# Extractivist Ethics

### Sarah Pink

## Introduction

Let me take you to the moment I started to write this chapter about AI, ethics, and people, following links through reports and papers on ethics I found the AI4People’s ethical framework. The intentions of such projects have societal wellbeing at heart, and contribute actively to the important argument that AI needs to be regulated and needs to do good. This made me feel uncomfortable as I began to read; knowing I was about to turn my anthropologist’s eye to critique the logics of an agenda that seeks to make AI ethical, when surely I should be on the same page (*Spoiler alert: hold onto the page metaphor and don’t click on the links until I ask you to*). To be clear, my critique is not of the AI4People’s ethical framework in particular. Rather, the framework exemplifies how the metaphors, narratives, and structures that commonly frame good intentions towards ethical AI and People in dominant discourses betray a logic that is misaligned with everyday realities. My agenda, and my work as a whole, focuses on creating collaborative partnerships to work toward ethical futures, rather than simply writing endpoint outlines of what is wrong. But to make the connections/relations required for collaboration we need to make visible the cracks between approaches, disciplines, and logics.

The ‘AI4People’s Ethical Framework for a Good AI Society: Opportunities, Risks, Principles and Recommendations’report is ‘committed to the development of AI technology in a way that secures people’s trust, serves the public interest, and strengthens shared social responsibility’ and it presents a set of recommendations towards ensuring this.[[1]](#footnote-1) It is authored by a range of scholars and industry contributors who specialize in ethics from fields including philosophy, law, and computer science, but excluding the social sciences such as anthropology and qualitative sociology. While the report presents perfectly reasonable principles from the perspective of the societal structures in which we presently operate, the terminologies and concepts it uses to communicate its ideas are difficult to reconcile with those of an anthropological approach to people, ethics, and emerging technologies. First, the idea that people’s trust can be ‘secured’ by particular developments in AI technology requires trust to be a fixed quality that can be extracted from people and captured. Second, the notion of ‘the public interest’ invokes a one-dimensional framing of people, rather than actual people in the messy contingency of their lives where real interests and everyday ethics play out. Third, by referring to ‘social responsibility,’ it focuses on a sociological unit or level of analysis, rather than on the experiential domain of life. That is, the AI4People agenda is inhabited by a striking absence of the experience and activity of actual people.

In this chapter I argue for a people-focused approach which must be surfaced through engagement with theory and research in everyday worlds. Elsewhere I have defined ‘techno-solutionist approaches to ethics as extractivist, where they seek to identify and capture human ethics values and invest them in machines with the intention that such ethical machines will engender trust.’[[2]](#footnote-2) Here, I extend this argument, along with the premise that to be ethical, AI should not simply be for people, but be designed with people, attentive to diversity, specificity and locality. It should not seek to secure, gain or win people’s trust or anything else after the event of its design, but should be created already within relations of trust, attentive to ‘everyday ethics’.[[3]](#footnote-3) By everyday ethics I mean ethics as they are lived out in and contingent on the circumstances of everyday life; where ethics are not necessarily fixed in such a way that they can be applied consistently across all situations, and are nuanced by the relationships between people, things and environment. My definition builds on ethics as understood in phenomenological anthropology, which ‘reveals ethical life as a condition marked by ontological indeterminacy and ethical overload;’[[4]](#footnote-4) in anthropology (as I practice it) ethics are indeterminate, ‘contingent, emergent from the everyday worlds and circumstances of life.’[[5]](#footnote-5) Such an understanding of ethics permeates anthropological research ethics as well as how we understand other people’s ethics, and thus concerns equally the actions of the reflexive ethnographer.[[6]](#footnote-6)

The AI4People’s agenda is of course not alone in its approach, and it seeks to offer a solution to the question of how to make AI ethical, which aligns with the ways AI is being developed. In one sense, this is a step in an ethical direction. But it is also emblematic of a consistent and glaring gap in the dominant discourses advanced in the technology industry, government, and the engineering and computer sciences about how to make AI ethical. I believe that there is a common concern about people and AI across these disciplines and stakeholders, which we can better address by bringing people (encountered through collaborative ethnographic research practices, engagements, and interventions) into the debate. This means we need to ensure that everyday ways of knowing and diversity are at the forefront of the ways we consider ethics. But a glance at the AI4People report’s web page[[7]](#footnote-7) banner (*go to the link now*) assures me that the page is visibly gendered. Although the report itself was authored by men and women, this appears to be erased in the visual representation of the banner. The substantially outed[[8]](#footnote-8) and critiqued ‘manel’ (all male panel)[[9]](#footnote-9) lives on, and highlights that the topic of extractivist ethics I am about to address also surfaces the gendered politics of technology, data, ethics, and academia.

