# Economies of Virtue

# Thao Phan, Jake Goldenfein, Declan Kuch, and Monique Mann

What does ‘AI ethics’ do? Who does it serve? What is its purpose? If we ask those invested in questions of technology and social, political, and environmental justice, or those engaged in on the ground inquiry with communities affected by automated decision-making, or working inside of regulatory institutions, the answer we might expect is ‘nothing at all.’ The fast-moving wave of critique that defined the ‘‘tech lash’’ is now being subsumed by an even larger wave of what Elettra Bietti calls ‘‘ethics bashing.’’[[1]](#footnote-1) Where the tech lash centered on exposing the bad behavior of some of the world’s largest tech corporations, this next wave addresses the rise of ‘ethics’ as a new industrial agenda, focusing on actions such as the establishment of AI ethics boards, hiring AI ethics teams, and funding AI ethics research. Critics have fairly interpreted these actions as efforts to instrumentalize ethics,[[2]](#footnote-2) to reduce it to another form of industrial capital,[[3]](#footnote-3) or to co-opt and capture researchers as part of efforts to control public narratives.[[4]](#footnote-4)

Many critical accounts aim to bring transparency and accountability to fields that are themselves concerned with questions of transparency and accountability in AI systems. The turn within AI ethics conferences, most notably the Association of Computing Machinery’s annual Fairness, Accountability, and Transparency conference (FAccT), to include whole streams that turn the gaze of inquiry back onto the organisation and its values, its processes of governance, and its sponsorship policies, signals a significant reflexive turn within these communities.[[5]](#footnote-5) The intended audience for these discussions are people who, in many cases, work under the broad umbrella of AI ethics. Although an increasing number of scholars and activists now seek to disavow and indeed challenge both the terms ‘AI’ and ‘ethics’, for those who work outside the academy—the people who do not spend their days on Twitter, reading articles on Medium, Wired or the MIT Technology Review—‘AI ethics’ is broadly what the general public think we do. If we consider ethics as ‘an internalised aspirational mode of enquiry that aims at a better world’ and ‘a more just society,’[[6]](#footnote-6) then many would not object to this title. Rather the objection stems from what ethics has become. In short, what it is made to do, who it serves, and the purposes it *does*, rather than *should*, fulfil.

When posed with this question, ‘what does AI ethics do?’, the answer, in reality, is very rarely ‘nothing at all.’ As the growing corpus of critical literature shows, the ‘ethics’ of AI ethics is much more malleable, more pliable, and more amenable to strategic operationalization than anticipated. Ethics, it’s said, can be washed[[7]](#footnote-7) and bashed.[[8]](#footnote-8) It can operate as a rubber stamp[[9]](#footnote-9) or an empty gesture.[[10]](#footnote-10) It can be a fig leaf,[[11]](#footnote-11) a seductive diversion,[[12]](#footnote-12) or, as Lily Hu writes in her searing analysis, it can be a performative theatre: ‘Silicon Valley, with its long financial strings, plays the tech ethics marionette; “ethics” is a show, and they know it.’[[13]](#footnote-13) Put simply, ethics as deployed by Big Tech does worse than nothing. It is divisive, contested, and more often than not enables the troubling practice of ‘business as usual.’ At its worst, AI ethics is not just useless but ‘dangerous, hoarding expertise and funding that should be devoted to more effective work.’[[14]](#footnote-14) For these reasons, celebrated tech journalists like Karen Hao[[15]](#footnote-15) have scolded the AI ethics community:

[t]alk is just that—it’s not enough. For all the lip service paid to these issues, many organizations’ AI ethics guidelines remain vague and hard to implement. Few companies can show tangible changes to the way AI products and services get evaluated and approved. We’re falling into a trap of ethics-washing, where genuine action gets replaced by superficial promises.

This collection takes its point of departure from here. In our initial call for contributions, we invited scholars and activists to respond to the growing trend of AI ethics bashing by focusing specifically on the financial structures that support AI ethics and its effects on the people who perform this work. We—the editors of this collection—asked: who funds research into the ethics of AI technologies? How might these funding arrangements cement or exacerbate hierarchies of power? And how does an industry-sponsored agenda on ethics impact the production of knowledge about AI systems?

In our call, we suggested that in response to campaigners, industry insiders, and civil society actors raising concerns about ‘bad AI,’ there was now a wave of social science–led research around AI ethics. Joining the growing chorus, we expressed concern that, as the field of inquiry and practice grew, it too was at risk of an ethical crisis, one that brought into relief the complex positionality of researchers in this area. The social and political pressures that initially compelled companies to behave more ethically has since morphed into varying infiltrations of research culture in the form of conference sponsorship, formal research partnerships, joint industry-academic appointments, and more.[[16]](#footnote-16)

This influx of industry funding is, of course, a symptom of a broader set of problems: the politicization of tertiary education, the cultivation of a hostile culture towards a perceived ‘left intellectualism,’ and the one–two neoliberal punch of funding withdrawal for public institutions the world over, coupled with growing incentives to acquire external industry funding. In Australia, for instance, a broken funding system has meant that 1) the few government grant sources that are available offer less than they did a decade ago and are mostly accessible to already wealthy institutions, 2) the grants are primarily structured to incentivise research deemed ‘commercializable’ or of value to industry, and 3) even if researchers make it through the highly competitive process,[[17]](#footnote-17) project success can still be subject to political interference.[[18]](#footnote-18) The slow withdrawal of public funding for public tertiary education has meant that universities must turn to other forms of income for survival. Again, to use the example of Australia, universities now approach education as a secondary enterprise to the real business of managing investment portfolios alongside other forms of profiteering, hiring former industry executives to manage universities as they would a large corporation. The result is that universities value teaching and research to the extent that they can be translated into metrics informing international rankings, which can then be cashed out in international student fees. At the same time, the trickle-down effect of financial discipline pressures universities to extract as much from workers as possible, including systematic wage theft from staff.[[19]](#footnote-19)

These factors place AI ethics in the awkward position of being financially dependent on the organisations they seek to hold to account. As Hu argues, ‘[i]t is an uncomfortable fact that however much external advisory boards and universities claim to be ‘third parties’, ethical tech institutions are in fact parasitic on the continual moral failures and disappointments of a hegemonic tech industry.’[[20]](#footnote-20) For Hu, these relationships are mutually reinforcing, with corporations using AI ethics as means to dissimulate political demands for change, and tech advocacy groups capitalising on the clout and resources that industry partnerships afford.[[21]](#footnote-21) STS scholar Lee Vinsel has referred to this phenomena as ‘criti-hype’—the kinds of self-serving criticism ‘that both feeds and feeds on hype’ concerning emerging technologies.[[22]](#footnote-22)

Indeed, this volume itself began its life from a small research grant ($9,000 AUD) targeted at early career researchers in Australia. This ‘workshop programme grant’ was funded by the Academy of the Social Sciences in Australia, an organization supported primarily by Australian Commonwealth grants and Fellowship subscriptions.[[23]](#footnote-23) As early and mid-career researchers ourselves, the institutional expectations to attract funding neatly collided with the realization that topics on AI, and especially the ethical dimensions of AI, were eminently fundable. While we were always motivated by our commitment to earnestly explore the topic of funding, labour, and AI ethics, the practical demands of requiring resources and institutional support were never far away.

