madsen, J. K., *The Psychology of Micro-Targeted Election Campaigns* (Basingstoke: Palgrave Macmillan, 2019).
- Maslen, H., *Remorse, Penal Theory and Sentencing* (London: Hart Publishing, 2015).
- Manent, P., *A World beyond Politics? A Defense of the Nation-State* (Princeton: Princeton University Press, 2006).
- Mayer, C., *Prosperity: Better Business Makes the Greater Good* (Oxford: Oxford University Press, 2018).
- Metz, C. & Satariano, A., 'Facebook Restricts Live Streaming After New Zealand Shooting'. *New York Times* (14 May 2019).
<https://www.nytimes.com/2019/05/14/technology/facebook-live-violent-content.html>.
- Milbank, J. & Pabst, A., *The Politics of Virtue: Post-Liberalism and the Human Future* (London: Rowman & Littlefield, 2016).
- Mill, J. S., *On Liberty* (New York: Bedford, 2008 [1859]).
- Mischel, W. & Ebbesen, E. B., 'Attention in Delay of Gratification'. *Journal of Personality and Social Psychology*, Vol. 16, No. 2 (1970), pp. 329-337.
- Montesquieu, *The Spirit of the Laws* (Cambridge: Cambridge University Press, 1989 [1748]).
- Murphy, H., 'How Facebook could target ads in age of encryption'. *Financial Times* (27 Mar 2019). <https://www.ft.com/content/0181666a-4ad6-11e9-bbc9-6917dce3dc62>.
- Murphy, J. B., 'Nature, Custom, and Reason as the Explanatory and Practical Principles of Aristotelian Political Science'. *The Review of Politics*, Vol. 64, No. 3 (2002), pp. 469-495.
- Murungi, J., 'The Question of an African Jurisprudence: Some Hermeneutic Reflections'. Ch 43 of Wiredu, K. (ed.), *A Companion to African Philosophy* (Malden: Blackwell Publishing, 2004).
- Nash, K., *Contemporary Political Sociology: Globalization, Politics, and Power* (Chichester: Wiley-Blackwell, 2010), 2nd Ed.
- Newman, D., *Community and Collective Rights: A Theoretical Framework for Rights Held by Groups* (Oxford: Hart Publishing, 2011).
- Newport, C., *Deep Work: Rules for Focused Success in a Distracted World* (London: Piatkus, 2016).
- North, D. C., *Structure and Change in Economic History* (New York: W. W. Norton, 1981).
- North, D. C., *Institutions, Institutional Change and Economic Performance* (Cambridge: Cambridge University Press, 1990).

- North, D. C., 'Institutions'. *Journal of Economic Perspectives*, Vol. 5, No. 1 (1991), pp. 97-112.
- Nyabola, N., *Digital Democracy, Analogue Politics: How the Internet Era is Transforming Politics in Kenya* (London: Zed Books, 2018).
- Online Library of Liberty, 'Montesquieu thought that commerce improves manners and cures "the most destructive prejudices" (1748)' (2019). <https://oll.libertyfund.org/quotes/85>.
- Ostrom, E., *Understanding Institutional Diversity* (Princeton: Princeton University Press, 2005).
- Owen, T., *Disruptive Power: The Crisis of the State in the Digital Age* (Oxford: Oxford University Press, 2015).
- Packard, V., *The Hidden Persuaders* (London: Longmas, Green & Co., 1957).
- Papacharissi, Z., *Affective Publics: Sentiment, Technology, and Politics* (Oxford: Oxford University Press, 2015).
- Penney, J., *The Citizen Marketer* (Oxford: Oxford University Press, 2017).
- Pettit, P., *Republicanism: A Theory of Freedom and Government* (Oxford: Oxford University Press, 1997).
- Picard, R. W., *Affective Computing* (Cambridge, MA: MIT Press, 1997).
- Piketty, T., *Capital in the Twenty-First Century* (Cambridge, MA: Harvard University Press, 2014).
- Plato, *Complete Works* (Indianapolis: Hackett Publishing Company, 1997), edited by John M. Cooper.
- Points, L. & Potton, E., 'Artificial Intelligence and Automation in the UK'. *House of Commons Library, Briefing Paper*, No. 8152 (21 Dec 2017).
- Polanyi, K., *The Great Transformation: The Political and Economic Origins of Our Time* (Boston: Beacon Press, 1944).
- Posner, R. A., 'The Right of Privacy'. *Georgia Law Review*, Vol. 12, No. 3 (1978), pp. 393-422.
- Putnam, R. D., *Making Democracy Work* (Princeton: Princeton University Press, 1993).
- Putnam, R. D., *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon & Schuster, 2000).
- Quinn, M., 'One Difference Between 2009 vs 2018 Iran Protests? 48 Million Smartphones'. *Voice of America* (3 Jan 2018). <https://www.voanews.com/usa/one-difference-between-2009-vs-2018-iran-protests-48-million-smartphones>.
- Rai, S., 'Alarming Lessons From Facebook's Push to Stop Fake News in India'. *Bloomberg* (20 May 2019). <https://www.bloomberg.com/news/articles/2019-05-20/alarming-lessons-from-facebook-s-push-to-stop-fake-news-in-india>.