## Extractivist ethics (mining for ethics)

The metaphors we use to refer to data, as well as automated, connected, and intelligent emerging technologies and systems are themselves sites of contestation, and with that they also constitute possible sites of investigation and intervention. Sally Wyatt suggests that as critical social scientists we need to contest the extractive metaphors used by industry and policy makers which frame data as a resource that can be mined.[[10]](#footnote-10) In this chapter I outline a related mode of contestation, which instead of switching the metaphors involves applying them to ethics in order to interrogate how ethics are situated within approaches to data and emerging technologies such as AI, which scholars have already labelled as being extractivist.[[11]](#footnote-11) I discuss how similar logics of extraction, which have been critiqued by critical data scholars, are applied to ethics in AI, and what this suggests regarding industry, policy, and research. I call this *extractivist ethics.*

So how are extractivist ethics constituted? On one level, as exemplified above, ethics become the bait through which trust in technology is extracted from publics or users. On another, as I have proposed elsewhere, in relation to the relationship between trust and ethics, techno-solutionist approaches to ethics can be defined as extractivist where ‘they seek to identify and capture human ethics values and invest them in machines with the intention that such ethical machines will engender trust.’[[12]](#footnote-12) This kind of causality typifies renderings of ethics in the engineering sciences. A good example is the well-known MIT moral machine experiment,[[13]](#footnote-13) which I (and others) have discussed elsewhere,[[14]](#footnote-14) but here take up in a new direction.

The moral machine is a game which serves as survey seeking to extract the moral judgments of thousands of people across the world in relation to a set of future self-driving car scenarios based on the ‘Trolley Problem’ (a philosophical conundrum where the person playing the game needs to decide whom from a choice of possible victims the train car should kill in an accident) by judging a series of scenarios presented online. There is perhaps nothing surprising that a game should be used to extract ethics from its players. Writing more generally of social media, Sheila Jasanoff notes how such technologies ‘profit from people’ by ‘mining their thoughts, words, habits, bodies and emotions as resources to create new marketable goods.’[[15]](#footnote-15) Edmond Awad and colleagues (all men), have good intentions, as the AI4People authors discussed above. They are attentive to cultural difference and suggest that ‘we can embrace the challenges of machine ethics as a unique opportunity to decide, as a community, what we believe to be right or wrong; and to make sure that machines, unlike humans, unerringly follow these moral preferences.’[[16]](#footnote-16) Again, my quarrel is not so much with the sentiment but with an understanding of ethics which is disconnected from all the anthropological evidence of how ethics actually play out in the contingent circumstances of the everyday. As the philosopher Onora O’Neill has highlighted, the use of surveys or polls to quantify human sentiment, affective states, and contingent decisions is limited. For instance, polls on public trust ‘offer no evidence about the judgements that people make when they decide to trust or refuse trust to particular individuals or institutions for particular matters, in which they often differentiate cases with some care.’[[17]](#footnote-17) Equally questionable is the status of knowledge about ethics derived from responses to improbable ethical dilemmas which are subsequently suspended from their sources, rather than situated within realistic situations in which they actually unfold.

