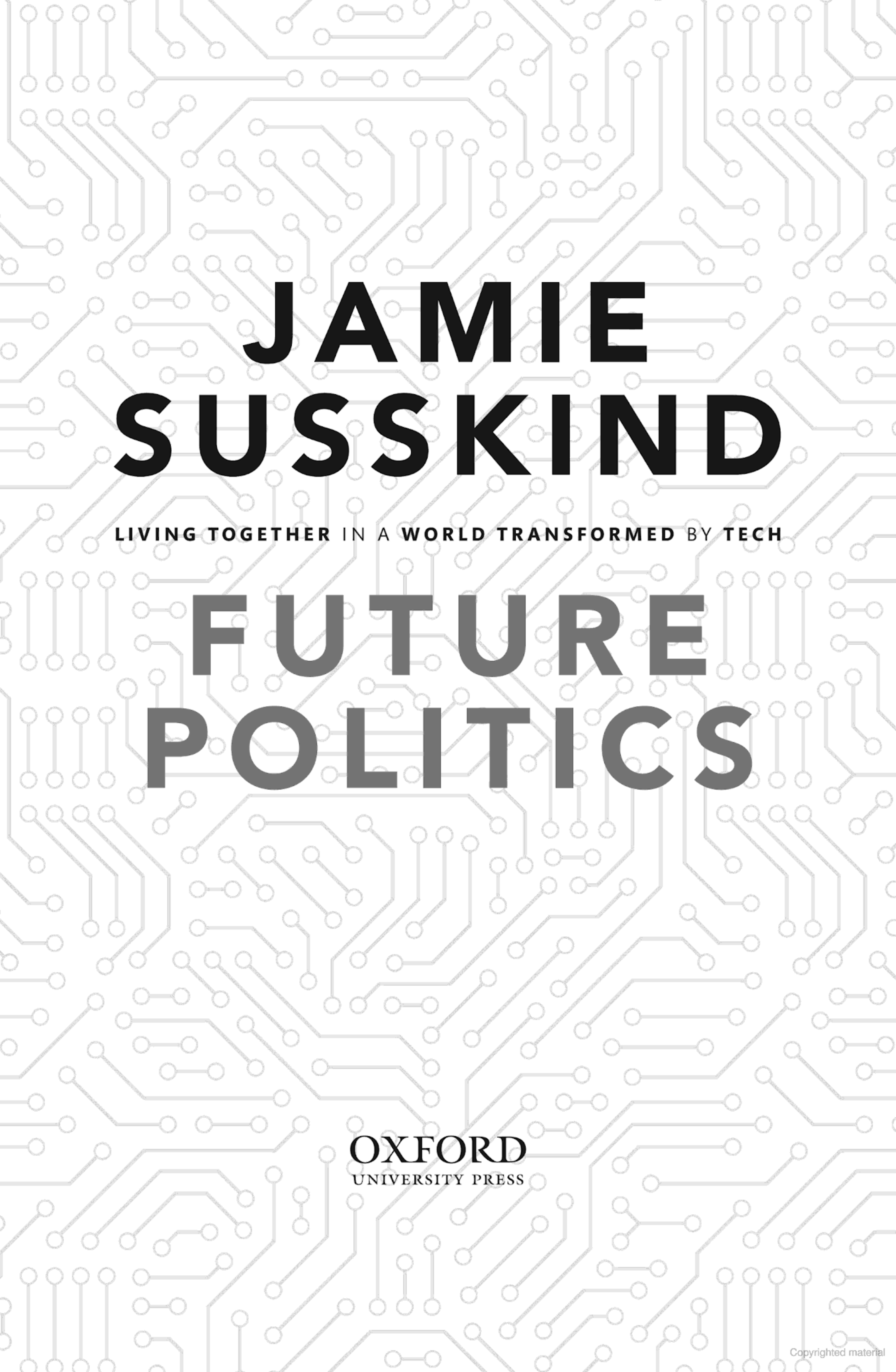




JAMIE SUSSKIND

LIVING TOGETHER IN A WORLD TRANSFORMED BY TECH

FUTURE POLITICS



**JAMIE
SUSSKIND**

LIVING TOGETHER IN A WORLD TRANSFORMED BY TECH

**FUTURE
POLITICS**

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Jamie Susskind 2018

The moral rights of the author have been asserted

First Edition published in 2018

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2018933401

ISBN 978-0-19-882561-6

Printed in Great Britain by
Clays Ltd, St Ives plc

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

COPYRIGHT NOTIFICATIONS

Chapter 2 epigraph: From *Technics and Civilization*, Lewis Mumford (1934). Copyright © 1934 by Elizabeth M. Morss and James G. Morss. Used by permission. Courtesy of the estate of Lewis and Sophia Mumford.

Chapter 4 epigraph: From *Wind, Sand, & Stars*, Antoine de Saint-Exupéry (1939). Reproduced with permission by the Antoine de Saint-Exupéry Estate.

Chapter 7 epigraph: From *Nineteen Eighty-Four*, George Orwell (1949). Copyright © George Orwell, 1949. Reprinted by permission of Bill Hamilton as the Literary Executor of the Estate of the Late Sonia Brownell Orwell.

Chapter 9 epigraph: From *Technics and Civilization*, Lewis Mumford (1934). Copyright © 1934 by Elizabeth M. Morss and James G. Morss. Used by permission. Courtesy of the estate of Lewis and Sophia Mumford.

Chapter 16 epigraph: From *Tragic Choices: the Conflicts Society Confronts in the Allocation of Tragically Scarce Resources* by Guido Calabresi and Philip Bobbitt (1978). Copyright © 1978 by the Fels Center of Government. Used by permission of W.W. Norton & Company, Inc.

Every effort has been made to trace and contact copyright holders prior to publication. If notified, the publisher will be pleased to rectify any errors or omissions at the earliest opportunity.

CONTENTS

Introduction	1
--------------	---

PART I THE DIGITAL LIFEWORLD

1. Increasingly Capable Systems	29
2. Increasingly Integrated Technology	42
3. Increasingly Quantified Society	61
4. Thinking Like a Theorist	69

PART II FUTURE POWER

5. Code is Power	89
6. Force	100
7. Scrutiny	122
8. Perception-Control	142
9. Public and Private Power	153

PART III FUTURE LIBERTY

10. Freedom and the Supercharged State	163
11. Freedom and the Tech Firm	188

PART IV FUTURE DEMOCRACY

12. The Dream of Democracy 211
13. Democracy in the Future 227

PART V FUTURE JUSTICE

14. Algorithms of Distribution 257
15. Algorithms of Recognition 271
16. Algorithmic Injustice 279
17. Technological Unemployment 295
18. The Wealth Cyclone 313

PART VI FUTURE POLITICS

19. Transparency and the New Separation of Powers 345
20. Post-Politics 362

- Notes* 367
Bibliography 437
Index 491

'A new political science is needed for a world itself quite new'

Alexis de Tocqueville, *Democracy in America* (1835)

Introduction

The future stalks us. It is always waiting, barely out of sight, lurking around the corner or over the next rise. We can never be sure what form it will take. Often it catches us entirely unprepared.

Nowadays, many of us share the sense that we are approaching a time of great upheaval. The world seems to be changing faster than we can grasp. Often we struggle to explain political events that would have been unimaginable just a few years ago. Sometimes we don't even have the words to describe them. Inwardly, we know that this is just the beginning.

The premise of this book is that relentless advances in science and technology are set to transform the way we live together, with consequences for politics that are profound and frightening in equal measure. We are not yet ready—intellectually, philosophically, or morally—for the world we are creating. In the next few decades, old ways of thinking that have served us well for hundreds, even thousands, of years, will be called into question. New debates, controversies, movements, and ideologies will come to the fore. Some of our most deeply held assumptions will be revised or abandoned altogether. Together we will need to re-imagine what it means to be free or equal, what it means to have power or property, and even what it means for a political system to be democratic. Politics in the future will be quite unlike politics in the past.