As researchers based in the humanities and social sciences, we were lucky in that the questions we sought to answer did not require expensive hardware, compute power, access to large datasets, or server space—requirements that often force the hand of many of our colleagues in STEM faculties. Rather, our demands were more rudimentary: time. Both *our* time and the time of the various interlocutors we have engaged with throughout the lifespan of this project. In most circumstances, funding can do more than secure resources. It can also work to accumulate clout and authority, bestowing legitimation and accolades on those who receive it. With funding, we could pay for catering and copyediting, as well as administrative and research assistance to host workshops and develop publications. But in its most valuable form, the funding gave us the prestige necessary to ask for others to donate their time to our project and to give ourselves permission to take our own time back from the institution.

We mention this here in some detail because it is precisely these kinds of mundane, practical negotiations and institutional demands that dictate the conditions of labour for those engaged in the work of AI ethics. Rampant forms of precarity, insecurity, and normalized cultures of exhaustion and overwork have meant that these kinds of reflexive and critical conversations can rarely take place unless they can also be mobilized in service of strategic goals and personal metrics. These pressures are palpable in areas like the humanities and social sciences, who have been on the receiving end of faculty funding cuts and who have, for decades, struggled against a conservative-led culture war.[[24]](#footnote-24) In this context, it’s no surprise that many turn to profitable topics, like AI ethics, and to industry-funded grants as a means for survival.

It’s under these disheartening conditions inside of universities that Big Tech succeeds, in the famous words of Fred Moten and Stefano Harney, in ‘[turning] insurgents into state agents.’[[25]](#footnote-25) That ‘AI ethicist’ has now become a legitimate job title precisely illustrates Moten and Harney’s argument. Former critics now make their wage through teaching specialized courses and degrees on the social, political, and ethical dimensions of AI, producing a new generation of professional technocrats that ostensibly do AI better than the generation previous. As Ben Tarnoff has incisively argued, one of the unexpected effects of the techlash has been ‘a mass credentialing event for a new class of experts as “AI ethics,” “responsible innovation,” and similar pursuits attract significant funding and visibility.’[[26]](#footnote-26) These forms of professionalization and credentialing only serve in the interests of corporate actors who were once the targets of critique. To riff on Moten and Harney, AI ethics is ‘more than an ally’ to Big Tech’s corporate agenda, ‘it is its attempted completion.’[[27]](#footnote-27)

In the following sections, we turn more explicitly to the themes and concepts that underpin this anthology. We begin with a brief account of the rise of AI ethics before discussing what we see as the contemporary reification of ethics into a commodity form. We then introduce each of the chapters, outlining the ways in which they illustrate the vast and variegated network of circulations that define what we call an ‘economy of virtue’ before closing with a brief discussion on the trajectory for AI ethics from here.

## The Rise of AI Ethics

The term ‘tech-lash’ entered the mainstream in 2013, with *The Economist*’s Adrian Wooldridge suggesting the public mood was shifting against Silicon Valley’s tech elite.[[28]](#footnote-28) Public perceptions of ‘Big Tech’ were shaken by the industry’s growing contributions to material inequality, failed (and decidedly untrendy) consumer products like Google Glass, repeated privacy failures, unjustified tax breaks, and problematic politics.[[29]](#footnote-29) Google’s effort to ‘organize the world’s information,’[[30]](#footnote-30) for instance, was reinterpreted as harboring the capacity to produce profound social harm. Analysts were realizing that companies once seen as standard-bearers for liberatory rhetorics like the ‘open internet’ had now leveraged their control over online services and data flows into worrying forms of profiteering and domination. Far from facilitating a revolutionary break from the prevailing socio-economic situation, Silicon Valley’s promises of liberation were exposed as just another expression of industrial capitalism. Within the academy, critical scholars had been describing the contradictions between Silicon Valley’s counter-cultural self-presentation and its relentless pursuit of capital for some time.[[31]](#footnote-31) But the legitimation of AI ethics as a field of inquiry, alongside and as part of the tech-lash, incentivised additional high-profile academics across social and technical disciplines to produce new critical analyses of tech industry services and structures.

One of Big Tech’s primary products—machine learning (ML)—was at the forefront of this re-evaluation. There was growing recognition that it frequently generates biased outcomes, deeply affecting the lives of marginalised people. In 2013, mainstream news began reporting on how ML technologies, by virtue of being trained on historical (and historically biased) data, reproduced forms of gendered and racialised discrimination.[[32]](#footnote-32) Awareness was growing around the mistakes, misrecognitions, and problematic profiling performed by algorithmic scoring systems in domains like employment, social services, and policing.[[33]](#footnote-33)

With clear capacities for harm and massive commercial interest, machine-learning became the set of techniques around which AI ethics oriented itself as a field. Early technical work on fairness in machine-learning classification, for instance by Cynthia Dwork et. al. (including Moritz Hardt, an early organizer of the FATML conference),[[34]](#footnote-34) explored ways to do ‘fair’ classification, drawing on notions of equality and fairness in the political theory of H. Peyton Young, John Roemer, and John Rawls. Once the political and social relevance of these technical issues was exposed, questions of fairness and bias were quickly targeted by the humanities and social sciences.

In March 2013, Latanya Sweeney, Founder and Director of the Harvard University Data Privacy Lab, published an influential paper on racial discrimination in online ad delivery, bringing critical attention to how commercial platforms reproduced patterns of racial discrimination via their ad delivery system.[[35]](#footnote-35) Soon after, Kate Crawford, then a Principal Researcher at Microsoft Research, published ‘The Hidden Biases in Big Data’ in *The Harvard Business Review*.[[36]](#footnote-36) Crawford framed fixing (big) data science as the next frontier in studying the relationship between technology and society, recommending technologists interface with social scientists and that computational sciences come to terms with qualitative methods. In *The Atlantic*, Nicholas Diakopoulas connected the problem of bias with that of opacity in machine learning.[[37]](#footnote-37) Additionally, Safiya Noble helped bring many of these issues regarding data-driven systems, commercial interest, and racial bias together in her research highlighting the racial and gender bias in search engines.[[38]](#footnote-38) In August 2014, Solon Barocas and Andrew Selbst made available their article ‘Big Data’s Disparate Impact,’[[39]](#footnote-39) intricately outlining the ways big data technologies generated discriminatory treatment in legally meaningful ways.