The implication of understanding ethics as contingent (as argued earlier) is that everyday ethics are slippery, they are incredibly difficult to capture, to invest in either organizations or machines, or to regulate. Thus, it follows that the assumptions it is possible to solve ethical problems that emerge after AI has become embedded in everyday life are limited. Typical solutions involving either designing ethical machines, or introducing regulation and governance, construct risk mitigation processes, based on logics which follow causal chains created externally to everyday life and its ethics. When instead we turn the focus to what it actually means to be human in the everyday—that is, the experience of being, feeling, and doing, ethics cannot be abstracted, fixed, or predetermined externally to the everyday. Instead, through giving ‘primacy to first- and second- person positions’, phenomenology draws our attention to the intersubjectivity and intercorporeality of ethics.[[18]](#footnote-18) As Mattingly and Throop put it, ‘Far from being a site of culturally well-articulated obligations or the imposition of normative moral orders that create docile subjects, scholars have empirically documented ways that the ethical can pose excessive demands that render lived experience uncanny.’[[19]](#footnote-19) This means that while there are ongoing attempts to abstract ethics into regulations, such ethics are unlikely to ever be aligned with the ethical requirements of everyday life.

As such, in this section I have offered two very different answers to questions of the kind invoked by legal and STS scholar Sheila Jasinoff when she asks: ‘Whose duty is it in today’s complex societies to foresee or forestall the negative impacts of technology, and do we possess the necessary tools and instruments for forecasting and preventing harm?’[[20]](#footnote-20) The MIT moral machine experiment tries to answer this question head on by both taking responsibility for forestalling the negative and creating tools through which to create ethical self-driving cars that people will subsequently trust and adopt, and in doing so reduce traffic deaths and carbon emissions. But Mattingly and Throop ask a different question, which complicates both this response and the mode of responsibility it assumes: ‘What is at stake in emphasizing the underdetermined nature of ethical life […]? What does it mean to portray the human as characterized by potentiality or possibility rather than actuality? What does it mean to claim that there is an excessiveness to the ethical demand such that it cannot be reduced to following prescriptive norms or rules?’[[21]](#footnote-21) In this context the answer is that because the ethics that will characterize the relations between people and self-driving cars (and by extension AI in general) are indeterminate, they can neither be extrapolated to machines nor be engaged to forecast and prevent harm.

## Extractivist Ethics and Anticipatory Audits (Attracting Investment)

In this section I investigate how extractivist ethics could participate in the anticipatory visions of capitalism. These visions are key to capitalism occupying the future to maintain its structural hold on everyday life. As a resource that can be extracted, or as a bait to capture trust, ethics can be invested in trustworthy intelligent and automated machines, thus serving as the catalyst in causal chains of human trust, acceptance, and adoption of AI. Here, extracted ethics could participate in creating an anticipatory infrastructure through which ethical AI and ethical machines are seen as a technological solution to situations where public acceptance of automated technologies is perceived as a challenge,[[22]](#footnote-22) and thus to attract investment in the technologies that are envisioned as solutions to societal problems.

To make investment and markets for AI plausible and realistic, extractivist ethics are also aligned with what I call the *anticipatory audit*. Anticipatory audits are part of what anthropologists have long since referred to as ‘audit cultures.’[[23]](#footnote-23) Many elements of audit culture are anticipatory by nature. Take, for example, university ethics committees.[[24]](#footnote-24) This is an example I have discussed often but it is worth repeating here because it both connects with the academic research and funding context mentioned below, and is likely part of the experience of academic readers. Usually regulated by an institutional (or in some case national) body which sets the rules which define ethical research conduct, ethical approval involves ensuring that any risks of what is defined as unethical happening in our research are identified and mitigated *in advance.* However, when ethics are the subject of an anticipatory audit, the only way that ethics can be accounted for is by fixing them still, capturing them for measurement against ethical regulations. The result is to reassure our institutions that our research will in fact be ethical (and that they have minimized the possibility of conduct that would be interpreted as unethical). The case of ethical AI is similar in that the ethical conditions that AI should manifest are prescribed in advance through ethics frameworks, and can therefore, like the ethics of researchers, the ethics of AI can also be audited before they are let loose into the world—that is, into everyday life environments.