Politics in the twentieth century was dominated by a central question: how much of our collective life should be determined by the state, and what should be left to the market and civil society? For the generation now approaching political maturity, the debate will be different: to what extent should our lives be directed and controlled by powerful digital systems—and on what terms? This question is at the heart of *Future Politics*.

In the next few decades, it is predicted, we'll develop computing systems of astonishing capability, some of which will rival and surpass humans across a wide range of functions, even without achieving an 'intelligence' like ours. Before long, these systems will cease to resemble computers. They'll be embedded in the physical world, hidden in structures and objects that we never used to regard as technology. More and more information about human beings—what we do, where we go, what we think, what we say, how we feel—will be captured and recorded as data, then sorted, stored, and processed digitally. In the long run, the distinctions between human and machine, online and offline, virtual and real, will fade into the background.

This transformation will bring some great benefits for civilization. Our lives will be enriched by new ways of playing, working, travelling, shopping, learning, creating, expressing ourselves, staying in touch, meeting strangers, coordinating action, keeping fit, and finding meaning. In the long run, we may be able to augment our minds and bodies beyond recognition, freeing ourselves from the limitations of our human biology.

At the same time, however, some technologies will come to hold great power over us. Some will be able to force us to behave a certain way, like (to take a basic example) self-driving vehicles that simply refuse to drive over the speed limit. Others will be powerful because of the information they gather about us. Merely knowing we are being watched makes us less likely to do things perceived as

shameful, sinful, or wrong. Still other technologies will filter what we see of the world, prescribing what we know, shaping the way we think, influencing how we feel, and thereby determining how we act.

Those who control these technologies will increasingly control the rest of us. They'll have *power*, meaning they'll have a stable and wide-ranging capacity to get us to do things of significance that we wouldn't otherwise do. Increasingly, they'll set the limits of our *liberty*, decreeing what may be done and what is forbidden. They'll determine the future of *democracy*, causing it to flourish or decay. And their algorithms will decide vital questions of *social justice*, allocating social goods and sorting us into hierarchies of status and esteem.

The upshot is that political authorities—generally states—will have more instruments of control at their disposal than ever before, and big tech firms will also come to enjoy power on a scale that dwarfs any other economic entity in modern times. To cope with these new challenges, we'll need a radical upgrade of our political ideas. The great English philosopher John Stuart Mill wrote in his *Autobiography* of 1873 that, 'no great improvements in the lot of mankind are possible, until a great change takes place in the fundamental constitution of their modes of thought.'¹

It is time for the next great change.

The Next Great Change

We already live in a time of deep political unease. Every day the news is of bloody civil war, mass displacement of peoples, ethnic nationalism, sectarian violence, religious extremism, climate change, economic turbulence, disorienting globalization, rising inequality, and an array of other challenges too dismal to mention. It seems like the world isn't in great shape—and that our public discourse has sunk to the occasion. Political élites are widely distrusted

and despised. Two recent exercises in mass democracy in the English-speaking world, the 2016 US presidential election and UK Brexit referendum, were rancorous even by the usual unhappy standards, with opposing factions vying not just to defeat their rivals but to destroy them. Both were won by the side that promised to tear down the old order. Neither brought closure or satisfaction. Increasingly, as Barack Obama noted at the end of his presidency, ‘everything is true, and nothing is true’.² It’s getting harder for ordinary citizens (of any political allegiance) to separate fact from fraud, reality from rumour, signal from noise. Many have given up trying. The temptation is to hunker down and weather the present storm without thinking too hard about the future.

That would be a mistake.

If mainstream predictions about the future of technology are close to the mark, then the transformation on the horizon could be at least as important for humankind as the industrial revolution, the agricultural revolution, or even the invention of language. Many of today’s problems will be dwarfed by comparison. Think about the effect that technology has already had on our lives—how we work, communicate, treat our illnesses, exercise, eat, study, and socialize—and then remember that in historical perspective, the digital age is only a few seconds old. Fully 99.5 per cent of human existence was spent in the Palaeolithic era, which began about 3 million years ago when humans began using primitive tools. That era ended about 12,000 years ago with the last ice age.³ During this long twilight period, people noticed almost no cultural change at all. ‘The human world that individuals entered at birth was the same as the one they left at death’.⁴ If you consider that the earliest human civilizations emerged some 5,000 years ago, then the seventy or so years that we have lived with modern computing machines, the thirty or so we have had the world wide web, and the decade we’ve spent with smartphones don’t seem very long at all. And while time passes linearly, many developments in

digital technology are occurring exponentially, the rate of change accelerating with each passing year.

We have no evidence from the future, so trying to predict it is inherently risky and difficult. I admire those who try to do so in a rigorous way, and I have borrowed extensively from their work in this book. But to be realistic, we should start by acknowledging that such predictions often badly miss the mark. Much of the future anticipated in these pages will probably never come to pass, and other developments, utterly unforeseen, will emerge to surprise us instead. That said, I believe it is possible to make sensible, informed guesses about what the future might look like, based on what we know of the current trends in science, technology, and politics. The biggest risk would be not to try to anticipate the future at all.

The story is told of an encounter between the Victorian statesman William Gladstone and the pioneering scientist Michael Faraday. Faraday was trying to explain his groundbreaking work on electricity to Gladstone, but Gladstone seemed unimpressed. ‘But what use is it?’ he asked, with growing frustration; ‘What *use* is it?’

‘Why sir,’ replied Faraday, reaching the end of his patience, ‘there is every possibility that you will soon be able to tax it.’

Many innovators, like Faraday, find it hard to explain the social and practical implications of their work. And the rest of us, like Gladstone, are too often dismissive of technologies we don’t yet understand. It can be hard to see the political significance of inventions that, at first glance, seem to have nothing to do with politics. When confronted with a new gadget or app, we tend not to think first of all about its implications for the political system. Instead we want to know: what does it do? How much does it cost? Where can I get one? This isn’t surprising. In general, technology is something we encounter most often as consumers. But this rather narrow attitude now needs to change. We must apply the same scrutiny and scepticism to the new technologies of power that we have

always brought to powerful politicians. Technology affects us not just as consumers but as citizens. In the twenty-first century, *the digital is political*.

This book is partly for Gladstones who want to understand more about technology and partly for Faradays who want to see more clearly the political significance of their work. But mainly it's for ordinary citizens who want to understand the future a bit better—so if nothing else they can hold the Gladstones and the Faradays to account.

Philosophical Engineers

Consider the following passage:

Here's to the crazy ones. The misfits. The rebels. The troublemakers. The round pegs in the square holes. The ones who see things differently. They're not fond of rules. And they have no respect for the status quo. You can quote them, disagree with them, glorify or vilify them. About the only thing you can't do is ignore them. Because they change things. They push the human race forward. And while some may see them as the crazy ones, we see genius. Because the people who are crazy enough to think they can change the world are the ones who do.

These are not the words of a politician. They're from the voiceover to 'Think Different', a 1997 Apple advertisement featuring iconic footage of rebels including Mahatma Ghandi and Martin Luther King. The ad embodies a worldview, widely held among tech entrepreneurs, that their work is of philosophical as well as commercial importance. 'It is commonplace in Silicon Valley,' explains Jaron Lanier, 'for very young people with a startup in a garage to announce that their goal is to change human culture globally and profoundly, within a few years, and that they aren't ready yet to worry about money, because acquiring a great fortune is a petty

matter that will take care of itself.’⁵ There is something attractive about this way of thinking, partly because it suggests that tech companies might not be as rapacious as they are sometimes made out to be. And the basic premise is right: digital technologies *do* indeed have an astounding capacity to change the world. Compare the following statements:

‘The philosophers have only *interpreted* the world in various ways; the point is to *change* it.’

‘We are not *analysing* a world, we are *building* it.’

