CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

POLITICAL THEORY OF THE DIGITAL AGE

With the rise of far-reaching technological innovation, from artificial intelligence to Big Data, human life is increasingly unfolding in digital lifeworlds. While such developments have made unprecedented changes to the ways we live, our political practices have failed to evolve at pace with these profound changes. In this pathbreaking work, Mathias Risse establishes a foundation for the philosophy of technology, allowing us to investigate how the digital century might alter our most basic political practices and ideas. Risse engages major concepts in political philosophy and extends them to account for problems that arise in digital lifeworlds, including AI and democracy, synthetic media and surveillance capitalism, and how AI might alter our thinking about the meaning of life. Proactive and profound, *Political Theory of the Digital Age* offers a systemic way of evaluating the effect of AI, allowing us to anticipate and understand how technological developments impact our political lives – before it's too late.

MATHIAS RISSE is the Berthold Beitz Professor in Human Rights, Global Affairs, and Philosophy at Harvard University. He is the author of *On Global Justice* (2012), *Global Political Philosophy* (2012), and *On Justice: Philosophy, History, Foundations* (2020), as well as the coauthor of *On Trade Justice* (2019) and *Holding Together: The Hijacking of Rights in America* and *How to Reclaim Them for Everyone* (2022).

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

Political Theory of the Digital Age

WHERE ARTIFICIAL INTELLIGENCE MIGHT TAKE US

MATHIAS RISSE
Harvard University

CAMBRIDGE
UNIVERSITY PRESS

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE

Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781009255219

DOI: 10.1017/9781009255189

© Mathias Risse 2023

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press & Assessment.

First published 2023

A catalogue record for this publication is available from the British Library

A Cataloging-in-Publication data record for this book is available from the Library of Congress

ISBN 978-1-009-25521-9 Hardback
ISBN 978-1-009-25519-6 Paperback

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

For Kozue

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE

Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

Contents

Preface *page* ix
Acknowledgments xxv

1 **Introduction: Digital Lifeworlds in Human History** 1

2 **Learning from the Amish: Political Philosophy as Philosophy of Technology in the Digital Century** 23

3 **Artificial Intelligence and the Past, Present, and Future of Democracy** 47

4 **Truth Will Not Set You Free: Is There a Right to It Anyway? Elaborating on the Work Public Reason Does in Life 2.0** 73

5 **Knowing and Being Known: Investigating Epistemic Entitlements in Digital Lifeworlds** 96

6 **Beyond Porn and Discreditation: Epistemic Promises and Perils of Deepfake Technology** 119

7 **The Fourth Generation of Human Rights: Epistemic Rights in Life 2.0 and Life 3.0** 137

8 **On Surveillance Capitalism, Instrumentarian Power, and Social Physics: Securing the Enlightenment for Digital Lifeworlds** 160

9 **Data as Social Facts: Distributive Justice Meets Big Data** 183

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

10 God, Golem, and Gadget Worshippers: Meaning of Life in the Digital Age 207

11 Moral Status and Political Membership: Toward a Political Theory for Life 3.0 231

Epilogue 255

Bibliography 259
Index 291

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

Preface

It seems probable that once the machine thinking method has started, it would not take long to outstrip our feeble powers ... At some stage therefore we should have to expect the machines to take control.

—Alan Turing[1]

The world of the future will be an ever more demanding struggle against the limitations of our intelligence, not a comfortable hammock in which we can lie down to be waited upon by our robot slaves.

—Norbert Wiener[2]

1 THE NEED TO DO POLITICAL THEORY FOR THE DIGITAL AGE

Political thought explores how we should live together. Our lives increasingly unfold in digitally interconnected ways, and so at this stage, political theory must investigate how to inhabit this digital century. Much innovation in the digital domain is driven by machine learning, a set of methods that analyze the myriads of available data ("Big Data") for trends and inferences. Unlike conventional programs, machine-learning algorithms learn by themselves, drawing on their supply of data. These algorithms are based on so-called "neural networks," programs that imitate the ways in which brain cells interact with each other. Typically, such algorithms are what efforts at creating artificial intelligence (AI) amount to today. Owing to their sophistication and sweeping applications, these techniques are poised to alter our

[1] Alan Turing, "'Intelligent Machinery: A Heretical Theory,' a Lecture Given to '51 Society' at Manchester," AMT/B/4, The Turing Digital Archive, https://turingarchive.kings.cam.ac.uk/publications-lectures-and-talks-amtb/amt-b-4, last accessed June 27, 2022.

[2] Wiener, *God & Golem*, 69.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

world dramatically. In some circles, there is much enthusiasm about what might be possible: "No objective answer is possible to the question of when an 'it' becomes a 'who,'" writes one engineer, "but for many people, neural nets running on computers are likely to cross this threshold in the very near future."[3]

At the time of writing, the production of AI models appears to be moving into its own kind of industrial age, much beyond earlier stages when these models were more artisanal and speculative. These advances have drawn on breakthroughs from around 2010 – in the words of a Google Senior Vice President for Research, "the 2010s were truly a golden decade of deep learning research and progress"[4] – when computers became powerful enough to run enormously large machine-learning models and the Internet started to provide the humungous amount of training data such algorithms require to go through their learning process. Since then, conceptual breakthroughs in programming have led to the creation of ever more complex and sophisticated software – and the supercomputers required to enable the most advanced AI models to unfold their full power have become so expensive that, short of well-funded governmental AI strategies in the wealthiest countries, the field is likely to end up being dominated by the research agendas of private companies with substantial resources.