Around 2014, civil society also turned its attention to social justice concerns associated with big data and AI. The Leadership Conference on Civil and Human Rights, for instance, published its ‘Civil Rights Principles for the Era of Big Data’ in February 2014,[[40]](#footnote-40) outlining the need for accountability and fairness to tackle embedded discrimination. That agenda was given more specificity by organisations like Upturn,[[41]](#footnote-41) who connected the potential harms of data driven decision-making with broader issues of social justice. Shortly after, Obama’s White House (the Executive Office of the President) published its 2014 report on Big Data, identifying bias and discrimination in big data as risks of significant material and immaterial harms.

Concerns around algorithmic bias and fairness in machine learning were also connected with related scholarly interests in algorithmic accountability. New York University hosted a conference on Governing Algorithms in May 2013, asking how to turn the ‘problem of algorithms into an object of productive inquiry?’[[42]](#footnote-42) Data & Society also articulated a research agenda around algorithmic accountability in March 2014.[[43]](#footnote-43) In February 2015, NYU hosted its Algorithms and Accountability Conference, building on the 2013 event, and expressly outlining its topic as the challenges of algorithmic power in terms of transparency, fairness, and equal treatment, hosting speakers primarily from the humanities and social sciences. A more expressly technical forum had also coalesced as the 2014 FATML workshop, interested in exploring ‘how to characterise and address these issues with computationally rigorous methods,’[[44]](#footnote-44) and offering a venue for broadly technical solutions within a complex and inter-disciplinary problem definition. This workshop ran until 2018 and was supplemented the same year with the larger and more interdisciplinary FAT\* Conference (which in 2019 became affiliated with ACM, and changed its name to FAccT in 2020), bringing the social scientific and technical fields together. There are now a multitude of specialized AI Ethics workshops, professional organizations, and conferences, both standalone and as part of broader technical events, investigating issues such as algorithmic bias, the falsehood of technical objectivity, and solutions in the form of accountability, transparency, and fairness.

Despite AI Ethics emerging as a response to the failures of industry, Big Tech has long played a central role in supporting the field from its earliest days, and continues to participate in the scholarly environment with significant funding and support. Unquestionably, industry support of AI Ethics research has enabled a great deal of high quality, independent scholarship. But at the same time, it raises complex questions around the positionality of researchers and the institutional dynamics that define the ‘value’ of research.

Considering the commercial applications and tools that constitute the core subject of AI ethics analysis, it is no surprise that industry and academia have coalesced around these particular problems. However, it is precisely the development of networks of stakeholders with markedly different interests and values that has enabled the new forms of *circulation* described here as an ‘economy of virtue.’[[45]](#footnote-45) As Lily Hu points out with respect to the flagship AI ethics conference: ‘FAccT researchers are, generally-speaking, not shouting into the void; quite the opposite, many are in fact meeting at post-conference corporate-sponsored cocktail parties to discuss collaborations across institutions and interests.’[[46]](#footnote-46) Indeed, events like FAccT in essence operate as an interface between academia and industry, providing a platform that enables the consolidation of these groups and their competing interests. They create social and professional networks that can either lead to partnerships and collaborations or act as pipelines to direct employment in industry. The scale of these flagship events has also meant that they rely on the funding of Big Tech to subsidise the logistical costs of hosting. This creates particular obligations between professional societies and the organizations that they seek to hold to account; obligations that ensure doors always remain open to the flow of industrial interests.

These social and intellectual exchanges have grounded the reification of research, researchers, and reputations into commodity forms capable of circulating between and through industrial and academic institutions. These circulations take a number of shapes, drawing in, producing, and exchanging both explicitly financial as well as reputational forms of value for all involved. For instance, scholarly ethics outputs might become levers for acquiring industry funding, industry supported scholarly platforms, or industry appointments. At the same time, universities endorse these developments, trading on the influx of external money, prestige, and ‘impact’ associated with industry engagement. All the while Big Tech benefits from the enhancement and legitimation of their ethical credentials and endeavours.

Such circulations harbour the contradiction that high-value ethics commodities only acquire and sustain value through having limited ‘ethical effects.’ That is, in order to circulate freely between the academy and industry, the ethical content (in terms of effects on the world—or at least the business to which they are directed) must be effaced or at least constrained. As the abhorrent treatment of prominent AI ethics researcher Timnit Gebru by her former employer Google demonstrates,[[47]](#footnote-47) when the ethics commodity fundamentally challenges business models, the appointments and status offered in exchange disappear (and worse). It may be that the use-value of ethics commodities acquired by industry manifests elsewhere—not in the guiding of meaningful change in industrial and technical practice, but in the construction of reputational edifices capable of 1) shielding commerce from structural critique that could impact profit, and 2) incentivizing the AI ethics field’s progression in congenial directions. The multitude of forms ethics might take, the channels of its circulation, and the work that it does or does not do, as well as the ways its imperatives intersect with the experience of researchers, are precisely the subjects explored in the contributions to this edition.

## Contributions to this volume

This volume is a collective response to the reification of ethics into commodity forms. It explores how industry participation in ethical AI research has created a new economy of virtue—a massive network of actors variously situated across industry, civil society, and universities producing and circulating ethics as a service and as a product. The authors bring both critical perspectives and firsthand experiences of the challenges, dilemmas, and opportunities that life within this economy affords. Their experiences are diverse, hailing from a range of disciplinary backgrounds, including law, anthropology, criminology, media and communication studies, STS, political economy, and more. Where some of the authors are seasoned academics, professors with decades of experience in the field, others are at the beginning of their careers, entering industry at the peak of the AI ethics funding frenzy. They do, however, have a shared investment in issues of technology, power, and social justice, and while few (if any) would call themselves ‘AI ethicists,’ they are nevertheless all intimately intertwined with the controversies and debates that have followed the development of the field.

In the chapters that follow, these authors give voice and testimony to the tactics and strategies of commodification and resistance. Each chapter explores these dynamics as they unfold across different sites and terrains. When these stories are placed together, we begin to see common trajectories and flows between actors, institutions, and interests. It is a sad irony that the more ethics circulates as a commodity, the less ethical work it is able to do. Yet, as discussed above and as the chapters will illustrate, even in its commodity form, AI ethics can always be put to work to do something, to serve someone's interests.

This collection is arranged into three sections: subjects, sites, and actions. In the first section, the authors draw on their own subjectivities and subjective experiences to narrate the contradictions and dilemmas that these complex arrangements of funding place workers within.

In ‘Your Thoughts for a Penny? Capital, Complicity and AI Ethics,’ Corinne Cath and Os Keyes describe the industry sponsorship of PhD scholarships and collaborative projects as a cunning investment. They provide vignettes from their experiences as PhD students across the U.S. and the U.K., providing illuminating detail on how people on the ground navigate and negotiate the tensions and discomfort that can arise when one’s wage is tied to one’s enemy. ‘Critique does not avoid complicity,’ they write, and as their examples demonstrate, in many cases it operates as a mode of recuperation. They end with a call for feminist refusal, a position that helps situate researchers in relations of power by acknowledging that no engagement—however critical—is outside of it.