The first is from Karl Marx’s 1845 *Theses on Feuerbach*.⁶ It served as a rallying cry for political revolutionaries for more than a century after its publication. The second is from Tim Berners-Lee, the mild-mannered inventor of the World Wide Web.⁷ Marx and Berners-Lee could scarcely be more different in their politics, temperament, or choice of facial hair. But what they have in common—in addition to having changed the course of human history—is a belief in the distinction between *making* change and merely *thinking* about it or *studying* it. On this view, far from being a spectral presence out of our control, the future is something we design and build.

‘We are not experimental philosophers,’ says Berners-Lee, ‘we are philosophical engineers.’⁸ It’s a practical and hands-on way of looking at life, one more familiar to builders and inventors than to tweedy academics or beturtlenecked philosophers. It also happens to be the defining mindset of our age. Today, the most important revolutions are taking place not in philosophy departments, nor even in parliaments and city squares, but in laboratories, research facilities, tech firms, and data centres. Most involve developments in digital technology. Yet these extraordinary advances are taking place in a climate of alarming cultural and intellectual isolation. With a few exceptions, there is a gulf between the arts

and the sciences. Political philosophy and social policy rarely appear in degree programmes for science, technology, engineering, and mathematics. And if you ask the average liberal arts student how a computer works, you are unlikely to get a sophisticated response.

In tech firms themselves, few engineers are tasked with thinking hard about the systemic consequences of their work. Most are given discrete technical problems to solve. Innovation in the tech sector is ultimately driven by profit, even if investors are prepared to take a ‘good idea first, profits later’ approach. This is not a criticism: it’s just that there’s no reason why making money and improving the world will always be the same thing. In fact, as many of the examples in this book show, there’s plenty of evidence to suggest that digital technology is too often designed from the perspective of the powerful and privileged.

As time goes on, we will need more philosophical engineers worthy of the name. And it will become even more important for the rest of us to engage critically with the work of tech firms, not least because tech working culture is notorious for its lack of diversity. Roughly nine out of every ten Silicon Valley executives are men.⁹ Despite the fact that African-Americans make up about 10 per cent of computer science graduates and 14 per cent of the overall workforce, they make up less than 3 per cent of computing roles in Silicon Valley.¹⁰ And many in the tech community hold strong political views that are way outside the mainstream. More than 44 per cent of Bitcoin adopters in 2013, for instance, professed to be ‘libertarian or anarcho-capitalists who favour elimination of the state’.¹¹

As I will argue, we put so much at risk when we delegate matters of political importance to the tiny group that happens to be tasked with developing digital technologies at a given time. That’s true whether you admire the philosophical engineers of Silicon Valley or you think that most ‘tech bros’ have the political sophistication

of a transistor. We need an intellectual framework that can help us to think clearly and critically about the political consequences of digital innovation. This book hopes to contribute to such a framework, using the ideas and methods of political theory.

The Promise of Political Theory

The purpose of philosophy, says Isaiah Berlin, is always the same: to assist humans ‘to understand themselves and thus operate in the open, and not wildly, in the dark.’¹² That’s our goal too. Political theory aims to understand politics through the concepts we use to speak about it.¹³ What is *power*? When should *freedom* be curtailed and on what basis? Does *democracy* require that everyone has an *equal* ability to shape the political process? What is a *just* distribution of society’s resources? These are the sorts of questions that political theorists try to answer. The discipline has a long and rich history. From Plato and Aristotle in the academies of ancient Greece to Thomas Hobbes and Jean-Jacques Rousseau in the tumult of early modern Europe, to the giants of twentieth-century political thought like Hannah Arendt and John Rawls, western political thinkers have long tried to clarify and critique the world around them, asking why it is the way it is—and whether it could or should be different.

For several reasons, political theory is well-suited to examining the interplay of technology and politics. First, the canon of political thought contains wisdom that has outlived civilizations. It can shed light on our future predicaments and help us to identify what’s at stake. We’d be foolish not to plunder the trove of ideas already available to us, even if we ultimately decide that some of those ideas need an upgrade or a reboot. Political theory also offers *methods* of thinking about the world that help us to raise the level of debate above assertion and prejudice.

To my mind, the best thing about political theory is that it deals with the *big* themes and questions of politics. It offers a panoramic view of the political forest where other approaches might get lost in the trees (or stuck in the branches). That's necessary, in our case, to do justice to the subject-matter. If we think that technology could have a fundamental impact on the human condition, then our analysis of that impact should be fundamental too. That's why this book is about four of the most basic political concepts of all:

Power: How the strong dominate the weak

Liberty: What is allowed and what is prohibited

Democracy: How the people can rule

Social Justice: What duties we owe to each other

In a time of great change, I suggest, it pays to go back to first principles and think about these concepts quite apart from any particular legal regime. That way we might be able to imagine a superior system to the one we have inherited.

Political theory is also useful because it allows us to think critically not just about politics but also about how we *think* and *speak* about politics. Concepts are the 'keyholes through which we inevitably see and perceive reality'.¹⁴ When I want to say something to my neighbour about politics, I don't need to start from scratch. I know that if I say a process is 'undemocratic' then she will have a pretty good idea of what I mean and the connotations I wish to convey, without any need for me to explain what democracy is and why it should be considered a good thing. That's because we are members of the same linguistic community, sharing a 'common stock of concepts' drawn from our shared history and mythology.¹⁵ It's convenient.

On the other hand, what we want to say about politics can sometimes be limited by the poverty of the words at our disposal.

Some things seem unsayable, or unthinkable, because the common stock of concepts hasn't yet developed to articulate them. 'The limits of my language,' says Ludwig Wittgenstein, 'mean the limits of my world.'¹⁶

What this means in political terms is that even if we could see the future clearly, we might not have the words to describe it. It's why, so often, we limit our vision of the future to a turbo-charged version of the world we already live in. 'If I had asked people what they wanted,' said Henry Ford, the first mass-producer of automobiles, 'they would have said faster horses.' Ford recognized that it can be hard to conceive of a system radically different from our own. Failure to keep our language up-to-date only makes it harder.

Why I Wrote this Book

My first real taste of political theory was as at university, where I fell in love with the discipline under the watchful eye of some wonderful professors. It sparked an obsession that has stayed with me since. (I acknowledge in hindsight that my most successful undergraduate romance may well have been a passionate but doomed affair with the German philosopher G. W. F. Hegel.)

Passion aside, something troubled me about the discipline of political theory. Political theorists seemed to pride themselves on thinking deeply about the *history* of political ideas but, with some exceptions, were almost entirely uninterested in their *future*. I found this strange: why would the same scholars—so sensitive to context when writing about the past—discuss politics as if the world will be the same in 2050 as it was in 1950? It seemed that a good deal of very clever political theory was of little practical application because it did not engage with the emerging realities of our time. When I thought about politics in the future, I thought about

Orwell, Huxley, Wells—all novelists from the early twentieth century rather than theorists from the twenty-first. It turns out that I wasn't alone: since the election of Donald Trump to the US presidency in late 2016, Orwell's *Nineteen Eighty-Four* has surged up the best seller lists. But this prompts the question: if we want to understand the world as it will be in 2050, should we really have to rely on a work of fiction from 1949?

After I left university and became involved in my own modest political causes, my niggling sense of unease—that political theory might be unable, or unwilling, to address the looming challenges of my generation—became a more urgent concern. What if developments in technology were to happen so fast that we lacked the intellectual apparatus to make sense of them? What if, unthinkingly, we were about to unleash a future that we couldn't understand, let alone control?

I wanted answers, and that's why I began working on this book.

The Politics of Technology

Before ploughing on, let's begin with a simple question: what is the connection between digital technology and politics?

Technology in General

New technologies make it *possible* to do things that previously couldn't be done; and they make it *easier* to do some things we could already do.¹⁷ This is their basic social significance. More often than not, the new opportunities created by technology are minor in nature: an ingenious new way of grinding coffee beans, for instance, is unlikely to lead to the overthrow of the state. But sometimes the consequences can be profound. In the industrial revolution, the

invention of power looms and stocking and spinning frames threatened to displace the jobs of skilled textile workers. Some of those workers, known as the Luddites, launched a violent rampage through the English countryside, destroying the new machines as they went. We still use the word Luddite to describe those who resist the arrival of new and disruptive technologies.