Regarding *specialized AI*, at the high end, one may think of algorithms winning at chess or Go – where the point is not only that AI beats human players, but the stunning progression of how that has happened: Initially AI drew lessons from the history of human play, then it played against itself, yet later AI taught itself the rules of the games, and eventually it created systems that could learn and win at multifarious games (all of which happened within a few years). Still at the high end, one might also think of speech recognition and natural language processing, including the emergence of large language models capable of generating human-like products. But leaving such high-end technology aside, specialized AI already operates in multifarious everyday devices. In contrast to such distinctly specialized operations, *general AI* approximates human performance across domains. Once there is general AI smarter than us, it might well produce something smarter than itself, and so on, perhaps very fast. That moment is known as the *singularity*, an intelligence explosion that could alter the course of human history in ways nothing else has ever done. In opening his 2021 BBC Reith Lectures ("Living with Artificial Intelligence"), Stuart Russell declares "the eventual emergence of general-purpose Artificial Intelligence [to be] the biggest event in human history" (regardless of whether such emergence is precipitated by an intelligence explosion).[5]

[3] Agüera y Arcas, "Do Large Language Models Understand Us?" 183. For the state of the art in AI research and of the discussions of its ethical dimensions and impact on societies as of spring 2022, see Manyika, *AI & Society*, *Spring 2022 Issue of Daedalus*; Liang and Bommasani, "On the Opportunities and Risks of Foundation Models."

[4] Dean, "A Golden Decade of Deep Learning," 69.

[5] "The Reith Lectures: Living with Artificial Intelligence," BBC, www.bbc.co.uk/programmes/m001216k, last accessed June 27, 2022.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

To be sure, the possibility, nature, and likelihood of a singularity remain intensely disputed, and we are nowhere near anything like it. But "nowhere near" might mean in terms of engineering capacities rather than time. A few major breakthroughs – and the recent past is certainly full of breakthroughs, and here one might think again of the advances in game-playing and language-processing just mentioned – could radically transform the field. (As a Microsoft Chief Technology Officer wrote in 2022, "I have been surprised so many times over the past two decades by what AI scientists and researchers have been able to accomplish that I have learned to heed the second half of Arthur C. Clarke's first law: When a distinguished but elderly scientist states that something is possible, they are almost certainly right. When they state that something is impossible, they are very probably wrong."[6]) Engagement with these developments from a standpoint of political theory is inevitably constrained by their evolving nature. But while an intelligence explosion would undoubtedly revolutionize politics (to put it mildly), many questions for political thought already arise from the increasing presence of digital technology across domains all around us.

To be sure, often these questions are not fundamentally new but have assumed fresh relevance or need to be rethought. We must do political theory for the digital age. And eventually this age might transform into an entirely new world populated with new kinds of superintelligences. My starting point is the liberal-egalitarian outlook articulated by John Rawls and others. Using this outlook *as a starting point* does not mean that at any moment in this book it would be sensible to ask, "And what would Rawls say about this?" There is too much novelty to the problems we encounter here for that to be sensible. At the same time, every journey into new territory must start from somewhere. Notwithstanding its attractions as an innovative approach to the problems of our time, the liberal-egalitarian outlook has neglected two themes that are becoming increasingly important in our digital century.[7] The first (more general) theme is the distinctive importance of technology for political thought. The second (rather specific) one is the roles of citizens as *knowers* and *knowns* – that is, as both holders and providers of "data" and "information," to use two of the defining terms of our age. And, of course, the possible advent of superintelligences is not on the radar of Rawls's theory at all.

There is nothing to bar liberal-egalitarian thought from paying more attention to technology, to recognize it as a domain that creates philosophical problems all its own, rather than derivative ones. Putting these problems on the agenda of liberal-egalitarian thought ushers in authors from traditions that have made technology

[6] Scott, "I Do Not Think It Means What You Think It Means," 83.

[7] That outlook must be supplemented in other ways, especially through the addition of racial themes, which many authors have noticed and on which I have written myself in Risse, *On Justice*.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

central: phenomenology and especially Marxism, which has put technology front and center all along in ways that I think all traditions of political thought should now. The Marxist tradition makes frequent appearances in this book. Similarly, nothing in liberal-egalitarian thought sits uneasily with an explicit articulation of matters of epistemic justice and epistemic rights (justice and rights to the extent that they are concerned with what people may know and what may be known of them). But epistemic justice has been theorized largely outside of the liberal-egalitarian tradition, and epistemic rights is a rather new topic anyway. Here too an adequate articulation within liberal-egalitarian thought requires turning to authors who have done relevant work, in this case especially Michel Foucault and writers in the Science, Technology, and Society (STS) tradition. The possible advent of super-intelligences requires new conversations altogether, no matter how speculative they would be for now.

This book means to help set an agenda in a new domain of inquiry where things have been moving fast. Michael Rosen once referred to political theory as "the oasis where the caravans meet."[8] Political *philosophy* differs from political *theory* by focusing on the depth of arguments made in and about the political domain, holding at bay not only much of the societal background but also other kinds of inquiry that might bear on the questions at hand. Political theory is more broadly oriented and manages its trade-offs differently, tending to integrate more breadth at the expense of in-depth investigations of ideas. At this stage, in the process of transferring political thought to the digital domain, it is important to take the more comprehensive view of political theory rather than the narrower one of political philosophy to make sure we transfer the debates in their full richness into this new era. Accordingly, "political theory" is in the title of this book.