In ‘Extractivist Ethics,’ Sarah Pink describes ethics as ‘the bait through which trust in technology is extracted from publics or users.’ She argues that techno-solutionist approaches to design creates a disconnection between producers and everyday people, which in turn cultivates an instrumentalist approach to ethics. Human ethical values have purpose insofar as they can be used to make ‘ethical machines’ that can then be showcased to engender trust. She calls for a return to everyday ethics, which by their nature are slippery and unstable and therefore less amenable to the forms of capture and investment, at least as industry actors would imagine it.

In the final essay in this section, Rodrigo Ochigame describes the framing of ethics as a kind of ‘amicable criticism’ that ‘can serve as a ‘leverage for entering into business relationships.’ Like Cath and Keyes, he describes his time as a graduate student researcher working in the AI ethics groups the MIT Media Lab. He provides candid commentary on the ethical scandals that shrouded the lab and its former director, Joichi Ito, offering an unflinching assessment of the role of the Media Lab in sustaining the agendas of Silicon Valley. Originally published on the investigative journalism website *The Intercept*, Ochigame’s essay is a stunning example of speaking out against the sordid dynamics that most only hear through whispers. We are honoured to republish the essay with his permission here.

In the second section of this volume, ‘sites,’ we turn to situated case studies to understand how ‘ethical’ practice is leveraged by Big Tech within specific domains of AI application.

In ‘Ecocide Isn’t Ethical: Political Ecology and Capitalist AI Ethics,’ Sy Taffel, Laura Bedford and Monique Mann describe AI ethics as a way for corporations to frame ethical practice away from anthropogenic forms of planetary harm. They discuss in detail the social and ecological impacts of AI and ML, moving between sites of resource extraction, data centres, and e-waste processing. This contribution provides a damning critique of how the fantasy of ‘green AI’ operates to sustain unjust, unethical, and decidedly ecocidal corporate practices.

For Angela Daly, the corporate agenda driving AI ethics has made it abstract, disconnected and apolitical. Like Sarah Pink, she also calls for an analytic return to sites in everyday life to understand the complex lived realities of people on the ground and to highlight the inadequacy of abstracted ethical principles. Turning to the specific example of facial recognition, Daly departs from the hifalutin world of ethical principles and instead brings attention to a world of ethical negotiation that is enacted through protest, dissensus, and organizing.

In their chapter on global standards and standard-setting, Tsvetelina Hristova and Liam Magee outline how ethics is used as a vehicle for the socialization of risk, allowing it to scale up and be turned into a form of value. Their work expands the sites and practices typically associated with economies of virtue, turning to the highly bureaucratized zones of ISO (International Standards Organisation) frameworks and subcommittees. Here, they examine how middle-power countries approach standards setting as an economic and political strategy, the effect of which is the transformation of ethics into a literal exportable commodity.

Finally, Michael Richardson closes this section with his chapter on the quintessential site for the study of paradoxes in ethics: defence and military AI. He describes how a focus on ethics insulates researchers working with departments of defence from the squeamish questions of lethal violence while at the same time justifying the use of autonomous weapons to intensify impact and injuries. He writes, ‘Defence researchers and companies can not only *be* virtuous, but also can *make* war virtuous too.’ He underscores how decreases in state funding push universities to diversify income streams through corporate and military partnerships. ‘Universities,’ he reminds us, ‘are, after all, institutions of empire and colony even more than they are sites of learning, knowledge-making and dissent.’

The third and final section, ‘action,’ concludes the volume with two interviews with prominent scholars and activists, reflecting on moments of direct action against the politics of industry money and influence. Through these discussions we’re given a glimpse at the alternatives, into what activism might achieve, and the kinds of reflexivity that scholars need to understand the complex forces and imperatives shaping their working lives and subjectivities. In ‘Open Secrets’ with Meredith Whittaker, Jathan Sadowski and Thao Phan, and ‘Dropouts’ with Lilly Irani, Alex Hanna, J. Khadijah Abdurahman, and Jake Goldenfein, scholars with rich experience from across different positions in the Economy of Virtue describe the different ways institutions hijack our sense of self as researchers, and how our efforts to (re)imagine and/or (re)define the university may be better spent on organising ourselves as university workers as in any other industrial enterprise.

With their rich and diverse experience across sectors, these scholars describe the complexities of funding in the context of community, collegiality, and supporting people to earn a wage, as well as how they’re leveraged into tools of internalised discipline. With personal understanding of the internal mechanics of industry influence over the academy, and how alliances form around funding prerogatives, they describe the way these institutions encourage certain types of critical work, while simultaneously subverting radical efforts that might undermine industrial interests or funding relationships. The alternatives they offer is workplace organization as a tool to break scholars away from prerogatives of prestige that serve the interests of those with power over us, and in so doing to de-commodify the work performed by scholars in the economy of virtue.

## Conclusion: From AI Ethics to the Economy of Virtue

This anthology is part of a growing body of literature within the AI ethics/AI and social responsibility literature that attempts to produce ‘self-reflexive critiques of the conditions of knowledge creation.’[[48]](#footnote-48) This style of critique builds on a long tradition in which researchers approach their own practices, communities, and institutions as objects for critical analysis and reflexivity. Exemplary forms of this style are found in areas such as critical race studies, Indigenous studies, Black Feminist theory, and other strands of feminist scholarship.[[49]](#footnote-49) While the tenor of this work is often critical, the intention is to instigate positive change and to hold organizations and research communities accountable to the standards and values that they espouse. In recent years, many scholars have turned their attention to the growing interface between industry and the academy, and the impact on the topic of ethics—both as a domain of research (e.g. bioethics, AI ethics) and as a set of guiding principles that manage the creation of knowledge (e.g. diversity, equity, and inclusion programs).[[50]](#footnote-50)

The contribution offered in this edition is an effort to highlight the complexities of researcher positionality in this field, and expose the dilemmas we all face when making choices about the research, methods, and partners we pursue. Many of the contributors have direct experience in industry, research-focused arms of industry, or university faculties and research institutes that receive funding and other kinds of support from industry.[[51]](#footnote-51) Several of the contributions build from personal experience, provide insight into the institutional and organizational arrangements of Big Tech, and give testimony to the conditions of labour that shape research practice within these contexts. We have no interest in calling out colleagues for the choices they are compelled to make under capitalism. Our interest is understanding the ways industry both establishes and takes advantage of the incentive structures operating in this research environment. We hope this facilitates further reflexive analyses by university workers on the conditions of their own work in this domain.