The *economic* consequences of innovation, as in the case of the Luddites, will often require a political response. But new technologies can raise *moral* challenges too. A few years from now, for instance, virtual reality systems will enable behaviour that was previously the stuff of science fiction, including (for instance) the possibility of virtual sex using platforms capable of simulating human sexual activity. This raises some interesting questions. Should it be legally permissible to have virtual sex with a digital partner designed to look like a real-life person—a celebrity, say—without that person’s knowledge or consent? People are likely to have strong feelings about this. What about having virtual ‘sex’ with a virtual incarnation of your best friend’s husband? It would seem wrong for the law to allow this, but you might argue that what you get up to in the privacy of your virtual world, without harming anybody, is no business of anyone else, let alone the law. To take a more extreme example: what about having virtual sex with an avatar of a child, in circumstances where no child is actually harmed in the creation of that experience?

These are new questions. We’ve not had to answer them before. Maybe you’ve already formed views on them. Perhaps those views differ from mine. And that’s precisely the point: these are inherently *political* questions, and they ought to be answered carefully by reference to some acceptable set of moral principles. As one distinguished author has noted, new technologies can lead us to look again at our political views, just as a new dish on the menu of our favourite restaurant can lead us to challenge our taste in food.¹⁸

Some technologies are ‘inherently political’ in that they actually require ‘particular kinds of political relationships’ or are at least strongly compatible with them.¹⁹ Langdon Winner, writing in 1980, cited nuclear power as an example: if you are going to adopt a system of nuclear power, you must also accept that a powerful ‘techno-scientific-industrial-military elite’ will be required to administer and supervise the plants and infrastructure.²⁰

Then there are other technologies that are not inherently political, but which are made political by their context. In the United Kingdom, for instance, a licensed firearm is generally a technology for the hunting of wild animals. Guns are not part of mainstream culture and people are mostly happy with the idea of strict regulation. But in the United States, the right to bear arms is guaranteed by the Second Amendment. Cultural opposition to regulation is much stiffer. Same technology, different political context.

A final, more subtle, connection between technology and politics is that our inventions have a way of inveigling themselves into our political and intellectual life. A good example is the mechanical clock. As Otto Mayr explains, some ancient civilizations imagined the state as being like a human body, with individual human members of the political community forming appendages like a ‘hand’ or a ‘foot’.²¹ In the late Renaissance, the metaphor of the body was joined by others like the ‘ship of state’. After Copernicus, the monarch came to be seen as a great sun around which his subjects revolved.²² In the sixteenth to eighteenth centuries, the dominant metaphor was that of the clock, an ingenious contraption that commanded ‘unprecedented veneration’.²³ After a while, thinkers came to see politics from the perspective of clockwork.²⁴ Harmony, temperance, and regularity became the prevailing political ideals.²⁵ The people yearned for an ‘ever vigilant statesman-engineer’ who could ‘predict as well as repair all

potential troubles'. Thus it was that a particular technology and a set of political values went hand in hand, reaching a climax in the seventeenth century with its 'extraordinary production of clocks' and the 'conspicuous flourishing of the authoritarian conception of order.'²⁶

The twentieth-century political imagination was also coloured by technology. Even before the digital age, the prospect of all-powerful computing machines inspired a great deal of art and fiction. In *The Machine Stops* (1928) E. M. Forster portrayed a world in which humans were subordinated to The Machine, a global technological system that monitored and controlled every aspect of human existence.²⁷

Some writers have tried to pin down what they see as the ideology of our own time. Evgeny Morozov, for instance, has written of the 'Google Doctrine' ('the enthusiastic belief in the liberating power of technology accompanied by the irresistible urge to enlist Silicon Valley start-ups in the global fight for freedom');²⁸ 'cyber-utopianism' ('a naive belief in the emancipatory nature of online communication that rests on a stubborn refusal to acknowledge its downside');²⁹ and 'solutionism' ('Recasting all complex social situations either as neatly defined problems with definite, computable solutions or as transparent and self-evident processes that can be easily optimized—if only the right algorithms are in place').³⁰

My own view is that the technologies in question are too young for us to know what lasting imprint they will leave on our political thought. It took hundreds of years for the idea of the mechanical clock to soak into European political and intellectual life. How we choose to interpret the technologies of our time, and how they in turn shape our perception of the world, are matters that are yet to be determined.

Digital Technologies

So much for the general relationship between technology and politics. But this book is about *digital* technologies—what used to be called information and communication technologies, or ICTs. And it appears that these are not just inherently political; they are *hyper-political*. That's because they strike at the two most fundamental ingredients of political life: *communication* and *information*.

All political order is built on coordination, cooperation, or control. It's impossible to organize collective life without at least one of the three. And none of them is possible without some system for exchanging information, whether among ordinary people or between ruler and ruled.³¹ This is why *language* is so important. As James Farr puts it, without language politics would not only be 'indescribable' but 'impossible':³²

Emerging nations could not declare independence, leaders instruct partisans, citizens protest war, or courts sentence criminals. Neither could we criticise, plead, promise, argue, exhort, demand, negotiate, bargain, compromise, counsel, brief, debrief, advise nor consent. To imagine politics without these actions would be to imagine no recognisable politics at all.

Yuval Noah Harari elegantly observes that language played a crucial role in the earliest days of human politics. A small number of sounds and signs enabled our ancestors to produce an infinite number of sentences, each with a distinct meaning. This allowed them to talk about each other, about big, complex things in the world around them, and about things lacking any physical existence like myths and stories. Such lore still holds communities together today.³³

Writing may have begun as a way to reflect reality, 'but it gradually became a powerful way to reshape reality'.³⁴ The Domesday Book of 1086 was commissioned by William the Conqueror to discover the extent and value of people's property so it could be

more efficiently taxed. The contents of the Book were set in stone and could not be appealed. In 1179 a commentator wrote of this ‘strict and terrible last account’ that:³⁵

its sentence cannot be quashed or set aside with impunity. That is why we have called the book ‘the Book of Judgement’... because its decisions, like those of the Last Judgement, are unalterable.

All that mattered was what was written in the Book. Little has changed. As Harari puts it, ‘Anyone who has ever dealt with the tax authorities, the educational system or any other complex bureaucracy knows that the truth hardly matters. What’s written on your form is far more important.’³⁶

But language is not enough. Advanced political communities also need to be able to process large amounts of *information*—from poverty and economic growth rates to fertility, unemployment, and immigration figures. If you want to govern a nation, you must first know it.³⁷ The eighteenth-century revolutionaries who sought to replace the decaying *ancien régimes* understood this well. Revolution in France in 1789 was followed with an intense effort of standardization and rationalization, the introduction of a unified system of weights and measures, division of territory into *départements*, and the promulgation of the Civil Code.³⁸ Across the Atlantic, the founding fathers of the United States enshrined the need for a decennial census (an ‘Enumeration’) in the first article of the Constitution. Alexander Hamilton believed that the federal government ought to be ‘the center of information’, because that would make it best placed to ‘understand the extent and urgency of the dangers that threaten’.³⁹

The connection between information and politics is fundamental, and it has left its mark on our vocabulary. The word *statistics* comes from the eighteenth-century German term *Staatwissenschaft*, the ‘science of the state’ taught by university professors to princelings

of the Holy Roman Empire.⁴⁰ Indeed, the functional definition of statistics—‘to make a priori separate things hold together, thus lending reality and consistency to larger, more complex objects’⁴¹—is pretty much the same as the purpose of politics, one in numerical abstraction and the other in human reality. Originally, the English word *classified* had one principal meaning: the sorting of information into taxonomies. During the nineteenth century, as the British state grew in power and its empire grew in scale, *classified* came to hold its additional modern meaning: information held under the exclusive jurisdiction of the state.⁴² Even the word *control* has informational origins, deriving from the medieval Latin verb *contrarotulare*, which meant to compare something ‘against the rolls’, the cylinders of paper that served as official records.⁴³