This book aims to create a better footing for the philosophy of technology and for discussions around epistemic rights and justice in the liberal-egalitarian outlook, as a way of helping to bring into the digital era – the era of AI and Big Data, and possibly the age of the singularity – the debates that have traditionally preoccupied political thinkers. In this spirit, my focus is on themes rather than authors. Inevitably, the kind of work that needs to be done here creates some strange-bedfellow phenomena that generate questions about the compatibility of these authors. But one feature of the way in which the advent of AI affects political thought – and this is one of the main points I hope to make, by the very manner I approach my topics – is that the relationship among various traditions must be reassessed. Still, the grounding in the Rawlsian approach is the guide to creating coherence and sets the stage for the broader conversations to unfold.

[8] Rosen, *Dignity*, xvi.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

2 THE RAWLSIAN STARTING POINT (AND THE CONNECTIONS TO MARX)

Let me explain the basics of the Rawlsian starting point, with an eye on the relationship of its central elements to Marxist thinking. Rawls is of historical importance for political theory because he offers a complex account of distributive justice that integrates many ideas about how to organize society from the preceding two centuries or so. During this period, industrialized societies were becoming sufficiently interconnected to generate a new level of questioning about the absolute and relative statuses of different groups in society. "Social justice" emerged as an area for both political mobilization and intellectual inquiry.

Originally developed in his 1971 book *A Theory of Justice*, Rawls's theory of *justice as fairness* envisions a society of free and equal citizens holding equal basic rights and cooperating within an economic system that is egalitarian in certain respects.[9] Citizens are free in that each person feels entitled to make claims on institutions. Citizens are also free in being able to take responsibility for their lives (if provided with suitable opportunities and resources). Citizens are equal by virtue of having no artificial constraints imposed on their capacities to participate in social cooperation over a lifetime. Rawls aims to describe a just arrangement of major political and social institutions of a liberal society: political constitution, legal system, economy, family, and so on. Their arrangement is the *basic structure*. Individuals are supposed to cooperate within its confines. Compliance is enforced, through police, courts, and other agencies.

There are multifarious ways of organizing the basic structure. Different rules favor different groups by increasing that group's chances of seizing larger shares of the social product. Rawls insists that nobody *deserves* to be born with social advantages or disadvantages, or to be more or less gifted than others. For this reason, citizens are not at a deeper level entitled to having rules of cooperation that let them benefit from such assets. To be sure, any social system generates legitimate expectations: If someone performs in accordance with standards within the system, they can legitimately demand the rewards anticipated by those standards. But nobody is entitled to a society arranged in such a way that *their* talents generate large shares of the social product. After all, society's rules should be acceptable to everybody. In light of the characteristics of the basic structure that we just observed, Rawls submits that all social primary goods – cooperatively produced under conditions where compliance is enforced – are to be distributed equally *unless* unequal distribution works to everyone's benefit.

A division of labor occurs within the institutions that make up the basic structure. These institutions distribute the main benefits and burdens of social life, the *social primary goods*. These goods are basic rights and liberties; freedom of movement and free choice among a wide range of occupations; powers of offices and positions of

[9] The 1999 revised edition is now commonly used. See Rawls, *A Theory of Justice*.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

responsibility; income and wealth; as well as the social bases of self-respect, the recognition by social institutions that gives citizens a sense of self-worth and confidence to execute their plans.[10] Recognizing social primary goods as the subject (or the *distribuenda*) of a contemporary theory of distributive justice reveals the sheer range of things we provide to each other through joint activities at the level of society as a whole.

For Rawls, distributive justice requires substantially equal political and civil liberties and fair equality of opportunity in education, while economic inequalities within these constraints are permitted only if they are to everybody's benefit, especially to the benefit of the least advantaged. More specifically, Rawls proposes the following principles of distributive justice. The first – which deals with political status – states that each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, a scheme that is compatible with the same scheme of liberties for all. The second principle – which deals with relative economic standing – falls into two parts. The first part states that social and economic inequalities are to be attached to offices and positions that are open to all under conditions of *fair equality of opportunity*. The second states that remaining social and economic inequalities should be to the greatest benefit of the least-advantaged members of society (*Difference Principle*).[11] Fulfilment of the first principle, on political status, takes priority over the second. Within the second, fair equality of opportunity takes priority over the Difference Principle.

Theorizing social justice in an era of political and economic interdependence means recognizing the extent to which our lives are shared. Rawls's principles combine and rank different ideas. Rawls does not defend one idea as central to social justice but integrates various approaches from the nineteenth and twentieth centuries. Ideas about what people deserve and ideas about what people need are present, but so are liberty and equality. A recognition of the breadth of the domain to which considerations of distributive justice apply comes with a sense that no unitary criterion can guide the distribution. The first principle secures equal citizenship for each person by protecting civil and political liberties, a protection that also goes far in securing for each citizen similarly effective influence on political processes. The first part of the second principle makes sure that one's prospects do not vary given the segment of society one happens to be born into. The second part guarantees that the group that makes the largest concession vis-à-vis a baseline of equality makes as small a concession as feasible. The way of ascertaining whether such an arrangement has been achieved is to identify other feasible arrangements and determine if the respectively least advantaged *in one of those arrangements* would be better off than they are in the current arrangements. If they are, we should switch to that arrangement.

