Against the Conflation of Corporate Strategy, Ethics, and the Politics of AI

Henrik Skaug Sætra[1]

Abstract

Ethics as we know it is ill equipped to resist abuse by technology companies, Van Maanen argues in a recent article. Ethics is too malleable, provides too many different theories, and allows for a plethora of ethical guidelines to be created, and ethicists purportedly end up enabling ethics shopping and ethics washing. But is ethics really the problem, and is the best solution to "repoliticise" ethics? In this commentary, I instead claim that we must distinguish between politics and ethics, as ethics has a separate and crucial role to play if we are to face the challenges related to controlling technology. This is important if we are to avoid undermining political involvement and democracy. Whilst ethicists certainly have the power to change the world, as Van Maanen argues, they are not—and should not be—provided with political authority.

1 Introduction

Ethics as we know it is ill equipped to resist abuse by technology companies, Van Maanen (2022) argues. Such abuse, we are told, tends to be rooted in the fact that ethics is too malleable, provides too many different theories, and allows for a plethora of ethical guidelines to be created. Consequently, ethicists end up enabling tech companies to cherry-pick some framework that is easily implemented without the need for cumbersome concessions and changes.

But is ethics really the problem, and is the solution, as Van Maanen argues, to "repoliticise" ethics and introduce a "question, rather than theory or principle-based ethical data practice"? This proposed alternative is based on "casuistry" and realistic political theory, with a particular focus on Raymond Geuss. Being a political

---

✉ Henrik Skaug Sætra
   Henrik.satra@hiof.no

1   Faculty of Computer Science, Engineering and Economics, Østfold University College,
     1757-N Remmen, Halden, Norway

⌂ Springer

theorist, one might be inclined to think I would endorse such a proposal. However, whilst I appreciate engagement with the discipline of political theory, I instead support the realist notion that politics is a distinct "sphere of life" (Galston, 2010), and that ethics has a somewhat separate and crucial role to play if we are to face the challenges related to controlling technology (Sætra & Fosch-Villaronga, 2021).

Let me begin, however, by stating that I fully agree with some of Van Maanen's conclusions, such as the need for ethics that is not "list-producing, and regulation-driven". Furthermore, I agree that ethics need not be an individualistic and abstract enterprise, and that the ethicist has a certain power to change the world and might benefit from reflecting on how and where they are situated. Finally, I agree that ethics washing is a challenge and that we as societies would do well if we managed to find some way to make sure that technology worked to promote our common interests.

Whilst agreeing with this, I disagree on how to get there, and will in the remainder focus on (a) how corporations at times seem to subsume and subvert ethics, (b) the need to preserve a distinction between the world of development and engineering, ethics, and politics, and (c) the importance of taking politics—and concepts such as democracy—seriously.

2 Corporate and Academic Ethics

Early on, Van Maanen writes:

> Ethics, also, allows one to strategically 'shop' for the principles that limit one's action as little as possible while simultaneously presenting oneself as contributing towards the common good (Van Maanen, 2022).

Whilst this is true, he stands in danger of confusing the discipline of ethics with strategic corporate behaviour and communication. Is it a problem of *ethics* that corporations cherry-pick what suits their purposes, and potentially even misrepresent it to avoid any heavy lifting whilst being perceived as ethical? And, not least, is it a problem of ethics as an academic disciple if corporations take it upon themselves to make such guidelines?

Part of the problem is that Van Maanen relatively early on determines that a very specific form of ethics is now dominant and mainstream, and then proceeds to use the term "ethics" to refer to this type. The most popular type is claimed to be the one that produces "lists of principles, values, and maxims"—what Bolte et al. (2022) refer to as "checklist ethics".

However, is it correct that checklist ethics is the dominant form of academic AI ethics? All the big tech companies have their own AI ethicists, and the links between these corporation and academia are becoming quite strong (Hagendorff, 2020; Sætra et al., 2021). Whilst there are indeed many checklists and guidelines, this is not proof that academic philosophers are now mainly doing this. Looking at journals

such as *AI & Ethics*[1], most articles are not work of this kind, and the same goes for most other journals on technology and ethics. The same applies to conferences like FAccT[2] and AIES[3], who are increasingly asking for operationalization and deeper engagement with ethics than what is possible through checklists. In sum, Van Maanen's case—Floridi and Cowls (2019)—does not necessarily reflect academia or AI ethics in general. AI ethics is both a more diverse field than Van Maanen seems to believe, and it is also a field in rapid development.