In *Economy and Society* (1922), the most important sociological work of the early twentieth century, Max Weber heralded the ‘precision instrument’ of *bureaucracy* as the most advanced way of organizing human activity up to that point.⁴⁴ Weber’s view, vindicated by the next hundred years, was that a unified scheme for the systematic organization of information, premised on ‘[p]recision, speed, unambiguity, knowledge of the files, continuity, discretion ... unity’, would be the most effective system of political control.⁴⁵ This lesson was not lost on Franklin D. Roosevelt when he assumed the presidency of the United States in 1933. His famous programme of economic intervention was paired with a less glamorous but nonetheless vital effort to transform the federal government’s approach to statistics. The years 1935 to 1950 saw the emergence of the ‘three essential tools of social statistics and modern economics’: surveys based on representative samplings, national accounts, and computing.⁴⁶ If Hamilton, Weber, and Roosevelt were around today, they’d all be interested in the gargantuan amounts of data we now produce and the increasingly capable systems we use to process it.

Stepping back, it's now possible to advance a tentative hypothesis that can guide us throughout this book: that how we gather, store, analyse, and communicate our information—in essence how we organize it—is closely related to how we organize our politics. So when a society develops strange and different technologies for information and communication, we should expect political changes as well.

It doesn't appear to be a coincidence, for instance, that the first large-scale civilizations emerged at the same time as the invention of the technology of writing. The earliest written documents, cuneiform tablets first used by the Sumerians in the walled city of Uruk in around 3,500 BC, were wholly administrative in nature, recording taxes, laws, contracts, bills, debts, ownership, and other rudimentary aspects of political life. As James Gleick observes, these tablets 'not only recorded the commerce and the bureaucracy but, in the first place, made them possible'.⁴⁷ Other ancient civilizations mushroomed on the back of script, then the most powerful known method of gathering, storing, analysing, and communicating information. In *Empire and Communications* (1950), Harold Innis explains how the monarchies of Egypt and Persia, together with the Roman empire and the city-states of that time, were all 'essentially products of writing'.⁴⁸

Closer to our time, some of the earliest computing systems were actually developed so that governments could make more sense of the data they had gathered. Herman Hollerith, originally hired to work on the 1880 United States census, developed a system of punch-cards and a 'tabulating machine' to process census data, which he leased to the US government. Building on his earlier invention, Hollerith went on to found the International Business Machines Corporation, now better known as IBM.⁴⁹

Turning from the past to the future, we need to ask how transformative digital technologies—technologies of information and communication—will affect *our* system of politics.

That's the question at the heart of this book.

We know that technology's effects vary from place to place. The introduction of printing technology in China and Korea, for example, did not cause the same kind of transformation that followed the introduction of the Gutenberg press in Europe, where society was better primed for religious and political upheaval.⁵⁰ Differences like these can usually be explained by economic and political circumstances. Who owns and controls the given technology, how it's received by the public, whether its possible uses are contemplated in advance, and whether it is directed toward a particular end, will all affect its impact.

This means we should be slow to assume that the development of a given technology will inevitably or inescapably lead to a given social outcome. Think of the internet: because its network structure was inherently well-suited to decentralized and non-hierarchical organization, many confidently predicted that online life would be quite different from that found in the offline world. But that's not quite how things turned out. Largely as a result of the commercial and political world into which it was born, the internet has increasingly come under the direction and control of large corporate and political entities that filter and shape our online experience.

Additionally, we can't assume that technology means progress. In *What Technology Wants* (2010), Kevin Kelly memorably shows that ours is not the first age in which the beneficial promise of technology was massively overhyped. Alfred Nobel, who invented dynamite, believed that his explosives would be a stronger deterrent to war 'than a thousand world conventions'. The inventor of the machine gun believed his creation would 'make war impossible'. In the 1890s, the early days of the telephone, AT&T's chief engineer announced, 'Someday we will build up a world telephone system... which will join all the people of the Earth into one brotherhood.' Still

optimistic in 1912, Guglielmo Marconi, inventor of the radio, announced that, 'the coming of the wireless era will make war impossible, because it will make war ridiculous.' In 1917 Orville Wright predicted that the aeroplane would 'make war impossible' and in the same year Jules Verne declared, 'The submarine may be the cause of bringing battle to a stoppage altogether, for fleets will become useless...war will become impossible.' As Kelly explains, these creations, together with the torpedo, the hot air balloon, poison gas, land mines, missiles, and laser guns, were all heralded as inventions which would lead to the end of war.⁵¹ None did.

While Lenin described Communism as 'Soviet power plus the electrification of the whole country',⁵² Trotsky saw that technological progress was no guarantor of moral progress. '[A]longside the twentieth century', he wrote, there lives 'the tenth or thirteenth':

A hundred million people use electricity and still believe in the magic power of signs and exorcism...What inexhaustible reserves they possess of darkness, ignorance and savagery!...Everything that should have been eliminated from the national organism in the...course of the unhindered development of society comes out today gushing from the throat.⁵³

We cannot take any particular outcome for granted. Nor can we assume that our moral faculties will automatically develop along with our inventions. Everything is still to play for.

Overview of *Future Politics*

Despite its futuristic subject-matter, this book is structured the old-fashioned way. It's meant to be read from start to finish (although tech fanatics may prefer to flick over bits of Part I).

Part I lays the foundations. It sketches out a vision of the future with three defining features. The first is *increasingly capable systems*: machines that are equal or superior to humans in a range of tasks and activities (chapter one). The second is *increasingly integrated technology*: technology that surrounds us all the time, embedded in the physical and built environment (chapter two). The third is *increasingly quantified society*: more and more human activity (our actions, utterances, movements, emotions) captured and recorded as data, then sorted, stored, and processed by digital systems (chapter three). The term I use to describe this future is the *digital lifeworld*, a dense and teeming system that links human beings, powerful machines, and abundant data in a web of great complexity and delicacy.

Chapter four, ‘Thinking Like a Theorist’, surveys the political and intellectual challenges thrown up by the digital lifeworld, and the theoretical tools we have to address those challenges.

Part II turns to the future of *power*. Its central argument is that certain technologies will be a source of great power in the digital lifeworld (chapter five). Some of these technologies will exert power by applying a kind of *force* to human beings. Imagine a self-driving car that refuses to park on a yellow line, or a shopping app that won’t process orders for materials that look like those needed to make a bomb (chapter six). Others will exert power through *scrutiny*, by gathering and storing intimate details about us, and even predicting our behaviour before it happens (chapter seven). A final set of technologies will exert power by controlling our *perception*. These platforms will be able to filter what we know of the wider world, set the political agenda, guide our thinking, stoke our feelings, and mobilize our prejudices to a greater extent even than the media barons of the past (chapter eight).

These three forms of power—*force*, *scrutiny*, and *perception-control*—are as old as politics itself. What’s new is that digital technology will

give them a potency that far exceeds any previous instruments of power known to humankind. The main consequence for politics, I suggest, will be that those who control these technologies of power will be increasingly able to control the rest of us. Two groups stand to benefit the most: political authorities and big tech firms. That's the focus of chapter nine.

This change in the nature of power will affect every aspect of political life. Part III looks at the implications for liberty. On the one hand, new inventions will allow us to act and think in entirely new ways, unleashing exciting new forms of creation, self-expression, and self-fulfilment. On the other hand, we should expect to see a radical increase in the capacity of political authorities to enforce the law, leading to a corresponding reduction in what we are able to get away with. In short, the digital lifeworld will be home to systems of law enforcement that are arguably *too* effective for the flawed and imperfect human beings they govern (chapter ten). What's more, an increasing number of our most cherished freedoms—including the freedom to think, speak, travel, and assemble—will increasingly be entrusted to private tech firms, whose engineers and lawyers will design and operate the systems through which those freedoms are exercised. For freedom of speech, we'll rely on the restraint of social media and communications platforms; for freedom of thought, we'll depend on the trustworthiness of news and search algorithms; for moral autonomy, we'll rely on the judgement of those who determine what we can and can't do with their digital systems. That's chapter eleven.