[10] Rawls, *Restatement*, 58f.
[11] Rawls, 41–43. It is called the Difference Principle because it regulates differences among citizens.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

Rawls uses utilitarianism as a foil. Utilitarianism is associated with authors such as Jeremy Bentham, John Stuart Mill, and now Peter Singer and is influential among economists and policymakers. Utilitarianism identifies right action with maximization of collective well-being. By contrast, Rawls does not think such maximization is the correct response to the task of devising a society that free and equal citizens could respectfully coinhabit. Rawls connects to the Kantian tradition, in which respectful treatment of each person (rather than goals formulated at the collective level) is the driving theme. Social primary goods are goods generated through cooperation. Accordingly, their distribution should be justifiable to all. But no distribution would be so justifiable if any individual could be expected to make enormous sacrifices for the sake of social improvements, as utilitarianism implies.

Rawls distinguishes five types of institutional arrangements:[12] laissez-faire capitalism (a political economy defined around the private ownership of the means of production, with very limited state interference mostly to guarantee public safety and other basic forms of protection); welfare-state capitalism (which differs from laissez-faire capitalism by providing assistance or payments to low-income citizens, and thus includes redistributive elements that could come in a variety of forms); state socialism with command economy (where the state owns and regulates the means of production); democratic socialism (which shares with state socialism this commitment to no private ownership of means of production, but where control over firms is widely dispersed); and property-owning democracy (where it is private ownership of means of production that is widely dispersed). For Rawls, social justice requires either democratic socialism or property-owning democracy. Most tellingly, in addition to laissez-faire capitalism, he rejects welfare-state capitalism, because it condones substantial inequalities in the initial distribution of property and skill endowments, merely *re*-distributing some income ex-post. This undermines the ideal of equal citizenship.

Unlike the Marxist tradition, Rawls does not think capitalism's ills are beyond the reach of reform. The response to tensions within capitalist production is not necessarily to envisage a society that leaves behind capitalism. The response is to offer, and realize, a better ideal of justice based on a sensible understanding of society as an interconnected system in which benefits from cooperation must be assigned fairly. In terms of institutional reform, this might take us to either a property-owning democracy (a reformed capitalist system) or democratic socialism (a version of collective control that is highly emancipatory by distributing control widely). Also, Marx sees class conflict as the driving political theme in society and revolution as the only way of resolving class conflict. Marx has a simplistic notion of conflict in industrial societies. But he also thinks conflict can, and inevitably will, be overcome. As I explain later, Rawls has a more complex understanding of these conflicts that societies face today. And he does not think the conflicts can be, or even

[12] Rawls, Part IV.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

ought to be, overcome. Instead, they must be handled properly. The notion that does the crucial work in this regard is *public reason*.[13]

Public reason requires citizens to be able to justify decisions on fundamental political issues to each other using publicly available values and standards. In such contexts, the expectation is that citizens seek to find reasons that are compelling to others (whom they acknowledge as equals); in doing so, citizens should be both aware that others have different commitment and equipped with some sense of what those commitments are. The kind of fundamental issues Rawls has in mind here include questions about which religions to tolerate, who has the right to vote, who is eligible to own property, and what are suspect classifications for discrimination in hiring. These are what Rawls calls constitutional essentials and matters of basic justice, including the principles of distributive justice themselves. Public reason functions differently from "comprehensive moral doctrines," which are comprehensive in providing guidance across the whole range of questions that arise in human interactions. The major world religions are examples. The standpoint of public reason only applies to the much more limited context of fundamental issues about which citizens are expected to debate *as* citizens.

The crucial conflicts about such fundamental issues that societies face today are, first, disputes about how the goods of economic production and other advantages from interconnected life in modern societies are shared out, and then also disputes about the degree to which the rules of interaction favor conflicting conceptions of the good. The latter kind of conflict arises because public spaces must be shared among adherents of different doctrines with deep metaphysical and epistemological disagreements. After millennia of intellectual investigations and disputes (and now also with a deeper understanding of moral and evolutionary psychology that teaches much about how values enter people's lives), we realize that people inevitably and enduringly interpret human experiences differently. It is a significant step to think about conflicts that arise in living arrangements the way Rawls does: as clashes among competing ideas about both comprehensive doctrines and the sharing out of societal advantages that are unavoidable but not profoundly problematic if managed correctly, through an appeal to suitable conceptions of public reason and distributive justice.

This Rawlsian view is our starting point. It makes regular appearances throughout and thus brings coherence to the outlook presented here. It provides useful guidance for the digital age. But the subject matter demands that many other authors and viewpoints also enter, including some that normally do not appear in inquiries framed by this Rawlsian outlook. Again, this book deals with questions that could not be on Rawls's radar or on the radar of the generation of scholars after him that expanded the outlook. These questions were not much on my own radar when I took Rawls's theory as a starting point for perspectives on global justice.[14]

[13] Rawls, *Political Liberalism*.
[14] Risse, *On Global Justice*; Risse, *On Justice*.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

So, despite offering that starting point again in this book, there are chapters in which Rawlsian theory does not enter at all.