An alternative to the story told by Van Maanen is one in which ethics as a part of moral philosophy is very much alive in academia, but where regulators and businesses have a need to operationalise the philosophers' complex analyses and conclusions. This leads to simplifications, and quite often lists of principles and values. Microsoft, for example, relies on six ethical principles for the development and use of AI: fairness, reliability and safety, privacy and security, inclusivity, transparency, and accountability (Sætra, 2021). Whilst these reflect much of the work done in AI ethics, such lists are often produced by businesses due to an experienced need for "concrete and actionable guidance" (Microsoft, 2022). Law and norms, they argue, have not caught up, and in order to meet their needs they have developed goals, subgoals, and requirements related to six principles, based on their expert teams in "research, policy, and engineering" (Microsoft, 2022).

If "translations" of complex ethics into concrete and actionable principles is necessary, one might also argue that lists and concrete requirements such as those produced by Microsoft are better than vaguer and more complex ethical statements. Whilst ethical principles can be used for ethics washing, it is imperative that such lists are not equated with it. In lieu of legislation that factors in the various social and environmental impacts of AI, it seems unlikely that Van Maanen's solution of converting the army of academic ethicists into casuists with a "thorough understanding of social-political issues" will serve corporations' need for ethical direction.

AI ethics as a field comprises people working from a myriad of perspectives and all varieties of connections with and distance from industry, regulation, and politics. While frustration with checklist ethics is understandable, and while it would indeed be unfortunate if that was all academia was about these days, I do not believe that ethics as a discipline can be judged quite as harshly as Van Maanen does. Most disciplines have a variety of schools, theories, and approaches, and very few condemn, for example, political theory or international relations because someone decides to abuse insight from elements of these disciplines.

---

[1] https://www.springer.com/journal/43681

[2] https://facctconference.org

[3] https://www.aies-conference.com/

3 The Need for Both Ethics and Politics

In Section 2, Van Maanen (2022) discusses the approach of deriving regulation from the myriad of various ethical frameworks, which takes us to the relationship between ethics and politics. This is particularly important for Van Maanen, who argues that ethicists have a responsibility for how their theories, lists, and articles are used—or abused—by, for example, commercial actors. Since the work produced by ethicists have an influence on the world, AI ethics is, Van Maanen argues, *political.*

It would, however, be possible to argue that ethics is about the evaluation and elucidation of the effects of, for example, AI, and that politics and regulation is a distinct domain not necessarily subject to the same logic as the one that applies to ethics (Sætra & Fosch-Villaronga, 2021). If so, deriving politics from ethics could entail making a category mistake.

Lists of ethical principles can easily become "paper tigers", Van Maanen argues. Looks good, but not much bite. This is part of his argument against corporate ethical guidelines, and an interesting question becomes: is this a problem of corporate governance and the enforcement of ethics, or of ethics as a discipline? Various reporting standards, for example, require a company to report on their ethical standards and policies (Sætra, 2021), which generates the need to establish such policies. The policies themselves are perhaps not the problem—the real problem lies in developing and implementing internal routines for effecting such policies. I posit that the lack of enforcement often stems from a lack of motivation and the proper incentives to do so.

As Microsoft (2022) states, laws and norms lag behind development in AI, reflecting the pacing problem described by Downes (2009) and the general problem of establishing social control over technology (Collingridge, 1980). Microsoft has consequently opted to make their own "Responsible AI standard", but this is not externally enforced. The consequences of not internally enforcing this standard are mainly reputational, and if enough stakeholders perceive their actions as unethical enough, they could of course lose out in the market. This is, however, very different from what happens when the missing piece is introduced into the puzzle, namely *politics*. The European Union's approach to the regulation of AI serves as an example of relevant political action (Europan Commision, 2022).

As argued in Sætra and Fosch-Villaronga (2021), ethics is a domain in which the implications of technology are analysed, but this does—and should—not translate directly into policy. Politics is its own domain where society through political processes explores their foundational values and considers how technology might contribute to these through the analyses provided by the domain of ethics. Ethics is here about proposing and analysing, whilst politics is about determining what is desirable and enforcing this. Ethics as a domain does not and should not have regulatory power, and industry cannot and should not be expected to neither have sufficient insight into ethics or politics to restrict their actions on their own or to blindly follow the directions of academic ethicists. The latter is

because ethicists have expertise in ethical analysis, but they are not a source of political legitimacy. If we value democracy, pluralism, and the agonistic element of politics—and do not want a meritocracy of ethicists—*politics* is the domain in which people get involved and jointly decide on questions of technology (Sætra & Fosch-Villaronga, 2021).