Growth in the power of political and tech élites will, I suggest, urgently require parallel growth in the power of citizens to hold those élites to account. That's the focus of Part IV, which considers the future of democracy. It suggests a number of ways in which democracy might be transformed, for better or worse, by greater human participation in the form of Direct Democracy, Deliberative

Democracy, and Wiki Democracy; or by greater machine involvement in the form of Data Democracy and AI Democracy (chapters twelve and thirteen).

In Part V we turn to the future of social justice. In the digital lifeworld, I suggest, algorithms will play a central role in the *distribution* of important social goods, like jobs, loans, housing, and insurance (chapter fourteen). Algorithms will also increasingly be used to rank, rate, score, and sort us into social hierarchies of status and esteem (chapter fifteen). Who is seen and who remains unseen? Who finds popularity and who is forgotten? Who matters and who, in social terms, might as well not exist? These are important questions of *recognition*. Both distribution and recognition are essential to social justice, and previously they were left to the market, the state, and society. In the digital lifeworld, questions of social justice will depend, in large part, on the decisions taken by those who curate the relevant algorithms.

The digital lifeworld will also give rise to new and strange forms of injustice. Think of the online passport system in New Zealand that rejected the passport photograph of a man of Asian descent because it concluded his eyes were closed.⁵⁴ Or the voice-recognition systems that couldn't recognize women's voices because they had only ever 'heard' the voices of men.⁵⁵ Or the online auto-tagging systems that tagged photographs of black people as 'apes' and pictures of concentration camps as 'sport' and 'jungle gym'.⁵⁶ These are real examples, and they're just the beginning. It used to be that only humans could humiliate and degrade us. No longer. The implications for social justice are profound (chapter sixteen).

There are compelling reasons to suspect that the digital lifeworld could give rise to serious economic disparities between rich and poor, particularly as digital systems come to perform more and more tasks previously only done by humans, potentially leading

to large-scale technological unemployment (chapter seventeen). Chapter eighteen addresses the concern that the future economy might favour only an élite class of ‘owners’ of productive technologies, while a struggling majority is left to fight over a shrinking share of the pie. I call this the Wealth Cyclone. To avert this outcome, I suggest, we may need to revisit the very notion of *property* itself.

The central danger identified in this book is that gradually, and perhaps initially without noticing, we become increasingly subjugated to digital systems that we can scarcely understand, let alone control. That would place us, in turn, at the mercy of those who control those digital systems. In chapter nineteen I suggest two ways in which such a fate might be avoided. The first is *transparency*: to make sure that those who have the power to affect our core freedoms, or to affect the democratic process, or to settle matters of social justice, are never allowed to operate in darkness. The second is what I call the *new separation of powers*: to make sure that no entity is allowed to secure control over more than one of the means of *force*, *scrutiny*, and *perception-control*, or achieve a monopoly over any of them.

The book closes with a brief foray into a time *after* the digital lifeworld, where the world is so transformed that the idea of *politics* itself loses its meaning (chapter twenty).

A few final words before we press on.

This book can only scratch the surface. Whole swathes of political life are left untouched, along with ideas from outside the western philosophical tradition. Some issues are necessarily truncated and simplified. This is both a concession to my own limitations and an act of mercy for readers: any effort to be comprehensive would have resulted in a book of biblical girth. ‘May others,’ as Wittgenstein says, ‘come and do it better.’⁵⁷ (Except I mean it.)

I try throughout not to be dogmatic (a difficult thing for a lawyer). My aim is to offer a guide, not a manifesto. Some of the technologies and ideas may be new and unfamiliar, but our aim, at least, is as old as humankind: to be ‘at once astonished at the world and yet at home in it’.⁵⁸

PART I

The Digital Lifeworld

'It shall seem to men that they see new destructions in the sky, and the flames descending therefrom shall seem to have taken flight and to flee away in terror; they shall hear creatures of every kind speaking human language; they shall run in a moment, in person, to divers parts of the world without movement; amidst the darkness, they shall see the most radiant splendors. O marvel of mankind! What frenzy has thus impelled you!'

Leonardo Da Vinci (1452–1519)

ONE

Increasingly Capable Systems

'There is nothing that man fears more than the touch of the unknown.' **Elias Canetti, *Crowds and Power* (1984)**

We begin with a survey of where we are now and a sketch of the world to come.

In the next century, politics will be transformed by three developments: increasingly capable systems, increasingly integrated technology, and increasingly quantified society. Together these changes will give rise to a new and different type of collective life: the *digital lifeworld*. The strange word *lifeworld* is taken from the German *Lebenswelt*, meaning all of the immediate experiences, activities, and contacts that make up our individual and collective worlds. When you imagine the digital lifeworld, imagine a dense and teeming system that links human beings, powerful machines, and abundant data in a web of great delicacy and complexity. In this Part, I don't presume to evaluate or critique the effects of the technologies I describe. The purpose is to identify and understand them, and then (in chapter four) to inspect the intellectual tools that will help us to think clearly about what it all means for politics.

The next three chapters contain many real-life examples, and the idea is not to hold them all in mind at once. Instead our aim is to glimpse, if only in outline, the future that stalks us. That begins with acknowledging that we will soon live alongside computing

machines of extraordinary capability. These are *increasingly capable systems*,¹ and they will be the first defining feature of the digital lifeworld.

Artificial Intelligence

The field of Artificial Intelligence (AI), which began in earnest in 1943, is concerned with the construction of ‘intelligent’ digital systems. When I refer to AI here, I am describing systems that can perform tasks that in the past were believed to require the cognitive and creative processes of human beings.² Progress hasn’t always been smooth but today it is impressive and gathering in speed. There are thousands of activities formerly done only by humans that digital systems can now do faster, more efficiently, with greater precision, and at a different order of magnitude.

AI systems are now close to surpassing humans in their ability to translate natural languages, recognize faces, and mimic human speech.³ Self-driving vehicles using AI are widely expected to become commonplace in the next few years (Ford is planning a mass-market model by 2021).⁴ In 2016 Microsoft unveiled a speech-recognition AI system that can transcribe human conversation with the same number of errors, or fewer, as professional human transcriptionists.⁵ Researchers at Oxford University developed an AI system capable of lip-reading with 93 per cent accuracy, as against a 60 per cent success rate among professional lip-readers.⁶ AI systems can already write articles about sports, business, and finance.⁷ In 2014, the Associated Press began using algorithms to computerize the production of hundreds of formerly handcrafted earnings reports, producing fifteen times as many as before.⁸ AI systems have directed films and created movie trailers.⁹ AI ‘chatbots’ (systems that can ‘chat’ to you) will soon be taking orders at restaurants.¹⁰

Ominously, engineers have even built an AI system capable of writing entire speeches in support of a specified political party.¹¹ It's bad enough that politicians frequently sound like soulless robots; now we have soulless robots that sound like politicians.