- Rouse, M., 'Definition: Turing Test'. *TechTarget* (2019).
<https://searchenterpriseai.techtarget.com/definition/Turing-test>.
- Rousseau, J. J., *The Social Contract* (Cambridge: Cambridge University Press, 1997 [1762]).
- Russell, S. J., *Human Compatible: Artificial Intelligence and the Problem of Control* (United Kingdom: Allen Lane, 2019).
- Russell, S. J. & Norvig, P., *Artificial Intelligence: A Modern Approach* (Harlow: Pearson, 2016), 3rd Ed.
- Safiya Umoja, N., *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018).
- Sandel, M. J., *What Money Can't Buy: The Moral Limits of Markets*, (London: Penguin, 2012).
- Sanger, D. E. & Chen, B. X., 'Signaling Post-Snowden Era, New iPhone Locks Out N.S.A.' *New York Times* (26 Sep 2014).
<https://www.nytimes.com/2014/09/27/technology/iphone-locks-out-the-nsa-signaling-a-post-snowden-era-.html>.
- Schmidt, E. & Cohen, J., *The New Digital Age: Reshaping the Future of People, Nations and Business* (London: John Murray, 2013).
- Scott, J. C., *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* (New Haven: Yale University Press, 1998).
- Selinger, E. & Seager, T. P., 'Digital Jiminy Crickets: Do apps that promote ethical behaviour diminish our ability to make just decisions?' *Slate* (13 Jul 2012). <https://slate.com/technology/2012/07/ethical-decision-making-apps-damage-our-ability-to-make-moral-choices.html>.
- Selyukh, A., 'A Year After San Bernardino And Apple-FBI, Where Are We On Encryption?' *National Public Radio* (3 Dec 2016).
<https://www.npr.org/sections/alltechconsidered/2016/12/03/504130977/a-year-after-san-bernardino-and-apple-fbi-where-are-we-on-encryption>.
- Sen, A., 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory'. *Philosophy & Public Affairs*, Vol. 6, No. 4 (1977), pp. 317-344.
- Sidharth, A., 'Viral: Fake news of march by Muslims demanding release of Mandsaur rape accused'. *Alt News* (2 Jul 2018).
<https://www.altnews.in/viral-fake-news-of-march-by-muslims-demanding-release-of-mandsaur-rape-accused/>.
- Siedentop, L., *Inventing the Individual: The Origins of Western Liberalism* (London: Penguin Books, 2015).
- Sigalet, G., 'Waldron's Challenge to Aristotelians: On the Political Relevance of Moral Realism'. *Politics & Poetics*, Vol. 4 (2018), pp. 1-23.
- Singh, S., *The Code Book: The Secret History of Codes & Code-breaking* (London: Fourth Estate, 1999).

- Smith, A., *The Wealth of Nations* (New York: The Modern Library, 2000 [1776]).
- Smith, C., *What is a Person? Rethinking Humanity, Social Life, and the Moral Good from the Person Up* (Chicago: The University of Chicago Press, 2010).
- Srivastava, M., 'WhatsApp voice calls used to inject Israeli spyware on phones'. *Financial Times* (13 May 2019).
<https://www.ft.com/content/4da1117e-756c-11e9-be7d-6d846537acab>.
- Steane, A., *Faithful to Science: The Role of Science in Religion* (Oxford: Oxford University Press, 2014).
- Stephens-Davidowitz, S., *Everybody Lies: What the Internet Can Tell Us About Who We Really Are* (London: Bloomsbury, 2017).
- Stewart, C., *Driving Over Lemons: An Optimist in Andalucía* (London: Sort of Books, 2011).
- Taneja, H., 'The Era of "Move Fast and Break Things" Is Over'. *Harvard Business Review* (22 Jan 2019). <https://hbr.org/2019/01/the-era-of-move-fast-and-break-things-is-over>.
- Taylor, C., *Sources of the Self: The Making of the Modern Identity* (Cambridge: Cambridge University Press, 1989).
- Taylor, C., *A Secular Age* (Cambridge, MA: Belknap Press of Harvard University Press, 2007).
- Techworld, 'Best secure mobile messaging apps' (4 Jan 2019).
<https://www.techworld.com/security/best-secure-mobile-messaging-apps-3629914/>.
- Tegmark, M., *Life 3.0: being human in the age of artificial intelligence* (London: Allen Lane, 2017).
- Templeton, J., *Wisdom from World Religions: Pathways toward Heaven on Earth* (Philadelphia: Templeton Foundation Press, 2002).
- Templeton World Charity Foundation, 'Diverse Intelligences' (2018).
<https://www.templetonworldcharity.org/our-work/diverse-intelligences>.
- The Royal Society, *iHuman: Blurring lines between mind and machine* (London, Sep 2019).
- The Telegraph, 'Best of Matt' (28 Dec 2018).
<https://www.telegraph.co.uk/news/2016/03/31/best-of-matt-cartoon15>.
- Tocqueville, A., *Democracy in America* (London: Penguin Books, 2003 [1835]).
- Udrescu, S. & Tegmark, M., 'AI Feynman: a Physics-Inspired Method for Symbolic Regression' (27 May 2019). arXiv:1905.11481.
- Urbina, F. J., *A Critique of Proportionality and Balancing* (Cambridge: Cambridge University Press, 2017).