As one of many examples, the website of the top consultancy firm PricewaterhouseCoopers (PwC) takes up the question of ‘Responsible AI (RAI)’, which it states ‘is the only way to mitigate AI risks.’[[25]](#footnote-25) Their ‘Responsible AI Toolkit’ includes a focus on ethics. Yet while the possibility that getting the ethics right will mitigate the risks is tempting to believe, what that actually means still appears to be in the balance; a 2019 review of international ethics frameworks found ‘a global convergence emerging around five ethical principles [for AI] (transparency, justice and fairness, non-maleficence, responsibility and privacy), with substantive divergence in relation to how these principles are interpreted, why they are deemed important, what issue, domain or actors they pertain to, and how they should be implemented.’[[26]](#footnote-26) Moreover, in 2021 the Pew Research Institute issued the findings of their survey of technology experts, to suggest that most did not believe that ethical AI design would be broadly adopted by 2030.[[27]](#footnote-27)

However, of the most significant comments cited by the Pew were those by danah boyd, who pointed out that when AI systems are aligned with contemporary capitalism, ‘which fetishizes efficiency, scale and automation,’ they are antithetical to the ethical values of ‘augmentation, localized context and inclusion.’ boyd’s insight connects with the everyday ethics outlined above, which emphasizes precisely how ethics are contingent and specific. As these points show, it is not just a question of what the ethics of AI are, but also a question of where it gets its ethics from and whose values they align with. As I have shown through the example of the moral machine experiment, there have been attempts to extract the ethics from the everyday by aggregating individual responses, but these inevitably fail to generate ethics that align with ethics in the everyday because they are extractive. We cannot mine ethics. Rather we have to get in there with them. It is the opposite of extraction; it requires blending and collaboration both in place and with the ongoing emergence of life.

Indeed, the similarities between the anticipatory ethics audits we experience as university academics and those AI systems are subject to don’t stop at their common impulse to mitigate the risks related to what certain agents will do ‘in the wild.’ Both modes of anticipatory audit are also meant to account for, regulate, control, and mitigate any risks involved in the ethical behaviour of an active agent in the form of the AI or the researcher, over a passive agent in the form of a member of the public, a user or consumer, or a research participant. There are, however, several mismatches between anticipatory audits of ethics and the everyday ethics in which the academic researchers or the AI systems and technologies (who or which have been audited) will be let loose. The everyday worlds where their anticipated one-way ethical effects will be activated are in fact inhabited by very different ethics—the experiential ethics noted by Mattingly and Throop, which are contingent, contextual, embodied, intersubjective, and indeterminate.[[28]](#footnote-28)

## Funding Extractivist Ethics (The Gender of Funding)

Above I have outlined the inherent flaw in visions of human ethics as a determinate thing which can be extracted from society or garnered from experts as representing societal values, captured, and transferred into a machine.[[29]](#footnote-29) Yet such approaches to ethics offer a (deceptively) simple response to a complex problem, with a causal chain of guarantees which mitigate the risks of AI doing future harm, *as well as* mitigating a set of risks around the research needed to create the knowledge and technologies that will apply the solution.

Research that proposes to embed predetermined ethical values into AI is relatively not risky because it shows a clear route to impact. It might, of course, entail other risks relating to the difficulties or uncertainties related to the technological discoveries that the researchers wish to make, which is a different thing. However, if you already believe that ethical AI—AI infused with societally endorsed ethical values—will make people trust, accept, and adopt technology that will benefit society and the environment, it’s not a big leap to consequently assume that it’s a good idea to fund research that will aim to produce AI that will only act according to desirable human ethics, and that will be governed by an ethics framework approved by experts. Thus an extractivist ethics agenda would ultimately be appealing to well-intentioned organizations and researchers involved in narratives and practices of dominant innovation agendas. It would subsequently support the academic careers of those whose work is funded through them, and oil the wheels of the machines of research funding, outputs, and impact that govern success in academia. One of the factors that appears to govern success in academia, at least in funding outputs, is gender. In 2019 an EU H2020 funded project titled ‘Grant Allocation Disparities from a Gender Perspective’ reported a set of ‘indisputable facts:’ there are fewer women than men in STEM disciplines and in senior academic positions, and women get fewer research grants, less funding, and lower evaluations.[[30]](#footnote-30)

## Extractivist Ethics and Everyday Ethics at the Impasse

In this chapter, I have proposed the concept of *extractivist ethics*, which I suggest creates a category through which to reveal and contest the dominant narratives concerning the generation of trust and acceptance of and the constitution of markets for emerging technologies in society. I have explored the alignment of *extractivist ethics* to another concept—the *anticipatory audit.* I have suggested the risk mitigation paradigm that structures both extractivist ethics and the anticipatory audit, also aligns them to both corporate and research agendas, because they both promise paths to impact.

