

Book Review

‘Doing’ technology ethics: Review of Marc Steen’s ‘Ethics for people who work in tech’

Andrew P. Rebera^{1, 2,*}

¹ Institute of Philosophy, KU Leuven, Leuven, Belgium; andrew.rebera@kuleuven.be

² Royal Military Academy, Brussels, Belgium; andrew.rebera@mil.be

* Correspondence: rebera.andrew@gmail.com

Keywords: ethics; technology; consequentialism; deontology; virtue ethics; relational ethics; human-centred-design; value sensitive design; responsible innovation

Ten years ago, pundits claimed that 3-D printing would change the world (D’Aveni, 2013). That this is, to put it kindly, not true *yet*, goes to show that claims about the impact of new technologies are always highly speculative. Today, large language models (or tools based on them) are the latest technology to promise or threaten disruption in almost every walk of life, from work to education to art. But it remains to be seen how broad, deep, and lasting the impact will be. If uncertainty surrounds even the short-term impact of technologies that, though new, are *at large* in society, how much more difficult is it to anticipate the ways in which technologies *not yet* at large might fit into, reshape, or upturn our lives? There is no shortage of work on such issues. Yet Marc Steen’s *Ethics for people who work in tech*, published this year by CRC Press, is a welcome addition. Steen aims to empower tech professionals—an ill-defined but arguably underserved group—to manage, and take responsibility for, the impact of the technologies they develop.

Steen presents his credentials up front. Rightly so, for experience lends credibility, and Steen has plenty of both. The reader benefits from his 25 years’ experience in research, design, consultancy, and project management (p. 4). From this insider position, Steen addresses the reader directly, remaining constantly at or near the surface of the writing. This lends the text an informal, conversational feel that could easily become cloying, though to the author’s credit it does not. Nor is Steen coy in setting out *his* perspective on ethics and innovation. Again, such an approach could easily backfire, but here it works.

Steen locates himself within a humanistic approach which “takes human dignity and human autonomy as starting points for the design and application of technology” (p. 4). Ethics and responsibility are understood as “happening within people and between people” (p. 6). Ethics itself is a *process* (*‘doing ethics’*) rather than a set of norms, assertions, or answers (p. 8). It *may* be thought of as a domain of knowledge (p. 15), but such knowledge is best thought of as a kind of skill or know-how. Accordingly, for Steen, ‘doing ethics’ is less a search for answers than a process of pointed questioning—of ourselves, our assumptions, our situatedness, and so forth. Realism is rejected since “people construct morality through social interactions and processes” (p. 20). Such bold statements are apt to leave some readers cold (extensive background knowledge is not assumed), yet Steen is ever considerate: he begins the subsection immediately following, “Now, how is this relevant for your work?” (p. 18). The answer, or part of it, is that research and innovation are partly driven by economic and political forces (p. 19) and that, therefore, ethics—as subject and activity—sits naturally alongside science and technology, complementing them with processes of “reflection, inquiry, and deliberation” (p. 20) in which facts *and* values are on the table and, importantly, in which people of all backgrounds (‘hard’ or ‘soft’ domains) can collaborate.

Citation: Rebera, Andrew P. 2023.

Title. *Journal of Ethics and Emerging Technologies* 33: 1.

<https://doi.org/10.55613/jeet.v33i1.126>

6

Received: 30/05/2023

Accepted: 01/06/2023

Published: 30/06/2023

Publisher’s Note: IEET stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

The remainder of Part I seeks to broaden the reader's view, introducing different perspectives on socio-technological questions and issues: is technology neutral? how can it promote value and wellbeing? the limitations of thought experiments such as the trolley problem; and the distinction between privacy and data protection. Along the way, a variety of thinkers, thoughts, and theories are surveyed (the treatment is cursory, but the interested reader can follow up leads via the helpfully annotated bibliography).

Part II introduces the reader to consequentialism, deontology, relational ethics, and virtue ethics. These are not presented as exclusive, but combinable "according to what the project at hand requires" (p. 61). This 'buffet' approach is not without tensions: the perspectives, he acknowledges "are different" and we "do not want to confuse them" (p. 136). Yet what 'confusing them' would amount to is not clearly explained. Still, Steen is not advocating that one become, by turns, a committed consequentialist, deontologist, and so on, but simply that we exploit each view's distinctive insights. It is *perspectives* that Steen offers, not *theories*.