- Uslaner, E. M., *The Moral Foundations of Trust* (Cambridge: Cambridge University Press, 2002).
- Vallor, S., *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (New York: Oxford University Press, 2016).
- VanderWeele, T. J., 'On the promotion of human flourishing'. *PNAS*, Vol. 114, No. 31 (2017), pp. 8148-8156.
- VanderWeele, T. J., 'Measures of Community Well-Being: a Template'. *International Journal of Community Well-Being*, Vol. 2 (2019), pp. 253-275.
- Varma, T., 'The WhatsApp election: The fake news challenge to democracy'. *European Council on Foreign Relations* (30 Apr 2019).
https://www.ecfr.eu/article/commentary_the_whatsapp_election_the_fake_news_challenge_to_democracy.
- Vengattil, M., Kalra, A. & Phartiyal, S., 'In India election, a \$14 software tool helps overcome WhatsApp controls'. *Reuters* (15 May 2019).
<https://www.reuters.com/article/india-election-socialmedia-whatsapp/in-india-election-a-14-software-tool-helps-overcome-whatsapp-controls-idUSKCN1SLOPZ>.
- Vinge, V., 'The Coming Technological Singularity'. *Whole Earth Review* (Winter 1993).
- Volf, M., *The End of Memory: Remembering Rightly in a Violent World* (Cambridge: W. B. Eerdmans, 2006).
- Wachter, S., Mittelstadt, B. & Floridi, L., 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation'. *International Data Privacy Law*, Vol. 7, No. 2 (2017), pp. 76-99.
- Waterson, J., 'UK online pornography age block triggers privacy fears'. *The Guardian* (16 Mar 2019).
<https://www.theguardian.com/culture/2019/mar/16/uk-online-porn-age-verification-launch>.
- Weinshall-Margel, K. & Shapard, J., 'Overlooked factors in the analysis of parole decisions'. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 108, No. 42 (2011), p. E833.
- Weller, A., 'Where is AI going, and how will it promote flourishing?' Presentation at the conference 'Science, Philosophy, Religion & Human Flourishing', Ayia Napa, Cyprus, (Nov 2018).
- WhatsApp, 'WhatsApp Encryption Overview: Technical white paper' (19 Dec 2017). <https://www.whatsapp.com/security/WhatsApp-Security-Whitepaper.pdf>.
- Wilmer, H. H., Sherman, L. E. & Chein, J. M., 'Smartphones and Cognition: A Review of Research Exploring the Links between Mobile Technology Habits and Cognitive Functioning'. *Frontiers in Psychology*, Vol. 8, No. 605 (2017).

- Woolley, S. C. & Howard, P. N. (eds.), *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (Oxford: Oxford University Press, 2018).
- Zittrain, J., 'The Hidden Costs of Automated Thinking'. *The New Yorker* (23 Jul 2019). <https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking>.
- Zuboff, S., *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power* (London: Profile Books, 2019).
- Zuckerman, E., *Rewire: Digital Cosmopolitans in the Age of Connection* (New York: W. W. Norton, 2013).