Comparisons to Marx have already made an appearance. And while I use the Marxist tradition sometimes in contrast and often as a corrective to the liberal-egalitarian tradition, all this is instructive only because these traditions share quite a bit. One commonality is the conceptualization of society as an intensely interconnected endeavor with cooperative and coercive elements that must be arranged the right way (and whose proper arrangement can be understood from a suitable theoretical angle). In addition, Marx himself also acknowledges a recognizable public-reason standpoint. We can see this by looking at his important 1843 essay "On the Jewish Question."[15]

That essay responded to contemporary writer Bruno Bauer's claim that it was incoherent for Jews to ask for religious toleration. Bauer argued that, since religious toleration granted equal standing to multiple religions, Jews could demand toleration only if they saw Judaism as one religion among others. (A similar point would then apply to other religious or secular views about the good life, though that was neither Bauer's nor Marx's concern.) Marx replied that asking for toleration only implied accepting that the difference among views of the good life was politically and legally irrelevant. This standpoint involves distinguishing between what Marx calls "the point of view of a man" and "the point of view of a citizen." It is from the former (suitably generalized across genders) that one would make decisions about one's life. The point of view of the citizen should be adopted when making certain decisions in political life, such as, in Rawls's view, decisions about constitutional essentials and questions of basic justice. Marx's point of view of the citizen, accordingly, is a public-reason standpoint.[16]

3 SUMMARY OF THE BOOK, CHAPTER BY CHAPTER

1 *Introduction: Digital Lifeworlds in Human History*

This chapter takes stock of the current situation confronting political theory. I introduce the concept of *digital lifeworlds* and explain its relevance in the narrative of humanity. I use Max Tegmark's distinctions between Life 1.0, 2.0, and 3.0, respectively, for guidance in locating digital lifeworlds in history. We do not know

[15] McLellan, *Karl Marx: Selected Writings*, 46–70.

[16] To be sure, the similarities only go so far, since the second main point of Marx's essay is to insist that the political emancipation that comes from the adoption of the standpoint of the citizen is insufficient for genuinely human emancipation. For this standpoint left individuals divided within themselves (between those two outlooks) and separated from each other by differences of religion and class. True human emancipation, for Marx, requires elimination of both forms of alienation. By contrast, Rawls accepts such conflicts as unavoidable features of modern life. Chapter 8 returns to "On the Jewish Question." For discussion of the relationship between Marx and Rawls, see Scanlon, "Some Main Points in Rawls' Theory of Justice."

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

if Life 3.0 (the kind of life that designs both its culture and physical shape, the physical shape of individuals) will ever arise. But if it does, it will be from within digital lifeworlds – lifeworlds that already fundamentally change our lives and thus require intense scrutiny even if there will never be a Life 3.0. To understand these lifeworlds, we need appropriate notions of "data," "information," and "knowledge" and characterize the connections among them. To that end, we enlist Fred Dretske's understanding of knowledge in terms of flow of information. Such a notion of knowledge allows for a broader range of knowers than humans (to whom classical analyses were limited): It includes both animals and artificially intelligent beings as knowers. I also draw on Luciano Floridi's work on the philosophy of information for a related look at digital lifeworlds from a more detached standpoint ("infospheres populated by inforgs").

2 *Learning from the Amish: Political Philosophy As Philosophy of Technology in the Digital Century*

The Amish are an unusual case of a community intensely concerned with maintaining control over how technology shapes its future. Though the community's old-fashioned ways strike many people as perplexing, in the age of AI, there are good reasons as to why technology and its regulation should be just about as central to mainstream politics as they are to the way the Amish regulate their affairs. Technology is not neutral, as many still think, but is intensely political. This also means that political philosophy and philosophy of technology should be more closely related than they typically are. In fact, mainstream philosophy of technology has unfolded largely separately from mainstream political philosophy. The primary exception is the Marxist tradition that has long investigated the role of technology in the dialectical unfolding of history as Marx theorizes it (including, in the case of figures such as Herbert Marcuse, an investigation of the use of technology in the thwarting of Marx's own predictions). This chapter uses the Marxist tradition to identify three senses in which technology is political (the foundational, enframing, and interactive senses) and argues that the Rawlsian tradition also has good reason to recognize versions of these senses. In an era of AI and other technological innovation, political philosophy must always also be philosophy of technology.

3 *Artificial Intelligence and the Past, Present, and Future of Democracy*

Modern democracies involve structures for collective choice that periodically empower relatively few people to steer the social direction for everybody. As in all forms of governance, technology shapes how this unfolds. Political theorists have typically treated democracy as an ideal or an institutional framework, instead of

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

considering its *materiality*, the manner in which democratic possibilities are to some extent shaped by the objects needed to implement them. Specialized AI changes the materiality of democracy, not just in the sense that independently given actors now deploy different tools. AI changes how collective decision-making unfolds and what its human participants are like.

This chapter reflects on the past, present, and future of democracy and embeds into these basic reflections an exploration of the challenges and promises of AI for democracy in this digital century. We explore specifically how to design AI to harness the public sphere, political power, and economic power for democratic purposes. Thereby, this chapter also continues the discussion from Chapter 2 by developing how technology is political in the foundational sense. This chapter also investigates current questions about how AI could threaten or enrich the democratic processes of the present. Only in Chapter 11 do we ask if superintelligences might in due course themselves be part of a democratic process. Nonetheless, many themes of subsequent chapters already make an appearance here.