Central to knowledge creation in scholarly contexts are mechanisms for the evaluation of research and researchers.[[52]](#footnote-52) This has long been a contentious space, and indeed an area in which Big Tech has become increasingly central.[[53]](#footnote-53) In the curation and production of this edition, we adopted a model of collective editorship where (lead) chapter authors were allocated other chapters to review. Our intention in doing so was to implement a constructive approach to ‘peer-review’ that was designed to discourage disciplinary gatekeeping, and accommodate the challenging topics and themes explored throughout the edited collection. Our aim was to encourage reviewers to be accountable for their feedback while constructively working closely together to improve the quality of the text and argument. We also hoped to participate in mechanisms that enhance the scholarly community’s control over the conditions of its own work.

We—the editors—experienced this as a solidarity-building exercise far removed from the anonymous peer review processes through which commercial publishers extract academic labor in ways typically unrecognized by university incentives structures, and generally experienced as unappreciated and unenjoyable by workers. The same desire to (re)imagine and (re)define the university work experience influenced our choice of publisher—the Institute of Network Cultures (INC)—a press which has a history and reputation of multidisciplinary knowledge production and engagement that responds to urgent matters relating to digital networks through open access publishing and advocacy.

But these are, of course, marginal actions in the context of a research ecology defined by the dynamics outlined in the following chapters. How to find a path forward and navigate economies of virtue is an epic challenge. This collection of texts is part of the process of naming the dynamics of tech capture, co-optation, and compromise. The hope is for scholars, advocates, activists, and policy-makers to incorporate these reflexive critiques of the conditions of knowledge creation and dissemination, and the compromises and trade-offs faced by knowledge workers over whom interested institutions have power. This is certain to be uncomfortable given the politics of collegial proximity that inform academic prestige networks. But naming these dynamics is the only way to address them and to stage questions that allow us to envision and demand alternative futures.

## Acknowledgements

## There have been so many wonderful people who have helped us to bring this collection to life. The brilliant authors for not only contributing their time but for generously diving into the experimental “collective review” process. To Geert Lovink for his patience and support of our project, and for his commitment and vision to modes of publication that subvert the constraining norms of academic practice. To the editors and reviewers at *Science as Culture*, in particular Kean Birch and Kelly Bronson who edited its ‘Forum on Big Tech’ where we initially published on the ‘economy of virtue.’ To Elena Gomez for her swift and incisive editing. To the presenters who joined us at our first Economies of Virtue workshop in July 2021: Angela Daly, Sarah Pink, Brett Neilson, Mark Andrejevic, Julia Powles, Sy Taffel, Winifred Poster, Tsvetelina Hristova, Michael Richardson, Tanja Dreher, and Ed Santow. To Emma Kowal for her support and guidance in helping us to apply for the Academy of the Social Sciences in Australia workshop program grant. And finally, to the many interlocutors who have generously held conversations with us and more as we explored this thorny topic, especially Seth Lazar, Kate Devitt, Christine Parker, and Julia Powles.

## Funding Disclosures

## This work was supported by the Australian Research Council’s Centre of Excellence for Automated Decision-Making & Society [Grant Number CE200100005]. In particular, we wish to acknowledge the Melbourne Law School node of the Centre for its financial support in bringing the book to fruition. Declan Kuch acknowledges support of the Vice Chancellor’s Fellowship scheme at Western Sydney University for assistance with printing. This volume builds on the ‘Economies of Virtue: The Circulation of ‘Ethics’ in AI and Digital Culture’ workshop hosted at Deakin University 8–9 July 2021, funded by the Academy of the Social Sciences in Australia, Workshop Program Grant 2021**.**

## References

Abdurahman, J. Khadijah. ‘On the Moral Collapse of AI Ethics’ *Medium* blog,

8 December 2020,https://upfromthecracks.medium.com/on-the-moral-collapse-of-ai-ethics-791cbc7df872.

Abdurahman, J. Khadijah. ‘Fired for Speaking Out Against Anti-Blackness and Ethnic Cleansing: Open Letter to Cornell Tech’, *Medium* blog, 3 June, 2021,https://upfromthecracks.medium.com/fired-for-speaking-out-against-anti-blackness-and-ethnic-cleansing-open-letter-to-cornell-tech-3b3f669cf69f.

———. ‘On the Moral Collapse of AI Ethics’, *Medium* blog, 7 December, 2020,https://upfromthecracks.medium.com/on-the-moral-collapse-of-ai-ethics-791cbc7df872.

Ahmed, Sara. *On Being Included: Racism and Diversity in Institutional Life*. Durham: Duke University Press, 2012.

———. ‘The Language of Diversity’, *Ethnic and Racial Studies* 30 (2, 2007): 235–56.

Australian Research Council. ‘Selection Report: Discovery Early Career Research Award 2022 | Australian Research Council.’ Australian Government | Australian Research Council, 2022.https://www.arc.gov.au/funding-research/funding-outcome/selection-outcome-reports/selection-report-discovery-early-career-research-award-2022.

Barnes, Joel. ‘Defunding Arts Degrees Is the Latest Battle in a 40-Year Culture War’, *The Conversation*, 3 July 2020,http://theconversation.com/defunding-arts-degrees-is-the-latest-battle-in-a-40-year-culture-war-141689.

Barocas, Solon, and Andrew D. Selbst. “Big Data’s Disparate Impact.” *California Law Review* 104, no. 3 (2016): 671–732.

Benjamin, Ruha. *Race After Technology: Abolitionist Tools for the New Jim Code*. 1 edition. Medford, MA: Polity, 2019,https://ebookcentral-proquest-com.ezproxy-b.deakin.edu.au/lib/deakin/detail.action?docID=5820427.

Bietti, Elettra. ‘From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy’, *SSRN Electronic Journal*, 2021.

Birhane, Abeba, Ruane, Elayne, Laurent, Thomas, Brown, Matthew S., Flowers, Johnathan, Ventresque, Anthony, and Dancy, Christopher L. ‘The Forgotten Margins of AI Ethics’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 948–58. FAccT ’22. New York, NY, USA: Association for Computing Machinery, 2022.

Bourdieu, Pierre. *Science of Science and Reflexivity*, trans. Richard Nice, Chicago: University of Chicago Press, 2004.

Burn-Murdoch, John. ‘The Problem with Algorithms: Magnifying Misbehaviour’, *Guardian*, 14 August 2013, https://www.theguardian.com/news/datablog/2013/aug/14/problem-with-algorithms-magnifying-misbehaviour.

Crawford, Kate. ‘The Hidden Biases in Big Data’, *Harvard Business Review*, 1 April 2013,https://hbr.org/2013/04/the-hidden-biases-in-big-data.

Diakopoulos, Nicholas. ‘Rage Against the Algorithms’, *The Atlantic*, 3 October 2013,https://www.theatlantic.com/technology/archive/2013/10/rage-against-the-algorithms/280255/.