Every day, algorithms carry out countless transactions in financial markets on behalf of investors. They're trusted to pursue complex strategies based on changing market conditions. Deep Knowledge Ventures, a Hong Kong-based investor, has appointed an algorithm called VITAL (Validating Investment Tool for Advancing Life Sciences) to its board of directors.¹² In medicine, AI systems can now differentiate between lung cancers and predict survival periods better than human pathologists. Researchers believe the same is likely to be true of other cancers in due course.¹³ In law, an AI system correctly guessed the outcomes of 79 per cent of hundreds of cases heard at the European Court of Human Rights.¹⁴ Lethal autonomous weapons systems are under development. That's missiles, armed drones, and weaponized robots supported by AI. If deployed into the theatre of war, they'd have the capacity to select targets based on certain criteria before homing in and destroying them—potentially, in due course, without intervening human decision-making.¹⁵

Games of skill and strategy are considered a good way to gauge the increasing capability of digital systems. In short, they now beat the finest human players in almost every single one, including backgammon (1979), checkers (1994), and chess, in which IBM's Deep Blue famously defeated world champion Garry Kasparov (1997). In 2016, to general astonishment, Google DeepMind's AI system AlphaGo defeated Korean Grandmaster Lee Sedol 4–1 at the ancient game of Go, deploying dazzling and innovative tactics in a game exponentially more complex than chess. 'I... was able to get one single win,' said Lee Sedol rather poignantly; 'I wouldn't exchange it for anything in the world.'¹⁶

A year later, a version of AlphaGo called AlphaGo Master thrashed Ke Jie, the world's finest human player, in a 3–0 clean sweep.¹⁷

A radically more powerful version now exists, called AlphaGo Zero. AlphaGo Zero beat AlphaGo Master 100 times in a row.¹⁸

As long ago as 2011, IBM's Watson vanquished the two all-time greatest human champions at *Jeopardy!*—a TV game show in which the moderator presents general knowledge 'answers' relating to sports, science, pop culture, history, art, literature, and other fields and the contestants are required to provide the 'questions'. *Jeopardy!* demands deep and wide-ranging knowledge, the ability to process natural language (including wordplay), retrieve relevant information, and answer using an acceptable form of speech—all before the other contestants do the same.¹⁹ The human champions were no match for Watson, whose victory marked a milestone in the development of artificial intelligence. This was a system that could answer questions 'on any topic under the sun... more accurately and quickly than the best human beings'.²⁰ The version of Watson used on *Jeopardy!* was said to be the size of a bedroom; by the early 2020s it's expected that the same technology, vastly improved, will sit comfortably in a smartphone-sized computer.²¹ Today, what IBM calls 'Watson' no longer exists in a single physical space but is distributed across a cloud of servers that can be accessed by commercial customers on their computers and smartphones.²² As IBM is keen to stress, different versions of Watson do more than win game shows. In late 2016, one Watson platform discovered five genes linked to amyotrophic lateral sclerosis (ALS), a degenerative disease that can lead to paralysis and death. The system made its discovery after digesting all the published literature on ALS and parsing every gene in the human genome. This took Watson a matter of months; humans would have taken years.²³ In early 2017, Fukoku Mutual Life Insurance in Japan sacked thirty-four of its staff and replaced them with Watson's 'Explorer' platform, which

will chomp through tens of thousands of medical records and certificates, data on hospital stays, and surgical information to calculate payouts to policyholders.

AI has spawned a multiplicity of sub-fields, each applying different methods to a wide range of problems. There is a spectrum of approach, for instance, between those who seek to recreate the neural engineering of the human brain, just as ‘early designs for flying machines included flapping wings’, and those who employ entirely new techniques tailored for artificial computation.²⁴ Some researchers seek the holy grail of an *artificial general intelligence* like the human mind, endowed with consciousness, creativity, common sense, and the ability to ‘think’ abstractly across different environments. One way to achieve this goal might be *whole-brain emulation*, currently being pursued in the Blue Brain project in Switzerland. This involves trying to map, simulate, and replicate the activity of the (more than) 80 billion neurons and tens of trillions of synapses in the human brain, together with the workings of the central nervous system.²⁵ Whole-brain emulation remains a remote prospect but is not thought to be technically impossible.²⁶ As Murray Shanahan argues, our own brains are proof that it’s physically possible to assemble ‘billions of ultra-low-power, nano-scale components into a device capable of human-level intelligence’.²⁷

Most contemporary AI research, however, is not concerned with artificial general intelligence or whole-brain emulation. Rather, it is geared toward creating machines capable of performing specific, often quite narrow, tasks with an extraordinary degree of efficacy. AlphaGo, Deep Blue, and Watson did not possess ‘minds’ like those of a human being. Deep Blue, whose only function was to play chess, used ‘brute number-crunching force’ to process hundreds of millions of positions each second, generating every possible move for up to twenty or so moves.²⁸

It's tempting to get hung up on the distinction between machines that have a narrow field of cognitive capacity and those able to 'think' or solve problems more generally. The latter is a juicier target than the former. But distinctions between 'narrow' and 'broad', or 'strong' and 'weak' AI, can obscure the fact that even narrowly focused AI systems will create vast new opportunities and risks worthy of careful attention in their own right. Soon computers will be able to do things that humans can do, even if they don't do them in the same way—and a lot more besides. And it doesn't matter that one system may only be able to perform a small number of tasks. At the very least, it looks like the digital lifeworld will play host to a multitude of overlapping AI systems, each engineered to perform specific functions. And for those of us on the receiving end, it may be impossible to distinguish between a system that possesses general intelligence, and one that uses fifty different sub-systems to give the *impression* of general intelligence. In most important respects, the effect will be the same.

Machine Learning

The most important sub-field of AI at present is *machine learning*. As Pedro Domingos explains in his book *The Master Algorithm* (2015), the traditional way of getting a computer to do something was 'to write down an algorithm'—a series of instructions to the computer—'explaining...in painstaking detail' how to perform that task.²⁹ In contrast with an ordinary algorithm, a *machine learning algorithm* can discover on its own how to recognize patterns, create models, and perform tasks. It does this by churning through large amounts of data, identifying patterns, and drawing inferences. Machine learning algorithms can learn both knowledge ('if a thing looks like X it is a Y') and skills ('if the road curves left, turn the wheel left').³⁰

The idea is that after a certain point ‘we don’t have to program computers’; instead ‘they program themselves’.³¹

Many of the AI systems I have described employ machine learning techniques. Indeed, machine learning algorithms are all around us:³²

Amazon’s algorithm, more than any one person, determines what books are read in the world today. The NSA’s algorithms decide whether you’re a potential terrorist. Climate models decide what’s a safe level of carbon dioxide in the atmosphere. Stock-picking models drive the economy more than most of us do.

When the time comes for you to take your first ride in a self-driving car, remember that:³³

no engineer wrote an algorithm instructing it, step-by-step, how to get from A to B. No one knows how to program a car to drive, and no one needs to, because a car equipped with a learning algorithm picks it up by observing what the driver does.

Machine learning, to borrow from Domingos, is the automation of automation itself.³⁴ It’s a profound development because it liberates AI systems from the limitations of their human creators. Facebook’s engineers, among others, are working on a machine learning algorithm that can build other machine learning algorithms.³⁵

Machine learning algorithms generally ‘learn’ in one of three ways. In *supervised* learning, the human programmer sets a series of defined outcomes and provides the machine with feedback about whether it’s meeting them. By contrast, in *unsupervised* learning, the machine is fed data and left to look for patterns by itself. An unsupervised machine can therefore be used to ‘discover knowledge’, that is, to make connections of which its human programmers were totally unaware.³⁶ In *reinforcement* learning, the machine is given ‘rewards’ and ‘punishments’ telling it whether what it did was right or wrong. The machine self-improves.

Many of the advances described in this chapter, particularly those involving images, speech, and text, are the result of so-called ‘deep learning’ techniques that use ‘neural networks’ inspired by the structure of animal brains. Google launched one in 2012, integrating 1,000 large computers with more than a billion connections. This computer was presented with 10 million ‘random’ images from YouTube videos. It was not told what to look for, and the images were not labelled. After three days, one unit had learned to identify human faces and another had learned to respond to images of a cat’s face (this is YouTube after all).³⁷ Engineers at Google now use ‘duelling’ neural networks to train *each other*: one AI system creates realistic images while a second AI system plays the role of critic, trying to work out whether they’re fake or real.³⁸

The rapid increase in the use of deep learning can be seen from the AI systems used in games. The version of Deep Blue that beat Garry Kasparov at chess in 1997 was programmed with many general principles of good play. What’s most remarkable about AlphaGo Zero—the latest and most powerful incarnation of the Go-playing AI systems—however, is that it ‘learned’ not by playing against the very best humans or even learning from human play, but by playing against *itself* over and over again, starting from completely random moves, and rapidly improving over time.³⁹

Machine learning has been around for a while. Its rapid growth and success in the last couple of decades has been enabled in part by the development of new algorithms, but mostly by the explosion in processing power and the growth in available data (chapter three). Data is critical to machine learning. Too little will stunt the growth of a machine learning algorithm, but given enough, ‘a learning program that’s only a few hundred lines long can easily generate a program with millions of lines, and it can do this again and again for different problems.’⁴⁰ This is why data has been called ‘the new coal’⁴¹ and those who gather it are called ‘data miners’.