4 *Truth Will Not Set You Free: Is There a Right to It Anyway? Elaborating on the Work Public Reason Does in Life 2.0*

We first explore how damaging untruth can be, especially in digital lifeworlds. Digital lifeworlds generally create possibilities for spreading information at a pace and volume unheard of in analog contexts. But misinformation and disinformation spread the same way, which greatly enhances how individuals can tell stories about themselves or have them substantiated in echo chambers in the company of like-minded people. This set of considerations provides support for a right to truth. However, next we see that untruth is immensely important to people's lives. It is not just that people fail to have a preference for truth, but rather that untruth plays a significant role as an enabler of valued psychological and social dynamics. The considerations that pull into the opposite direction notwithstanding, there can therefore be no comprehensive right to truth.

Contrary to a well-known Bible verse, for most people it is anyway not the truth that sets them free. It is acceptance of worldviews in likeminded company that does so (worldviews, or comprehensive moral doctrines, which tend to contain plenty of untruths), in any event if being set free means having an orientation in the world. But that there can be no such *comprehensive* right is consistent with there being a right to truth *in specific contexts*. And to be sure, protecting the public sphere (as introduced in Chapter 3) for the exercise of citizenship from a public-reason standpoint means the state must protect truth telling and sanction untruth. But typically, as we conclude, the moral concern behind truthfulness is in this context not best captured in terms of an actual right to truth.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

5 *Knowing and Being Known: Investigating Epistemic Entitlement in Digital Lifeworlds*

Michel Foucault problematizes the relationship between knowledge and power in ways that more traditional epistemology has not, with power always already shaping what we consider knowledge. To capture the nexus between power and knowledge, he introduces the term "episteme" (for the totality of what is considered knowledge at a given time, how it is obtained, by what rules it is structured, and how all that is shaped by power relations of sorts). The significance of an era's episteme is easiest to see in terms of what it does to possibilities of self-knowledge. Therefore, I pay special attention to this theme by way of introducing the theoretical depth of Foucault's notion. I then develop Foucault's ideas further, specifically for digital lifeworlds.

With this vocabulary in place, I introduce the notion of *epistemic actorhood* that lets us capture the place of an individual in a given episteme. It is in terms of this place that we can turn to the notions of epistemic rights and epistemic justice. Epistemic actorhood comes with the four roles of individual epistemic subject, collective epistemic subject, individual epistemic object, and collective epistemic object. Using this vocabulary, we can then also articulate the notions of an epistemic right and of epistemic justice and develop them in the context of digital lifeworlds. Digital lifeworlds engage individuals both as knowers and knowns in new ways. The framework introduced in this chapter captures this point.

6 *Beyond Porn and Discreditation: Epistemic Promises and Perils of Deepfake Technology*

Deepfakes are a new form of synthetic media that broke upon the world in 2017. Bringing photoshopping to video, deepfakes replace people in existing videos with someone else's likeness. Currently most of their reach is limited to pornography and efforts at discreditation. However, deepfake technology has many epistemic promises and perils, which concern how we fare as knowers and knowns. This chapter seeks to help set an agenda around these matters to make sure that this technology can help realize epistemic rights and epistemic justice and unleash human creativity, rather than inflict epistemic wrongs of any sort. In any event, the relevant philosophical considerations are already in view, even though the technology itself is still very much evolving. This chapter puts to use the framework of epistemic actorhood from Chapter 5.

7 *The Fourth Generation of Human Rights: Epistemic Rights in Life 2.0 and Life 3.0*

British science fiction writer and social activist H. G. Wells was a major advocate for a universal declaration of human rights of the kind that was later passed in 1948.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

Wells paid much attention to the importance of knowledge for his era, more than found its way into the actual declaration – though, to be sure, recognizable epistemic rights do play a role in the human rights movement. However, at this stage in history, an enhanced set of epistemic rights that strengthen existing human rights – as part of a *fourth generation* of human rights – is needed to protect epistemic actorhood in those four roles introduced in Chapter 5.

Epistemic rights are already exceedingly important because of the epistemic intrusiveness of digital lifeworlds in Life 2.0, and they should also include a suitably defined right to be forgotten (that is, a right to have certain information removed from easy accessibility through internet searches). If Life 3.0 does emerge, we might also need a right altogether different from what is currently acknowledged as human rights, the right to exercise human intelligence to begin with. This right will become important again in Chapter 11. Human rights must expand beyond protecting "each of us from the rest of us" to protecting "us from them," much as such protection would have to prevail conversely. The required argument for the validity of the right to the exercise of human intelligence can draw on the secular meaning-of-life literature. I paint with a broad brush when it comes to the detailed content of proposed rights, offering them manifesto-style as the Universal Declaration of Human Rights (UDHR) does.

8 *On Surveillance Capitalism, Instrumentarian Power, and Social Physics: Securing the Enlightenment for Digital Lifeworlds*

"Surveillance capitalism" is a term coined by Shoshana Zuboff to draw attention to the fact that data collection has become so important for the functioning of the economy that the current stage of capitalism should be named for it. "Instrumentarian power" is a kind of power that becomes possible in such an economic system, power that deploys technology to obtain ever more knowledge about individuals to make their behavior predictable and thus monetizable. "Social physics" is a term used by computer scientist Alex Pentland to describe the potential of quantitative social science to put Big Data to beneficial use. The primary goal of this chapter is to discuss how surveillance capitalism in digital lifeworlds threatens the Enlightenment ideal of individuality itself (as discussed by Kant and Durkheim) and what it takes to secure the Enlightenment for digital lifeworlds. That is, I draw on democracy and epistemic rights to discuss how Enlightenment ideals can be secured in such lifeworlds.