Dormehl, Luke. ‘Algorithms Are Great and All, But They Can Also Ruin Lives.’ Accessed October 14, 2022.https://www.wired.com/2014/11/algorithms-great-can-also-ruin-lives/.

Dwork, Cynthia, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. ‘Fairness through Awareness.’ In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference on - ITCS ’12*, 214–26. Cambridge, Massachusetts: ACM Press, 2012.

Elliott, Carl. *White Coat, Black Hat: Adventures on the Dark Side of Medicine*, Boston: Beacon Press, 2010.

———. ‘Why Clinical Ethicists Are Not Activists’, *Hastings Center Report* 51 (4, 2021): 36–7.

Gansky, Ben, and Sean McDonald. ‘CounterFAccTual: How FAccT Undermines Its Organizing Principles’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1982–92. Seoul Republic of Korea: ACM, 2022.

Gebru, Timnit. ‘For Truly Ethical AI, Its Research Must Be Independent from Big Tech’, *Guardian*, 6 December 2021, https://www.theguardian.com/commentisfree/2021/dec/06/google-silicon-valley-ai-timnit-gebru.

Goldenfein, Jake, and Daniel Griffin. ‘Google Scholar – Platforming the Scholarly Economy’, *Internet Policy Review* 11, no. 3, 29 September 2022,https://policyreview.info/articles/analysis/google-scholar-platforming-scholarly-economy.

Hao, Karen. ‘A Leading AI Ethics Researcher Says She’s Been Fired from Google’, *MIT Technology Review*, 3 December 2020, https://www.technologyreview.com/2020/12/03/1013065/google-ai-ethics-lead-timnit-gebru-fired/.

———. ‘In 2020, Let’s Stop AI Ethics-Washing and Actually Do Something’, *MIT Technology Review*, 27 December 2019, https://www.technologyreview.com/2019/12/27/57/ai-ethics-washing-time-to-act/.

Hare, Julie. ‘Wage Theft Is “Systemic”: 21 Universities under Investigation’, *Australian Financial Review*, 20 October 2021, https://www.afr.com/work-and-careers/education/wage-theft-is-systemic-21-universities-under-investigation-20211020-p591kw.

Hoffmann, Anna Lauren. ‘Terms of Inclusion: Data, Discourse, Violence’, *New Media & Society*, 16 September 2020, 146144482095872.

hooks, bell. *Talking Back: Thinking Feminist, Thinking Black*, London: Routledge, 2014.

Hu, Lily. ‘Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?’ *Journal of Social Computing* 2, (3, 2021): 238–48.

Johnson, Bobbie, and Gideon Lichfield. ‘Hey Google, Sorry You Lost Your Ethics Council, So We Made One for You’, *MIT Technology Review* (blog), 8 April 2019,https://medium.com/mit-technology-review/hey-google-sorry-you-lost-your-ethics-council-so-we-made-one-for-you-28ee6c33576a.

Manjoo, F. ‘Silicon Valley Has an Arrogance Problem—WSJ’, *Wall Street Journal*, 3 November 2013, https://www.wsj.com/articles/SB10001424052702303661404579175712015473766.

McCarthy, Gregory Michael, and Jayasuriya, Kanishka. ‘A Review into How University Research Works in Australia Has Just Begun—It Must Confront These 3 Issues’, *The Conversation*, 20 September 2022,http://theconversation.com/a-review-into-how-university-research-works-in-australia-has-just-begun-it-must-confront-these-3-issues-190551.

Moreton-Robinson, Aileen. *Talkin’ Up to the White Woman: Aboriginal Women and Feminism*, Brisbane: University of Queensland Press, 2000.

Morozov, Evgeny. *To Save Everything, Click Here*, New York: PublicAffairs, 2014.

Moten, Fred, and Stefano Harney. ‘The University and the Undercommons’, *Social Text* 22 (2, 2004): 101–15.

Munn, Luke. ‘The Uselessness of AI Ethics’, *AI and Ethics*, 23 August 2022.

Noble, Safiya. *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York: NYU Press, 2018.

———. ‘Google Search: Hyper-Visibility as a Means of Rendering Black Women and Girls Invisible—InVisible Culture’, *Invisible Culture*, no. 19 (October 2013),https://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/.

———. ‘Missed Connections: What Search Engines Say About Women.’ *Bitch* magazine, Spring 2012,https://www.bitchmedia.org/issue/54.

Ochigame, Rodrigo. ‘The Invention of ‘Ethical AI’: How Big Tech Manipulates Academia to Avoid Regulation’, *The Intercept* (blog), 20 December 2019,https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/.

Phan, Thao, Jake Goldenfein, Monique Mann, and Declan Kuch. ‘Economies of Virtue: The Circulation of “Ethics” in Big Tech’, *Science as Culture*, 4 November 2021, 1–15.

Powles, Julia, and Helen Nissenbaum. ‘The Seductive Diversion of ‘Solving’ Bias in Artificial Intelligence’, *OneZero* (blog), 7 December 2018,

https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53.

Rosenblat, Alex, and Tamara Kneese. ‘Algorithmic Accountability’, n.d., 6.

Sweeney, Latanya. ‘Discrimination in Online Ad Delivery: Google Ads, Black Names and White Names, Racial Discrimination, and Click Advertising’, *Queue* 11 (3, 2013): 10–29.

Tarnoff, Ben. ‘Postcript to J. Khadijah Abdurahman’s “A Body of Work That Cannot Be Ignored”’, *Logic* magazine, 25 December 2021,https://logicmag.io/beacons/a-body-of-work-that-cannot-be-ignored/.

Turner, Fred. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism* (illustrated edition) Chicago: University of Chicago Press, 2008.

Upturn. ‘Civil Rights, Big Data and Our Algorithmic Future: A September 2014 Report on Social Justice and Technology’, 2014,https://bigdata.fairness.io/wp-content/uploads/2015/04/2015-04-20-Civil-Rights-Big-Data-and-Our-Algorithmic-Future-v1.2.pdf.

Vinsel, Lee. ‘You’re Doing It Wrong: Notes on Criticism and Technology Hype’, *Medium* blog, 1 February 2021,https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5.

Wagner, Ben. ‘Ethics As An Escape From Regulation. From “Ethics-Washing” to Ethics-Shopping?’ in *Ethics As An Escape From Regulation. From ‘Ethics-Washing’ To Ethics-Shopping?*, 84–9. Amsterdam: Amsterdam University Press, 2018.

Whittaker, Meredith. ‘The Steep Cost of Capture’, *Interactions* 28 (6, 2021): 50–5.

Wooldridge, Adrian. ‘The Coming Tech-Lash’, *The Economist*, 18 November 2013,https://www.economist.com/news/2013/11/18/the-coming-tech-lash.