Approaches to ethical AI that call for ethics frameworks and regulation have good intentions. Luciano Floridi is right to advocate that ‘Ethics-first is the right approach to set global standards for AI.’[[31]](#footnote-31) However, for ethics to really come first, more work is needed. At the moment the logics of ethics from above through regulation are not compatible with ethics as they occur in the everyday contexts that they ultimately seek to (ethically) impact on. Everyday ethics cannot entertain the certainties that extractivist ethics, as articulated in relation to ethical machines, desire. In part this concurs with another point boyd makes in the Pew Survey, that ‘[w]e misunderstand ethics when we think of it as a binary, when we think that things can be ethical or unethical;’ such binaries indeed coincide with the idea that machines can be made ethical. Seeing ethics participating in predictable causal sequences is similarly incorrect. Ethics cannot participate in predictable chains of reactions, simply because they are not static. These anthropological interpretations of ethics complicate the STEM models of ethical machines and their promise of beneficial impact on society. They are moreover difficult to work with, and teach, because they are slippery, tricky, and don’t stay still. But besides this they proffer another challenge to the societal structures that make STEM valued in research because as boyd puts it: ‘We cannot meaningfully talk about ethical AI until we can call into question the logics of late-stage capitalism.’

There is a politics to everyday ethics, which, as I hinted at the beginning of this chapter, which also requires us to attend to questions of gender. Gender is intersectional, not binary, meaning that both my own references to all male panels, research teams and the possibility of bias in research funding outcomes towards men, all need to be nuanced with other modes of difference, inequality, and inequity. However, the evidence suggests that AI is emerging within a gendered enterprise of research and development, which frequently favours men, and to move forward we need to empower other voices. A starting point could be to borrow Point 6 of the Feminist Data Manifest-no,[[32]](#footnote-32) to ‘refuse the expansion of forms of [data science] *ethics frameworks* that normalize[s] a condition of [data] *ethics* extractivism and is defined primarily by the drive to monetize and hyper-individualize the human experience.’ When ethics (as facets of human experience) are extracted from the everyday, or are used as bait to capture other everyday feelings like trust, in order to constitute anticipated markets, they are effectively being commodified. Like feminist data scholars we should instead: ‘commit to centering creative and collective forms of life, living, and worldmaking that exceed the neoliberal logics and resist the market-driven forces to commodify human experience’.

## Acknowledgments

This chapter has benefited from its original conceptualisation as part of the excellent Economies of Virtue workshop. I thank Thao Phan, Monique Mann, Jake Goldenfein, and Declan Kuch for their work and inspiration and I am especially grateful to Ellen Broad, Lorenn Ruster, Jake Goldenfein, and Declan Kuch for their wonderfully inspiring comments and review of the first version of this chapter.

## Funding Disclosure

My work on this chapter is undertaken in my role as a Chief Investigator in the Australian Research Council–funded Centre of Excellence for Automated Decision-Making and Society (CE200100005) (2020–2027).

## References

Awad, E., Dsouza, S., Kim, R. et al*.* ‘The Moral Machine Experiment‘, *Nature* 563 (2018): 59–64.

Cruz Castro, L. and Sans Menéndez, L. *Literature Review Synthesis Report.* CSIS Institute of Public Goods and Policies, Madrid, 2019.

Floridi, L. ‘Establishing the Rules for Building Trustworthy AI’, *Nat Mach Intell* 1 (2019): 261–262.

Floridi, L., Cowls, J., Beltrametti, M. et al. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds & Machines* 28,(2018): 689–707.

Jasanoff, S. *The Ethics of Invention*, New York: W. W. Norton & Company, 2016.

Jobin, A., Ienca, M. & Vayena, E. ‘The Global Landscape of AI Ethics Guidelines’, *Nat Mach Intell* 1 (2019): 389–99.