That said, his sympathies are apparent. Chapters on consequentialism and deontology include 'Limitations and critique' sections; chapters on relational and virtue ethics do not. Relational ethics is presented as a collection of views challenging Enlightenment assumptions. Approaches from Africa, Australasia, South America, and Asia are surveyed (pp. 101-103), alongside figures such as Levinas, Michael Sandel, Alasdair MacIntyre (pp. 105-106), and others from the ethics of care (pp. 106-108). The chapter ends with a nod to scholars applying relational approaches to technology, including Abeba Birhane, Mark Coeckelbergh, Virginia Dignum, Charles Ess, David Gunkel, and Kate Darling (p. 111). A helpful subsection on 'practical applications' is offered, in which Steen gives examples from his own projects. Of one he says, "Without knowing, we did something like relational ethics" (p. 101), which raises the question whether the insights of this part of the book needed to be presented in theoretical garb. If one can 'do' relational ethics without getting into the theory of it (e.g. without self-consciously questioning Enlightenment assumptions), then why not streamline the process by leaving theory aside? Yet this shows the spirit of Steen's project. The point is not merely to provide a method, but to draw the reader into a style of thinking. A little bit of knowledge, an orientation in the field, will help. Recognizing, for instance, that Enlightenment assumptions are just that, gives one license to interrogate, and maybe think outside and beyond them.

Virtue ethics is presented as robustly social and relational. *Eudaimonia* is defined as "living well *together*" (p. 118, my emphasis)—which is at least unusual. Subsections on the concepts *polis* and *telos* reinforce this view, while a subsection on *ethos* (but not *arete*, again unusually) introduces virtues as dispositions to think and act in ways that target excellences of particular kinds. The view is presented in a very positive light (especially compared to consequentialism and deontology). Indeed, two important objections are omitted. The first is the 'situationist critique' that virtues and character exert less influence on behavior than virtue ethicists suggest (e.g. Doris, 2010): factors such as mood or the presence of authority figures cannot be ignored. The second objection is that, whereas Steen praises virtue ethics for being "aspirational" (p. 123), others have noted that virtues are rare and hard to attain. Steen inadvertently gestures toward the worry when suggesting that "If [...] you find the vocabulary of *virtues* archaic, you can think of them as *superpowers*" (p. 126). Is it realistic to strive after superpowers? If we fall short of full virtuousness, what then? Steen mobilises Shannon Vallor's (2018) arguments against the applicability of consequentialism and deontology to *emerging* technologies, of which details concerning functionality and use are typically lacking (p. 124). But when the virtues are not fully developed, it remains unclear why they should be a better guide to decisions in technology development than, say, reflection on consequences, duties, or rights.

Part III introduces methods to integrate ethics into design processes: human-centred design (HCD), value sensitive design (VSD), and responsible innovation (RI). The chapter

on HCD focuses on the key concepts *iteration* (emphasizing the cyclical nature of defining and addressing problems) and *collaboration* (methods for stakeholder engagement and ‘participatory design’). The challenges to implementing such approaches, particularly in the face of organizations, work structures, or colleagues who prefer (or only know how) to work in a more linear fashion are not ignored. Indeed, awareness of such difficulties will help in finding ways to overcome them (p. 152). Steen says, “You will need to balance your own ideas with other people’s ideas. You can consciously decide to sometimes stick to your own ideas and at other times to privilege other people’s” (p. 152). When to concede and when to insist is, of course, a tricky question. Steen is likewise open about the difficulties of VSD. Values may be in tension; decisions on which and whose values to prioritize are always already ethically loaded; values may be mislaid or changed through design processes; and value-talk can be overly abstract (pp. 157-158). RI is initially introduced as having “the ambition to (better) align innovation projects to needs and concerns in society” (p. 162). Steen enriches this to a vision of innovation as creating “conditions in which people can extend relevant capabilities and cultivate relevant virtues, so they can live well together and flourish” (p. 162)—a further indication of his relational and virtue ethical sympathies. RI’s ‘key ingredients’ are *ethical reflection* and *societal engagement* (p. 162). Steen focuses on the latter, usefully providing a list and discussion of advantages of, and challenges to, achieving it. The ‘key dimensions’ of RI are *anticipation*, *responsiveness*, *diversity*, and *reflexivity* (p. 165). The discussion of these concepts is one of the most successful parts of the book, showcasing what is best about Steen’s approach and style: clear explanations, useful examples, and practical advice on implementation.

This is not a book for scholars—detailed analysis of figures or theories discussed in it must be sought elsewhere—but nor is it intended to be. For people who are new to addressing ethical and societal issues in tech, and for those who have experience but lack an overarching framework or method, this is a valuable resource. Its overall project—practical advice on how to *think about* and *implement* ethics in tech—is well-conceived. Steen delivers on the project too: ideas are presented accessibly, without dumbing-down; technical terminology is introduced but neither relied on nor hidden behind; in style, the author is present but not overbearing, advises without preaching, instructs without talking down.

A few limitations may be noted. While the book serves its target audience, it serves some sections more than others. This is partly because the target audience is vast: people in computer and data science; software development; engineering; design; business, marketing or management roles; procurement; and policy making (p. i). Moreover, what are the boundaries of ‘tech’ as used in the book’s title? *Technology* is an extremely broad category (a dovetail joint is a technology, but Steen isn’t addressing carpenters). Another way in which some audience sections are better served, is that many of the approaches Steen recommends will often only be implementable by people with direct responsibilities for ethical or societal issues, and then only with the support of relatively senior figures in their organization. Commenting on how to promote HCD, VSD, or RI he suggests: “You can allocate budgets to efforts like this” (p. 176). A common response is likely to be: ‘But I don’t have any control over budgets’.