Young, Meg, Michael Katell, and P.M. Krafft. ‘Confronting Power and Corporate Capture at the FAccT Conference’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1375–86. Seoul Republic of Korea: ACM, 2022.

1. Elettra Bietti, ‘From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy’, *SSRN Electronic Journal*, 2021: 211 & 217. [↑](#footnote-ref-1)
2. Bietti, ‘From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy’. [↑](#footnote-ref-2)
3. Thao Phan, Jake Goldenfein, Monique Mann, and Declan Kuch, ‘Economies of Virtue: The Circulation of “Ethics” in Big Tech’, *Science as Culture*, 4 November 2021, 1–15. [↑](#footnote-ref-3)
4. Meredith Whittaker, ‘The Steep Cost of Capture’, *Interactions* 28, no. 6 (2021): 50–5. [↑](#footnote-ref-4)
5. See, for example, Abeba Birhane, Elayne Ruane, Thomas Laurent, Matthew S. Brown, Johnathan Flowers, Anthony Ventresque, and Christopher L. Dancy, ‘The Forgotten Margins of AI Ethics’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 948–58. FAccT ’22. New York, NY, USA: Association for Computing Machinery, 2022; Ben Gansky, and Sean McDonald, ‘CounterFAccTual: How FAccT Undermines Its Organizing Principles’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1982–92. Seoul Republic of Korea: ACM, 2022; Young, Meg, Michael Katell, and P.M. Krafft. ‘Confronting Power and Corporate Capture at the FAccT Conference’ in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1375–86. Seoul Republic of Korea: ACM, 2022. For further discussion, see ‘Drop Outs’ in this volume. [↑](#footnote-ref-5)
6. Bietti, ‘From Ethics Washing to Ethics Bashing’, 211 & 217. [↑](#footnote-ref-6)
7. Ben Wagner, ‘Ethics as an Escape from Regulation. From “Ethics-Washing” to Ethics-Shopping?’ in *Being Profiled*, Amsterdam: Amsterdam University Press, 2018, 84–9. [↑](#footnote-ref-7)
8. Bietti, ‘From Ethics Washing to Ethics Bashing’. [↑](#footnote-ref-8)
9. Rashida Richardson in Bobbie Johnson and Gideon Lichfield, ‘Hey Google, Sorry You Lost Your Ethics Council, So We Made One for You’, *MIT Technology Review* blog, 8 April 2019.https://medium.com/mit-technology-review/hey-google-sorry-you-lost-your-ethics-council-so-we-made-one-for-you-28ee6c33576a. [↑](#footnote-ref-9)
10. Jake Metcalf in Johnson & Lichfield, ‘Hey Google, Sorry You Lost Your Ethics Council, So We Made One for You’. [↑](#footnote-ref-10)
11. Adam Greenfield in Johnson & Lichfield, ‘Hey Google, Sorry You Lost Your Ethics Council, So We Made One for You’. [↑](#footnote-ref-11)
12. Julia Powles and Helen Nissenbaum, ‘The Seductive Diversion of ‘Solving’ Bias in Artificial Intelligence’, *OneZero* blog, 7 December 2018.https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53. [↑](#footnote-ref-12)
13. Lily Hu, ‘Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?’ *Journal of Social Computing* 2 (3, 2021): 240. [↑](#footnote-ref-13)
14. Luke Munn, ‘The Uselessness of AI Ethics’, *AI and Ethics*, 23 August 2022: n.p. [↑](#footnote-ref-14)
15. Karen Hao, ‘In 2020, Let’s Stop AI Ethics-Washing and Actually Do Something,’ *MIT Technology Review*, https://www.technologyreview.com/2019/12/27/57/ai-ethics-washing-time-to-act/. [↑](#footnote-ref-15)
16. See Phan et al, ‘Economies of Virtue: The Circulation of “Ethics” in Big Tech’ and Whittakker, ‘The Steep Cost of Capture’. [↑](#footnote-ref-16)
17. The Australian Research Council’s DECRA (Discovery Early Career Researcher Award), the most prestigious grant available for Australian scholars in their early years, has an average 17 percent success rate. In 2022, there was a slight improvement, with 19.7 percent success rate (see Australian Research Council, ‘Selection Report: Discovery Early Career Research Award 2022 | Australian Research Council,’ Australian Government | Australian Research Council, 2022. https://www.arc.gov.au/funding-research/funding-outcome/selection-outcome-reports/selection-report-discovery-early-career-research-award-2022.). For further discussion on the Australian Research Council and funding see Richardson’s chapter in this volume. [↑](#footnote-ref-17)
18. Gregory Michael McCarthy and Kanishka Jayasuriya, ‘A Review into How University Research Works in Australia Has Just Begun—It Must Confront These 3 Issues’, *The Conversation*, 20 September 2022, https://theconversation.com/a-review-into-how-university-research-works-in-australia-has-just-begun-it-must-confront-these-3-issues-190551 [↑](#footnote-ref-18)
19. Julie Hare, ‘Wage Theft Is “Systemic”: 21 Universities under Investigation’, *Australian Financial Review*, 20 October, 2021,https://www.afr.com/work-and-careers/education/wage-theft-is-systemic-21-universities-under-investigation-20211020-p591kw. [↑](#footnote-ref-19)
20. Hu, ‘Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?’, 240. [↑](#footnote-ref-20)
21. Hu, ‘Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?’, 243. [↑](#footnote-ref-21)
22. Lee Vinsel, ‘You’re Doing It Wrong: Notes on Criticism and Technology Hype.’ *Medium* blog, 1 February, 2021,https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5. [↑](#footnote-ref-22)
23. https://socialsciences.org.au/publications/annual-report-2021/. [↑](#footnote-ref-23)
24. Joel Barnes. ‘Defunding Arts Degrees Is the Latest Battle in a 40-Year Culture War’, *The Conversation*, 3 July 2020,http://theconversation.com/defunding-arts-degrees-is-the-latest-battle-in-a-40-year-culture-war-141689. [↑](#footnote-ref-24)
25. Fred Moten and Stefano Harney, ‘The University and the Undercommons’, *Social Text* 22, (2, 2004): 101–15. [↑](#footnote-ref-25)
26. Ben Tarnoff, ‘Postcript to J. Khadijah Abdurahman’s “A Body of Work That Cannot Be Ignored”’, *Logic* magazine, 2021,https://logicmag.io/beacons/a-body-of-work-that-cannot-be-ignored/. [↑](#footnote-ref-26)
27. Moten and Harney, ‘The University and the Undercommons’, 106. [↑](#footnote-ref-27)
28. Alan Wooldridge, ‘The Coming tech-lash’, *The Economist*, 18 November 2013, https://www.economist.com/news/2013/11/18/the-coming-tech-lash. [↑](#footnote-ref-28)
29. A good example is Peter Thiel’s libertarian interest in sea-steading, see generally F. Manjoo, ‘Silicon Valley has an Arrogance Problem’, *Wall Street Journal*, 3 November 2013,