4 Politicised Ethics or a Stronger Division of Labour?

In his proposed alternative to ethics as now know it, Van Maanen relies on four questions as the basis for a "question-based" politicised ethics. The questions would indeed entail including broader social-political issues rather than just isolated technical questions in ethics. But this approach also turns the ethicist into a sort of political activist, as it requires them to consider their own "theory of change" and to navigate questions related to the timing of their actions, who they engage with, and who benefits from what (Van Maanen, 2022). Even if we accept that ethics is political since it can have effects on the world, we do not necessarily need to incorporate politics in ethics. This also holds for the domain of science and engineering. Technology is political (Winner, 1977), but this does not mean that all engineers must be guided by a "thorough social-political" understanding of the implications of their work (Van Maanen, 2022). This seems practically impossible, and I posit that it is also not desirable. Only through a proper division of labour between science, ethics, and engineering can we arrive at a situation in which technology contributes to achieving our goals as a society (Sætra & Fosch-Villaronga, 2021). By giving politics its due, we ensure that questions of value and ethics are not determined by arbitrary individuals, as also stressed by Van Maanen.

Furthermore, there is a risk that Van Maanen's form of "ethics"—which is quite simply a broadening of ethics to include political philosophy, STS, etc.—makes it even more problematic to assume that an engineer, for example, can be expected to base all their decisions on such fundamental questions. This is, however, more a practical question than a normative one. But is Van Maanen dealing with an ideal situation or a practical and "realistic" one? From an article explicitly arguing against ideal theory in favour of a realist counterpart, there is a need for further development of how the proposed politicised ethics would pan out in practice. Who would the casuists be, where would they be positioned (academia or industry), and what sort of authority to make political decisions would Van Maanen like to provide them with? This also relates to work in, for example, environmental ethics, where philosophers such as Næss (1989) warned of a form of "ecologism" emerging from making technical expertise the precondition of being able to say what is *right* or *valuable*, whilst stressing the need for political mobilisation and action.

In closing, I believe that Van Maanen's article provides interested practitioners and researchers alike with important theoretical input from political philosophy, which is both necessary and welcome in what could at times be described as an AI ethics bubble. As should be clear from this commentary, however, I do not accept the proposed conflation of ethics and politics. I accept the practical difficulties of demarcating and disentangling science, ethics, and politics, but it is nevertheless crucial to maintain politics as a distinct sphere, as highlighted by several of the political theory realists

(Galston, 2010). Politics requires judgement which is not determined by economic, legal, or moral principles (Galston, 2010). Furthermore, politics provides the basis for legitimacy and democratic decision-making, something that would not be achieved by replacing ethicists with political-moral experts left to engage in casuistry. Van Maanen (2022) states that ethics has a "disappointing capacity to stop, redirect, or at least slow down big tech's course", but this is more suggestive of the need for more political engagement with technology. On that, I very much agree with Van Maanen, but I do not think the ethicists are obsolete.

Materials and/or Code Availability Not applicable.

Data Not applicable.

Declarations

Ethics Approval Not applicable.

Consent Not applicable.

Competing Interests The author declares no competing interests.

References

Bolte, L., Vandemeulebroucke, T., & van Wynsberghe, A. (2022). From an ethics of carefulness to an ethics of desirability: Going beyond current ethics approaches to sustainable AI. *Sustainability, 14*(8), 4472. https://doi.org/10.3390/su14084472

Collingridge, D. (1980). *The social control of technology*. London: Frances Pinter.

Downes, L. (2009). *The laws of disruption: Harnessing the new forces that govern life and business in the digital age*. New York: Basic Books.

Europan Commision. (2022, Jun 22). A European approach to artificial intelligence. Retrieved from https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *HDSR*(1.1). https://doi.org/10.1162/99608f92.8cd550d1

Galston, W. A. (2010). Realism in political theory. *European Journal of Political Theory, 9*(4), 385–411. https://doi.org/10.1177/1474885110374001

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines, 30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Microsoft. (2022). *Microsoft Responsible AI Standard, v2: General requirements*. Retrieved from https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4ZPmV

Næss, A. (1989). *Ecology, community and lifestyle: Outline of an ecosophy*: Cambridge university press.

Sætra, H. S. (2021). A framework for evaluating and disclosing the ESG related impacts of AI with the SDGs. *Sustainability*.

Sætra, H. S., Coeckelbergh, M., & Danaher, J. (2021). The AI ethicist's dilemma: Fighting Big Tech by supporting Big Tech. *AI and Ethics*. https://doi.org/10.1007/s43681-021-00123-7

Sætra, H. S., & Fosch-Villaronga, E. (2021). Research in AI has implications for society: How do we respond? *Morals & Machines, 1*(1), 60–73.

Van Maanen, G. (2022). AI ethics, ethics washing, and the need to politicize data ethics. *Digital Society*.

Winner, L. (1977). *Autonomous technology: Technics-out-of-control as a theme in political thought*. MIT Press.