Mattingly, C. and J. Throop. ‘The Anthropology of Ethics and Morality’, *Annual Review of Anthropology* 47:1 (2018): 475–92.

O’Nell, O. ‘Accountable Institutions, Trustworthy Cultures’, *Hague J Rule Law* 9 (2017): 401–12.

Pels, P. ‘The Trickster’s Dilemma: Ethics and the Technologies of the Anthropological Self’ in M. Strathern (ed) *Audit Cultures: Anthropological Studies in Accountability*, London: Routledge, 2000, pp. 147–84.

Pink, S. ‘Ethics in a Changing World: Embracing Uncertainty, Understanding Futures, and Making Responsible Interventions’ in xS. Pink, V. Fors, T. O’Dell (eds) *Working in the Between: Theoretical Scholarship and Applied Practice.* Oxford: Berghahn, 2017, pp. 29–51.

——. ‘Trust, Ethics and Automation: Anticipatory Imaginaries in Everyday Life’ in Pink, S., Lupton, D., Berg, M. and Ruckenstein, M. (eds) *Everyday Automation*, London: Routledge, 2022, 44–58.

Pink, S., Raats, K., Lindgren T., Osz, K. and Fors, V. ‘An Interventional Design Anthropology of Emerging Technologies’ inHojer Bruun, M., Wahlberg, A., Brogaard Kristensen, D., Douglas-Jones, R.,  Hasse, C., Høyer, K., and Ross Winthereik, B. (eds) *The Handbook for the Anthropology of Technology*, London: Palgrave, 2021, pp. 183–200.

Stilgoe, J., T. Cohen, ‘Rejecting Acceptance: Learning from Public Dialogue on Self-Driving Vehicles’, *Science and Public Policy*, 48.6 (2021): 849–59.

Strathern, M. (ed). *Audit Cultures: Anthropological Studies in Accountability*, London: Routledge, 2000.

Wyatt, S. ‘Metaphors in Critical Internet and Digital Media Studies’, *New Media & Society*, *23*.2 (2021): 406–16.