    https://www.wsj.com/articles/SB10001424052702303661404579175712015473766. [↑](#footnote-ref-29)
30. Google’s famous self-stated mission since it was founded in 1998. For further discussion on this slogan, see ‘Open Secrets’ in this volume. [↑](#footnote-ref-30)
31. F. Turner, *From Counterculture to Cyberculture*, University of Chicago Press: Chicago, 2006; Morozov, *To Save Everything, Click Here*, Public Affairs: New York City, 2013b. [↑](#footnote-ref-31)
32. J. Burn-Murdoch, ‘The problem with algorithms: magnifying misbehaviour’, *Guardian*, 14 August 2013,

    https://www.theguardian.com/news/datablog/2013/aug/14/problem-with-algorithms-magnifying-misbehaviour. [↑](#footnote-ref-32)
33. L. Dormehl, ‘Algorithms Are Great and All, But They Can Also Ruin Lives’, *Wired*,19 November 2014, https://www.wired.com/2014/11/algorithms-great-can-also-ruin-lives/. [↑](#footnote-ref-33)
34. C. Dwork, M. Hardt, T. Pitassi, O. Reingold, & R. Zemel, ‘Fairness Through Awareness’, 2011, https://arxiv.org/abs/1104.3913. [↑](#footnote-ref-34)
35. L. Sweeney, ‘Discrimination in Online Ad Delivery: Google Ads, Black Names and White Names, Racial Discrimination, and Click Advertising’, *Queue* 11 (3, 2013): 10–29. 8 [↑](#footnote-ref-35)
36. K. Crawford, ‘The Hidden Biases in Big Data’, *Harvard Business Review*, 1 April 2013, https://hbr.org/2013/04/the-hidden-biases-in-big-data. [↑](#footnote-ref-36)
37. N. Diakopoulos, ‘Race Against the Algorithms’, *The Atlantic*, 4 October 2013, https://www.theatlantic.com/technology/archive/2013/10/rageagainst-the-algorithms/280255/. [↑](#footnote-ref-37)
38. S. Noble, ‘Missed Connections: What Search Engines Say About Women’, *Bitch* magazine, Spring 2012 (54), https://www.bitchmedia.org/issue/54; S. Noble, ‘Google Search: Hyper-visibility as a Means of Rendering Black Women and Girls Invisible—InVisible Culture’, *Invisible Culture*,

    19. https://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/; S. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, NYU Press: New York City, 2013. [↑](#footnote-ref-38)
39. S. Barocas, and A. Selbst, ‘Big Data’s Disparate Impact’, *California Law Review*,Vol. 104, No. 3, (June 2016): 671 - 732 [↑](#footnote-ref-39)
40. https://civilrights.org/2014/02/27/civil-rights-principles-era-big-data/# [↑](#footnote-ref-40)
41. Upturn, *Civil Rights, Big Data and Our Algorithmic Future: A September 2014 Report on Social Justice and Technology*, 2014,

    https://bigdata.fairness.io/wp-content/uploads/2015/04/2015-04-20-Civil-Rights-Big-Data-and-Our-Algorithmic-Future-v1.2.pdf. [↑](#footnote-ref-41)
42. See https://governingalgorithms.org/. [↑](#footnote-ref-42)
43. A. Rosenblat, T. Kneese, and d. boyd, ‘Algorithmic Accountability: A Workshop Primer Produced For: The Social, Cultural & Ethical Dimensions of “Big Data”’, 17 March 2014, New York, NY, https://www.datasociety.net/pubs/2014-0317/AlgorithmicAccountabilityPrimer.pdf. [↑](#footnote-ref-43)
44. See, for example, https://www.fatml.org/. [↑](#footnote-ref-44)
45. For further discussion, see Phan et al, ‘‘Economies of Virtue: The Circulation of “Ethics” in Big Tech’. [↑](#footnote-ref-45)
46. Lily Hu, ‘Tech Ethics: Speaking Ethics to Power, or Power Speaking Ethics?’ *Journal of Social Computing* 2 (3, 2021): 243. [↑](#footnote-ref-46)
47. J. Khadijah Abdurahman, ‘On the Moral Collapse of AI Ethics’ *Medium* blog,

    8 December 2020,https://upfromthecracks.medium.com/on-the-moral-collapse-of-ai-ethics-791cbc7df872.;

    Karen Hao, ‘A Leading AI Ethics Researcher Says She’s Been Fired from Google’, *MIT Technology Review*, 3 December 2020,https://www.technologyreview.com/2020/12/03/1013065/google-ai-ethics-lead-timnit-gebru-fired/; [↑](#footnote-ref-47)
48. Whittaker, ‘The Steep Cost of Capture’. [↑](#footnote-ref-48)
49. see, for example, Sara Ahmed, *On Being Included: Racism and Diversity in Institutional Life*. Duke University Press, 2012; bell hooks, *Talking Back: Thinking Feminist, Thinking Black*. 2 edition. Routledge, 2014; Aileen Moreton-Robinson, *Talkin’ Up to the White Woman: Aboriginal Women and Feminism*. Univ. of Queensland Press, 2000. [↑](#footnote-ref-49)
50. see Sara Ahmed, “The Language of Diversity.” *Ethnic and Racial Studies* 30, no. 2 (March 2007): 235–56; Ruha Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code*. 1 edition. Cambridge, UK; Medford, MA: Polity, 2019; Carl Elliott, *White Coat, Black Hat: Adventures on the Dark Side of Medicine*. 1st edition. Beacon Press, 2010; Anna Lauren Hoffmann, “Terms of Inclusion: Data, Discourse, Violence.” *New Media & Society*, September 16, 2020, 146144482095872 [↑](#footnote-ref-50)
51. see Powles and Nissenbaum, ‘The Seductive Diversion of ‘Solving’ Bias in Artificial Intelligence’; Ochigame, ‘The Invention of “Ethical AI”’; Gebru, ‘For Truly Ethical AI, Its Research Must Be Independent from Big Tech’; Abdurahman, ‘Fired for Speaking Out Against Anti-Blackness and Ethnic Cleansing: Open Letter to Cornell Tech’ and ‘On the Moral Collapse of AI Ethics’; Whittaker, ‘The Steep Cost of Capture’; Young, et al, ‘Confronting Power and Corporate Capture at the FAccT Conference’. [↑](#footnote-ref-51)
52. Pierre Bourdieu, *Science of Science and Reflexivity,* University of Chicago Press: Chicago, 2004. [↑](#footnote-ref-52)
53. Jake Goldenfein and Daniel Griffin, ‘Google Scholar: Platforming the Scholarly Economy’, *Internet Policy Review* 11(3) (online). [↑](#footnote-ref-53)