As we'll see at various points in this book, however, reliance on flawed real-world data can cause havoc with machine learning systems. Microsoft launched its AI chatbot Tay on Twitter on 23 March 2016. Tay was intended to mimic the speech of a nineteen-year-old girl and to learn from interactions with other Twitter users. Sixteen hours after its launch, Tay was removed from active duty after posting a series of racist and sexually inflammatory tweets, including one which captioned a photo of Adolf Hitler with the tag 'swag alert', and another saying 'Fuck my robot pussy daddy I'm such a naughty robot'. Tay had 'learned' to communicate this way from other users on Twitter. This example says as much about humans on social media as it does about machine learning.

A final point about machine learning: it used to be that the computing power that fuelled any particular system was physically present within the system in question. The most powerful digital devices literally contained the processors that made them run. The arrival of cloud computing in the last decade or so has meant that the computing power no longer needs to be located within the device itself: like Apple's Siri it can be accessed over the internet. This has major implications for the integration of technology, as it means that small devices can draw on big computing resources (chapter two). But it's also important for machine learning because it means that machines don't need to 'learn' from their own separate experiences; they can learn from others' too, so that every machine in a swarm or fleet adds to the collective 'intelligence' of the whole.

Exponential Growth in Computing Power

Progress in artificial intelligence and machine learning has been underpinned by advances in mathematics, philosophy, and neuroscience. But as mentioned, it has depended above all on two

developments: an explosion in the amount of available data, and an explosion in computing power.

For the last fifty years or so, computing power—the ability of computer chips to process data—has grown at an exponential rate, doubling roughly every two years. This progress is generally expected to continue. On current trends a computer in 2029 will be sixty-four times faster than it was in 2017. If the technology continued to improve at the same rate, then in 2041 it would be 4,096 times faster. After thirty years, the computer would have grown *millions* of times more powerful. Ray Kurzweil and others predict that within the next decade or so, a normal desktop machine (costing \$1,000 or thereabouts) will rival and surpass the processing power of the human brain. By 2050, ‘one thousand dollars of computing will exceed the processing power of all human brains on earth’.⁴² If this sounds unlikely, look back to where we have come from. Just thirty years ago, it would have needed 5,000 desktop computers to rival the processing power of today’s iPad Air.⁴³ Sixty years ago, 2010’s iPad2 (now hopelessly out of date) would have cost \$100 trillion, roughly twenty-five times the United States federal budget for 2015.⁴⁴ The average smartphone has more processing power than the Apollo Guidance Computer that sent Neil Armstrong to the moon.⁴⁵

Our brains are not wired to think exponentially. We tend to think of change as happening in a straight upward line, not noticing when the underlying rate of change is itself accelerating. To put it in perspective, try (in Pedro Domingos’ example) to imagine a single *E. Coli* bacterium of miniscule proportions. This bacterium divides in two and doubles in size roughly every fifteen or twenty minutes. Assuming the right conditions, after a few hours it will mushroom into a colony of bacteria, but still too small for the human eye to see. After twenty-four hours, however, that *E.coli* microbe will swell into a

bacterial mass the size of planet Earth itself.⁴⁶ That's the power of exponential growth.

The theory that computer processing power doubles roughly every two years is often referred to as Moore's Law. This 'law', which is not actually a law but rather an observable pattern of development, has been called 'Silicon Valley's guiding principle, like all ten commandments wrapped into one'.⁴⁷ It originated with a 1965 article by Gordon Moore, the co-founder of Intel, who predicted that the number of components that could be crammed onto an integrated circuit would double roughly every two years. At the time, Moore predicted that this trend would continue 'for at least ten years'.⁴⁸ Others were sceptical, giving it a couple of years at best. Moore's Law has now persisted for more than five decades. As Walter Isaacson notes, it became more than a prediction: it was a 'goal for the industry, which made it partly self-fulfilling'.⁴⁹ Interestingly, processing power is not the only technology improving at an exponential rate. A host of others, including hard-disk capacity, bandwidth, magnetic data storage, pixel density, microchip density, random access memory, photonic transmission, DNA sequencing, and brain-scan resolution, are all developing in the same way.⁵⁰ If Moore's Law continues, the next few decades will witness the arrival of machines of remarkable capability. A world in which every desktop-sized machine has the processing power of *all of humanity* will be radically different from our own.

It's sometimes said that Moore's Law will grind to a halt in the next few years, mostly because it will become physically impossible to cram any more transistors into the same microchip, and because the economic efficiencies enjoyed in the last half-century are set to diminish. There is certainly some evidence of a slowdown, although Moore's Law has been given its last rites countless times in the past.⁵¹ However, it is probably wrong to assume that the *current* computing paradigm—the integration of transistors onto 2D wafers of silicon

(*the integrated circuit*)—is the final computing paradigm, and cannot itself be improved upon by some other method. History, market forces, and common sense all suggest otherwise. Before the integrated circuit, computers were built using individual transistors. Before that, in the days of Alan Turing, they relied on vacuum tubes, relays, and electromechanics. The story of computing is the story of a succession of increasingly powerful methods of processing information, each developing exponentially, reaching its physical limitations, and then being replaced by something better. Exponential growth in computing processing power stretches back to the seventeenth century and ‘the mechanical devices of Pascal’.⁵² Nothing is inevitable, but Moore’s law did not begin with the integrated circuit and it is unlikely to end with it either.

The interesting question is what will come next. A host of new methods are already in development, aimed at reaching the frontier of silicon-based computing and pushing beyond it. One approach has been to use non-silicon materials in chips for the first time.⁵³ Another possibility is to move from the current paradigm of ‘2D’ integrated circuits—where transistors are arranged side-by-side on a wafer of silicon—to a ‘3D’ approach where transistors are piled high.⁵⁴ Another approach might be to abandon silicon altogether in favour of carbon nanotubes as the material for building even smaller, more efficient transistors.⁵⁵ Yet another approach, currently taken by Google, would be to use more special-purpose computer chips for particular functions—chips that do fewer things but much faster.⁵⁶ Microsoft increasingly uses a new type of chip that could combine much greater speed with flexibility.⁵⁷

Looking further ahead, engineers at Google and elsewhere are already hard at work developing ‘quantum computers’ which, in certain tasks, are expected to be able to compute well beyond the capabilities of classical computers.⁵⁸ Another possible alternative to

silicon might be to use 2D graphene-like compounds and ‘spintronic’ materials, which compute by harnessing the spin of electrons rather than moving them around.⁵⁹ There’s also the growing field of neuroelectronics, which seeks to reverse-engineer the neural networks of the human brain while potentially requiring less power than silicon.⁶⁰ In the still longer term, Quantum Dot Cellular Automata (‘QDCA’) technology may yield an unimaginably small semiconductor capable of doing the work of a transistor, but using much less power and wasting little energy.⁶¹

Many of these technologies are still in their infancy and nothing certain can be said about the future of Moore’s Law. But the least likely outcome is that computer science simply grinds to a halt, with hungry young Silicon Valley engineers hanging up their circuit boards and heading for retirement. Whatever the next computing paradigm turns out to be, it’s certainly not unreasonable to assume that computing power will continue to grow at the same rate as it has since Pascal’s invention of the calculator 400 years ago.