Since this chapter is the last in a row of chapters concerned with rights, we also discuss (and reject) the position that rights, especially human rights, are enough to articulate a promising normative vision for society. This discussion draws on insights from Max Horkheimer and Theodor W. Adorno's *Dialectic of Enlightenment*, which has synergies with Zuboff's work. It is important to be clear on the scope and limits of a rights-based vision for society since such visions have become prominent in our neoliberal world. But contrary to such a neoliberal understanding,

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

a strong view of democracy, as discussed in Chapter 3, is also required for a promising normative vision for society. And so is a plausible theory of distributive justice, a subject to which we turn in Chapter 9.

9 *Data As Social Facts: Distributive Justice Meets Big Data*

In the age of Big Data and machine learning, with its ever-expanding possibilities for data mining, the question of who is entitled to control the data and benefit from the insights that can be derived from them matters greatly for the shape of the future economy. Therefore, this topic should be assessed under the heading of distributive justice. There are different views on who is entitled to control data, often driven by analogies between claims to data and claims to other kinds of things that are already better understood. This chapter clarifies the value of approaching the subject of control over data in terms of (a notion of moral, rather than legal) ownership. Next, drawing on the work of seventeenth-century political theorist Hugo Grotius on the freedom of the seas, and thus on possibilities of owning the high seas, I develop an account of collective ownership of collectively generated data patterns and explore several important objections. This chapter also connects to my earlier work on distributive justice in *On Justice* and *On Global Justice*.

That a seventeenth-century figure would appear as we try to throw light on a twenty-first-century problem might startle or even irritate. But Grotius's account of the ownership of the seas formulates basic and immensely plausible ideas about what kind of thing should or should not be privatized. It offers lessons for our current debate. Since control over data matters enormously and is poorly understood, we should treat questions about it as *genuinely open*. This is a good time to bring to bear unorthodox thinking on the matter. Chapter 5 introduced epistemic justice, and Chapter 7 introduced new epistemic human rights. Both themes reenter here.

10 *God, Golem, and Gadget Worshippers: Meaning of Life in the Digital Age*

The question of the meaning of life is about how a human life is connected to other things and themes of value around it. Meaning of life and technology are not normally theorized together. But once we realize that all human activity is always technologically mediated, we see that any acts in pursuit of personal significance, too, are so mediated. However, this point then opens the possibility that technology enters the human quest for meaning *the wrong way*. This chapter explores what that possibility means and how to respond to it.

I use as my starting point Robert Nozick's proposal for how to think about the meaning of life. Nozick's account makes central the idea of "limited transcendence," essentially folding the kind of transcendence normally involved in interaction with divinity into a finite life. This understanding of meaning receives much plausibility from the fact that finite lives are the only sources of meaning open to

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

humans. Nozick's high-altitude view does not make sufficiently clear how technology enters. But once we bring in additional ideas from Don Ihde and Hannah Arendt, we can see clearly how it does. Next, we turn to Norbert Wiener's classic *God & Golem*. Wiener is concerned with "gadget worshippers," people who surrender control over their lives to machines in ways that are not appropriate to what these machines can do. Working with this notion, we can throw light on how technology can enter into the quest for meaning the wrong way and offer some advice for how to counterbalance that challenge.

11 *Moral Status and Political Membership: Toward a Political Theory for Life 3.0*

I introduce a distinction between "slow and relatively harmonious" and "fast and radical" as far as the integration of AI into human life is concerned. Regarding the "slow and relatively harmonious" scenario, I explore a set of questions about how it would make sense for humans to acknowledge some such status in machines (in a variety of ways). But we must then also ask whether self-conscious artificial intelligences would be fully morally equivalent to humans. I explore that issue by asking what an increase in moral status for machines would mean for the political domain. Chapter 3 already explored why AI would affect the democratic process in the near future. Here our concern is with a scenario further along when questions around political membership of intelligent machines would actually arise. One question is whether there is a cognitive capacity beyond intelligence and self-consciousness that is needed for involvement especially in the political domain. Paying attention to what is appropriate to say about animals in that regard turns out to be useful. I turn to Christine Korsgaard's as well as Sue Donaldson and Will Kymlicka's recent discussions of animals to investigate the matter.

As far as the "fast and radical" scenario is concerned, I first explore why philosophically we are so dramatically unprepared to deal with an intelligence explosion, with a focus on what kind of moral status a superintelligence might acknowledge *in us*. Finally, I attend to Tegmark's discussion of political scenarios that could arise after an intelligence explosion and add a public-reason scenario that, under certain circumstances, could offer a vision for a political context genuinely shared between humans and superintelligent machines.

4 A NOTE ON OMISSIONS

One set of topics I do not cover at length – though the topics do make regular appearances – is how inequality and fairness across groups in society are affected by AI and Big Data. The ways in which algorithms perpetuate patterns of discrimination and exacerbate economic inequality has become the single most covered issue in this field in recent years. Monographs authored by Ruha Benjamin, Safiya Noble,

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

Cathy O'Neil, Virginia Eubanks, Joshua Simons (whose work is forthcoming), and others and a substantial number of articles have advanced the debates around these topics in ways for which there is no counterpart among the topics that I do cover at length. And to be sure, my Rawlsian outlook (and my own development of it for the global domain in *On Global Justice* and *On Justice*) already comes with views on inequality that carry over into the digital century.