Steen is well-aware of such issues. He acknowledges, for instance, the tension between individual moral agency and structural factors (p. 57), noting the need to “identify and analyze enablers and barriers that are at play [...] and to find ways to *utilize* enablers and *reduce* barriers” (p. 176, my emphasis). For many people, especially those lacking job-security, that is easier said than done. As he rightly notes, “the tech sector can be particularly unwelcoming (toxic, even) to diversity, for example, in terms of hiring people or promoting people” (p. 170). It would have been of *significant* added value had Steen offered examples of how to approach such challenges. That said, there are no easy answers here, and it is to Steen’s credit that the problem is not ignored.

A second critical point concerns agency and responsibility. Steen gives several examples of beneficial, low-cost, low-disruption changes to working practices. We learn, for instance, of project managers making space in meetings to discuss and, crucially, *question* project assumptions and objectives (pp. 21-22). Such measures rely on people with the power and will to implement them. But in many innovation projects, even if the will is present, there is simply no scope for acting on any but the most conservative proposals for change. In many cases, the scope of work, tight deadlines, and contractual obligations to deliver *already-defined* outputs, preclude the kind of revisions that might be recommended. To have any chance of being funded, grant proposals often require a degree of precision in defining objectives, methods, and deliverables that forces applicants into premature decisions on important issues. The weight of sunk costs to organizations, and the extent of career capital invested in such projects, make changes of direction prohibitively costly. Once again, Steen does not conceal the difficulties (e.g. p. 146). But they point to a difficulty in balancing knowledge, agency, and responsibility. On these, Steen writes: “the more knowledge you have, the greater your responsibility for your actions, and the larger your agency, the more knowledge you will need to have in order to act responsibly” (p. 56). But this leaves the person who, having read Steen’s work, has much knowledge and little agency, in an awkward position. Where then does personal responsibility begin and end? What to do when one knows that change is required but is impotent to effect it?

A final critical point is that the relationship between ethics and law receives too little attention. People working in ethics will often find common cause with people working on legal issues – and may face similar challenges. On this theme, Steen bemoans the tendency to reduce issues of privacy to issues of data protection (p. 51). But the relevance of the ethics-law relationship to the practicalities of ‘doing ethics’ is not simply that ethical concepts get reduced to legal ones, but that ethics is mischaracterized as a compliance issue. In Steen’s defense, the disanalogy between legal and moral duties is briefly discussed in chapter 11; and his recommendations regarding cross-disciplinary collaboration may include tools with which misconceptions can be overcome. I am not unsympathetic to that suggestion. It would, nevertheless, have been helpful to focus more attention on a barrier that, plausibly, many people implementing Steen’s approach will encounter.

In closing I return to the opening chapter. One of Steen’s aims is “to help you ‘upgrade’ your moral sensitivities” (p. 6). This involves becoming more aware of and knowledgeable about certain kinds of issues; and becoming *more disposed to notice* them. Although the book discusses a range of issues raised by a range of technologies, it offers no systematic list of, say, *key issues in AI ethics* or the like. But it isn’t that kind of book. So, much of the ‘upgrading’ of the reader’s moral sensitivities will be due to their adopting ways of thinking, perspectives, and—crucially—*practices*, that better attune them to relevant considerations. It is this insistence on practice—on ‘doing ethics’—that is the book’s greatest strength. It is also why it would be counterproductive to simply list issues in, say, AI ethics. Lists are static. The point is to actively investigate and uncover issues, to create solutions to them, rethink them, and so on. Ethics, Steen notes at various points, is like innovation in being a process demanding curiosity and creativity. Both deal “with envisioning and creating possible, desirable futures” (p. 141). What he ultimately provides is a framework for implementing practices that insist on, and make space for, informed, collaborative, structured engagement in the creative process of aligning technology with people’s values. Of course, such a framework offers no guarantees; and, even fully implemented, there remain several significant challenges (some of which are mentioned above). But no book could resolve systemic difficulties of these kinds; and it is a strength of this one that it neither obscures the problems nor offers superficial solutions. Rather, it focuses on doing what one can do and doing it well: working on oneself, enabling collaboration, and aligning technological and business goals with societal values. Underlying the entire project of the book is a message of faith and hope: faith that people

in tech are motivated to do good; hope that they can manage it (p. 59). The book aims to support them in that increasingly important task. Does it succeed? I think it does.

Author Contributions: n/a

Funding: This research received no external funding.

Institutional Review Board Statement: n/a

Informed Consent Statement: n/a

Data Availability Statement: n/a

Acknowledgments: n/a

Conflicts of Interest: The author declares no conflict of interest.

References

- D'Aveni, R. A. (2013, March 1). 3-D Printing Will Change the World. Harvard Business Review. <https://hbr.org/2013/03/3-d-printing-will-change-the-world>
- Doris, J. M. (2010). Lack of character: Personality and moral behavior. Cambridge University Press.
- Vallor, S. (2018). Technology and the Virtues. Oxford University Press.