1. L. Floridi, J. Cowls, M. Beltrametti, et al. ‘AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations’, *Minds & Machines* 28, (2018): 22, https://www.eismd.eu/wp-content/uploads/2019/11/AI4People%E2%80%99s-Ethical-Framework-for-a-Good-AI-Society\_compressed.pdf. [↑](#footnote-ref-1)
2. S. Pink, ‘Trust, Ethics and Automation: Anticipatory Imaginaries in Everyday Life’ in S. Pink, D. Lupton, M. Berg & M. Ruckenstein (eds) *Everyday Automation*, London: Routledge, 2022. [↑](#footnote-ref-2)
3. Pink, ‘Trust, Ethics and Automation’. [↑](#footnote-ref-3)
4. C. Mattingly and J. Throop, ‘The Anthropology of Ethics and Morality’, *Annual Review of Anthropology* 47.1 (2018): 483. [↑](#footnote-ref-4)
5. Pink, ‘Trust, Ethics and Automation’. [↑](#footnote-ref-5)
6. M. Strathern (ed) *Audit Cultures: Anthropological Studies in Accountability*, London: Routledge, 2000; P. Pels, ‘The Trickster’s Dilemma: Ethics and the Technologies of the Anthropological Self’ in M. Strathern (ed) *Audit Cultures*. [↑](#footnote-ref-6)
7. https://www.eismd.eu/featured/ai4peoples-ethical-framework-for-a-good-ai-society/. [↑](#footnote-ref-7)
8. @allmalepanels, https://twitter.com/allmalepanels?lang=en. [↑](#footnote-ref-8)
9. J.K. Rodriguez and E.A. Guenther, ‘What’s Wrong With “Manels” and what Can We Do About Them?’ The Conversation, 15 October 2020, https://theconversation.com/whats-wrong-with-manels-and-what-can-we-do-about-them-148068. [↑](#footnote-ref-9)
10. S. Wyatt, ‘Metaphors in Critical Internet and Digital Media Studies’, *New Media & Society*, 23.2 (2021): 406–16. [↑](#footnote-ref-10)
11. S. Jasanoff, *The Ethics of Invention*. New York: W. W. Norton & Company, 2016; Wyatt, ‘Metaphors’. [↑](#footnote-ref-11)
12. Pink, ‘Trust, Ethics and Automation’. [↑](#footnote-ref-12)
13. E. Awad, S. Dsouza, R. Kim et al*.*, ‘The Moral Machine Experiment’, *Nature*, 563 (2018): 59–64. See also: https://www.moralmachine.net/. [↑](#footnote-ref-13)
14. S. Pink, K. Raats, T. Lindgren, K. Osz, and V. Fors, ‘An Interventional Design Anthropology of Emerging Technologies’ inMaja Hojer Bruun, Ayo Wahlberg,  Dorthe Brogaard Kristensen, Rachel Douglas-Jones,  Cathrine Hasse,  Klaus Høyer, and Brit Ross Winthereik (eds) *The Handbook for the Anthropology of Technology*, London: Palgrave, 2021; Pink, ‘Trust, Ethics and Automation’. [↑](#footnote-ref-14)
15. Jasanoff, 259. [↑](#footnote-ref-15)
16. Awad et al., 63. [↑](#footnote-ref-16)
17. O. O’Neill, ‘Accountable Institutions, Trustworthy Cultures’, *Hague J Rule Law* 9 (2017): 405. [↑](#footnote-ref-17)
18. Mattingly and Throop, 483. [↑](#footnote-ref-18)
19. Mattingly and Throop, 485-6. [↑](#footnote-ref-19)
20. Jasanoff, 7. [↑](#footnote-ref-20)
21. Mattingly and Throop, ‘The Anthropology of Ethics and Morality’, 486. [↑](#footnote-ref-21)
22. J. Stilgoe, T. Cohen, ‘Rejecting Acceptance: Learning from Public Dialogue on Self-Driving Vehicles’, *Science and Public Policy*, 48.6 (2021): 849–59. [↑](#footnote-ref-22)
23. For example, Strathern, *Audit Cultures*. [↑](#footnote-ref-23)
24. S. Pink, ‘Ethics in a Changing World: Embracing Uncertainty, Understanding Futures, and Making Responsible Interventions’ in (eds) S. Pink, V. Fors, T. O’Dell *Working in the Between: Theoretical Scholarship and Applied Practice*,Oxford: Berghahn, 2017. To be clear, I support ethical review processes because when they are done well, they provoke reflection as well as ethical conduct. [↑](#footnote-ref-24)
25. PwC, https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai.html. [↑](#footnote-ref-25)
26. A. Jobin, M. Ienca, and E. Vayena, ‘The Global Landscape of AI Ethics Guidelines’, *Nat Mach Intell* 1 (2019): 389. [↑](#footnote-ref-26)
27. https://www.pewresearch.org/internet/2021/06/16/experts-doubt-ethical-ai-design-will-be-broadly-adopted-as-the-norm-within-the-next-decade/. [↑](#footnote-ref-27)
28. Mattingly and Throop, ‘The Anthropology of Ethics and Morality’. [↑](#footnote-ref-28)
29. In this chapter I raise this issue specifically in relation to the question of the challenge of making ethical machines. This issue raises wider questions relating to how such a stance might be reconciled with the status of ethics, regulation, and governance in society, which is not within the scope of this chapter to address. [↑](#footnote-ref-29)
30. L. Cruz Castro and L. Sans Menéndez, *Literature Review Synthesis Report.* CSIS Institute of Public Goods and Policies, Madrid, 2019. [↑](#footnote-ref-30)
31. L. Floridi, ‘Establishing the Rules for Building Trustworthy AI’, *Nat Mach Intell* 1 (2019): 262. [↑](#footnote-ref-31)
32. ​​Feminist Data Manifest-No, https://www.manifestno.com/home; Point 6 of the Feminist Data Manifesto reads: ‘We refuse the expansion of forms of data science that normalizes a condition of data extractivism and is defined primarily by the drive to monetize and hyper-individualize the human experience. We commit to centering creative and collective forms of life, living, and worldmaking that exceed the neoliberal logics and resist the market-driven forces to commodify human experience.’ [↑](#footnote-ref-32)