One other omission of sorts that I should note is that this book does not seek to reach a bottom-line judgment about the existential risks posed by the development of AI.[17] We do encounter several large-scale dystopian scenarios throughout that offer dire assessments of the human use of technology, issued by authors such as Martin Heidegger, Lewis Mumford, Jacques Ellul, and Herbert Marcuse. We also occasionally touch on the more recent literature on existential risk. But for one thing, I have not been able to make up my mind about whether it is appropriate to call for a radical stop to or a temporary moratorium on technological advancement. Certainly, the aforementioned dystopian scenarios – while indeed insightful enough to offer warnings – are too bombastic in scope to allow for conclusive validation. At the same time, I take it as a given that technological advancement will continue anyway for the time being, one way or another, if only because of geopolitical rivalry. The task for political theory, then, is to think about the topics that will likely come our way, distinguish among the various timeframes (such as Life 2.0 and Life 3.0) in which they might do so, and make proposals for how a democratic society should prepare itself to deal with the changes in the technology domain that it might eventually have to address.[18]

[17] For a call for a moratorium on work that might lead to the creation of artificial consciousness, see Metzinger, "Artificial Suffering."

[18] Note that I capitalize "Chapter" when talking about specific chapters in this book (as in "Chapter 1"), but not when talking about chapters in other books. In some situations, that will help avoid confusion.

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

Acknowledgments

I am grateful to audiences at the University of Hamburg; the Wharton School of the University of Pennsylvania; the University of Tokyo; Harvard Law School the Universities of Freiburg, Heidelberg, and Cologne; University College London; the University of Alabama at Birmingham; Massachusetts Institute of Technology (MIT); and Northeastern University for discussions of some of this material over the years. I am grateful to my coauthors Catherine Kerner and Steven Livingston for earlier work on some of these issues. I owe thanks to Diana Acosta Navas, John Basl, Elettra Bietti, Gabriella Blum, Matthew Braham, Maria Carnovale, Hamza Chaudhry, Austin Choi-Fitzpatrick, Christoph Durt, Linda Eggert, Sam Gilbert, John Goldberg, Amanda Greene, Wilfried Hinsch, Philip Howard, Sheila Jasanoff, Peter Niesen, Sushma Raman, John Shattuck, Silja Vöneky, and Jim Waldo either for comments on parts of this material or for organizing events at which some of it was discussed (or both). Philosophical interactions with Tim Scanlon over the years have found their way into this project in more ways than might be apparent.

I owe much of my inspiration to the members of the Human Rights and Technology Fellowship at the Carr Center for Human Rights Policy at the Harvard Kennedy School (which I am fortunate enough to direct, alongside the Carr Center's formidable executive director, Sushma Raman) for ongoing discussions of the normative dimensions of technological innovation. I am also indebted to the speakers in our Toward Life 3.0 webinar series, including Nikita Aggarwal; Jay Aronson; Ron Deibert; Teresa Hodge and Laurin Leonard; Sheila Jasanoff; Alexa Koenig; Philip Howard; Sabelo Mhlambi; Safiya Noble; Frank Pasquale; Bruce Schneier; John Tasioulas; Shoshana Zuboff; and others.

I am also very grateful to the members of the Graduate Fellowship Colloquium at the Edmond J. Safra Center for Ethics (which I am fortunate enough to have directed over time jointly with Danielle Allen, Eric Beerbohm, and Meira Levinson, respectively) for discussions of some of my work and for creating often highly insightful cross-connections. And I am grateful to the participants in my course on the philosophy of

© in this web service Cambridge University Press & Assessment www.cambridge.org

CAMBRIDGE
Cambridge University Press & Assessment
978-1-009-25521-9 — Political Theory of the Digital Age
Mathias Risse
Frontmatter
More Information

technology that I have offered in recent years to students at the Harvard Kennedy School, the Harvard Graduate School of Design, and Harvard College for the many insights they have offered, and especially to my teaching fellow Cat Wade for doing such an amazing job twice with this course (as well as to course assistants Grace Stone and Stephen Dwyer for being excellent in their roles). As it has turned out, insights from my course on the meaning of life, which I have taught many times in various versions in different locations, have also been invaluable for this project. I am grateful to all who have contributed to this course over the years.

The penultimate version of the manuscript benefited enormously from Katherine (Kate) Williams's keen eye for grammatical subtleties and from her high standards of clarity of expression. Every chapter became noticeably better after Kate made her way through it. I am much indebted to Matt Gallaway at Cambridge University Press for approaching me several years ago about writing a book on this subject, and for his enduring interest in it over the years. Four anonymous reviewers for the press gave useful advice when the direction of this project needed it most.

In Chapter 6, I use material from "Beyond Porn and Discreditation: Epistemic Promises and Perils of Deepfakes Technology in Digital Lifeworlds" (coauthored with Catherine Kerner), *Moral Philosophy and Politics* 8, no. 1 (2021): 81–108. In Chapter 7, I use material from "The Fourth Generation of Human Rights: Epistemic Rights in Digital Lifeworlds," *Moral Philosophy and Politics* 8, no. 2 (2021): 351–78. I am grateful to De Gruyter publishing company for permission to use this material here. Throughout I use bits and pieces from "Human Rights and Artificial Intelligence: An Urgently Needed Agenda," *Human Rights Quarterly* 41 (2019): 1–16. I am grateful to Johns Hopkins University Press for permission to use this material.

© in this web service Cambridge University Press & Assessment www.cambridge.org