Acknowledgements

Dr Anna Alexandrova, University of Cambridge
Professor Jon Askonas, Catholic University of America
Professor Jesse Baldwin-Philippi, Fordham University
Dr Paul Billingham, University of Oxford
Professor Sarah-Jayne Blakemore, University College London
Michael Bloom, Now You Know Media
Professor Max Bonilla, University Francisco de Vitoria
Dr Jonathan Brant, University of Oxford
Dr Michael Burdett, University of Oxford
Sir Keith Burnett, University of Sheffield
Tara Isabella Burton
Dr Oliver Butler, University of Oxford
Dr Daniel Cheely, Collegium Institute
Professor Sarah Coakley, University of Cambridge
Professor Paul Davies, Arizona State University
Joyner D'Souza, University of Oxford
Professor Richard Ekins, University of Oxford
Professor Maura Jane Farrelly, Brandeis University
Professor Kevin Flannery, Pontifical Gregorian University
Professor David Ford, University of Cambridge
Professor Angela Franks, St John's Seminary
Dr Fiona Gatty, University of Oxford
Professor Nidhal Guessoum, American University of Sharjah
Professor Scott Hawley, Belmont University
Donald Hay, University of Oxford
Professor Rolf Heuer, CERN
Arthur Holland Michel, Bard College
Dr Peter Jordan, University of Oxford
Dr Samuel Kimbriel, Catholic University of America
Dr Jens Koed Madsen, University of Oxford
Dr Reid Maclellan, Harvard University

Nicola Macmichael, University of Oxford
Abrania Marrero, Harvard University
Dr J. Nathan Matias, Cornell University
Professor Tom McLeish, University of York
Dr Zeeya Merali
George Monbiot
Professor James Bernard Murphy, Dartmouth College
Professor Dave Myers, Hope College
Professor James Nolan, Williams College
Dr David O'Brien, Harvard University
Professor Zizi Papacharissi, University of Illinois-Chicago
Dr Danilo Petranovich, Abigail Adams Institute
Dr Dawid Potgieter, Templeton World Charity Foundation
Professor Lord Martin Rees, University of Cambridge
Dr Michael Sacasas, Christian Study Center of Gainesville
Dr Andrew Serazin, Templeton World Charity Foundation
Richard Sergay
Dr Afonso Seixas-Nunes, University of Oxford
Professor Tom Simpson, University of Oxford
Dr Beth Singler, University of Cambridge
Dr Bethany Sollereder, University of Oxford
Kippi Rai-Spraggon, University of Oxford
Dr Syman Stevens, Templeton World Charity Foundation
Professor Giuseppe Tanzella-Nitti, Pontifical University of the Holy Cross
Kristen VanBlargan, Now You Know Media
Professor Tyler J. VanderWeele, Harvard University
Professor Miroslav Volf, Yale University
Professor Graham Ward, University of Oxford
Michael Wear, Public Square Strategies
Archbishop Justin Welby
Dr Matthew F. Wilson, Harvard University
Dr Sam Wren-Lewis
Dr Bonnie Zahl, University of Oxford

External reviewers

The following were chosen for their expertise to act as external reviewers of a near-final version of the report. We thank them for the care and thought which they took to read the report, and for their valuable feedback.

Professor Sir Paul Collier, CBE, FBA

Blavatnik School of Government, University of Oxford

The Rt Revd Dr Steven Croft

Bishop of Oxford

Sir Andrew Dilnot, CBE

Warden, Nuffield College, University of Oxford

Professor Baroness Onora O'Neill of Bengarve, CH, CBE, FRS, FBA

House of Lords

Professor Weixing Shen

School of Law, Tsinghua University

Professor Nancy E. Snow

Director, Institute for the Study of Human Flourishing, The University of Oklahoma

Professor Lionel Tarassenko, CBE, FREng

President, Parks College, University of Oxford

Roger Taylor

Chair, Centre for Data Ethics and Innovation

'An impressive piece of work on an increasingly vital topic.'

Professor Mary Ann Glendon, Harvard Law School,
Chair, US Commission on Unalienable Rights

**'A stimulating and thought provoking read in an important and
developing area.'**

Dr Steven Croft, Bishop of Oxford

'There is much to engage with in this enormously rich document.'

Roger Taylor, Chair, UK Centre for Data Ethics and Innovation

**'The authors of "Citizenship in a Networked Age" take on the pressing
question of how to participate in democratic institutions as a dizzying
array of digital developments alter the fabric of our society. Answering
this is no easy task, but they do so with thoughtful arguments and
meticulous research on the past, present, and future. Readers will find
in its pages an abiding hope in humanity's ability to flourish with
technology, not in spite of it.'**

Kristen VanBlargan, Now You Know Media

**'A research report with a broad perspective and full of philosophical
thinking ... a profound analysis of the current status and future goals of
civic life in the internet age.'**

Professor Weixing Shen, Dean, School of Law, Tsinghua University

**'Commercially-driven new technology has plunged us into a networked
world that has huge implications for our societies. Public policy has
been slow to react and people are understandably anxious. We urgently
need advice based on scientific knowledge and impartial judgment, and
with this Report we have it. Both policy makers and concerned citizens
will find it invaluable.'**

Professor Sir Paul Collier, Oxford Blavatnik School of Government,
author of *The Future of Capitalism*

Citizenship in a Networked Age



TEMPLETON WORLD
CHARITY FOUNDATION

