oekom science

Book review: Nyholm, Sven (2023): This is technology ethics. An introduction

Michael W. Schmidt*,1

Have you been surprised by the recent development and diffusion of generative artificial intelligence (AI)? Many institutions of civil society have been caught off guard, which provides them with motivation to think ahead. And as many new plausible pathways of socio-technical development are opening up, a growing interest in technology ethics that addresses our corresponding moral uncertainties is warranted. In Sven Nyholm's words, "[t]he field of technology ethics is absolutely exploding at the moment" (p. 262), and so the publication of his introduction to the topic is well timed.

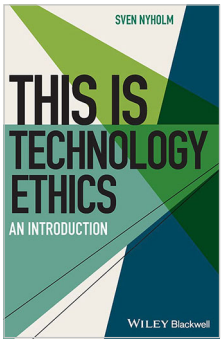
Nyholm starts off in good analytic fashion by clarifying the basic concepts of the subject matter, namely what we should understand by 'technology' and 'ethics' (chapters 1 and 2). In doing so he succeeds in fruitfully combining a reflective self-understanding of the field with classic topics from the philosophy of technology, such as a discussion of Heidegger's "The question concerning technology". However, an interesting and useful addition, especially for readers new to the field, might have been a discussion on how technology ethics relates to other fields such as the ethics of engineering or the philosophy of technology and engineering. These fields cannot be adequately differentiated with a purely conceptual perspective, since a proper understanding of them as social constructs requires knowledge of the recent history of ideas including the dynamics of research fields (e.g. Poznic 2024).

When introducing the main theoretical approaches in normative ethics such as deontology, consequentialism and virtue ethics, Nyholm is very sensitive to the diversity of moral thought, and also presents confucianism and ubuntu ethics which are explored further in later parts of the book (pp. 32-36). In my largely Rawlsian view, however, the due consideration of reasonable pluralism leads to further questions: When is it legitimate to opt for the specific but contested ethical approach one personally finds most plausible on reflection, and when should we instead refer to shared values? There are approaches in applied ethics that at least claim to be based on some kind of overlapping consensus that could have been added fruitfully, such as reasoning based on mid-level principles (Beauchamp and Childress 2019) or some version of the capabilities approach (Nussbaum 2007) but Nyholm doesn't include them.

A strong point of Nyholm's book is an explicit discussion of the methods of technology ethics (chapter 3). Nyholm is right in presenting the method of reflective equilibrium as a kind of basic methodological consensus and in declaring that it is a type of meta-method that can incorporate further methods such as the usage of thought-experiments or procedural approaches, that Nyholm refers to as "ethics by committee" (pp. 56-58). Of course, there is room left for criticism: As Nyholm suggests, it is not very controversial to commit oneself to the method of reflective equilibrium, however, this commitment is also not completely trivial as it excludes, for example, strong foundationalist approaches to ethics (Schmidt 2024). A useful extension of this point might have been that especially technology ethics has created interesting experimental approaches that apply the method of reflective equilibrium not purely from the armchair perspective but include stakeholders in workshop settings in the pursuit of reflective equilibrium (van de Poel and Zwart 2010; Doorn and Taebi 2018; Brandstedt et al. 2024). To me at least, this is of great importance, since I believe that technology ethics, ideally, should be practiced in such a way that it has direct transformative impact on the design of our socio-technical environment. All in all, however, Nyholm advocates a very plausible methodological pluralism for technology ethics by highlighting the benefits of a free division of epistemic labor.

A focus on digital technologies

The main part of the book is concerned with the discussion of contemporary topics in technology ethics, such as value alignment for AI (chapter 4), behavior change technologies and autonomy (chapter 5), responsibility gaps by automation (chapter 6), machines as moral agents and patients (chapters 7 and 8) or as "friends, lovers, and colleagues" (chapter 9), and ends with the discussion of trans- and posthumanist ideas (chapter 10). Of note is that these topics are all focused on digital technologies.



Nyholm, Sven (2023):
This is Technology Ethics: An Introduction.
Hoboken, NJ: Wiley-Blackwell.
288 pp.,
Paperback 33,99 EUR / E-Book 36,99 EUR,
ISBN 9781119755579

* *Corresponding author: michael.schmidt@kit.edu*
[1] *Institute for Technology Assessment and Systems Analysis, Karlsruhe Institute of Technology, Karlsruhe, DE*

https://doi.org/10.14512/tatup.7152
Published online: 12. 12. 2024

However, between rather ancient technologies that are, according to Nyholm, romanticized by Heidegger, and contemporary digital technologies, there are other – sometimes rather mundane – technologies, that also need ethical scrutiny. Nyholm includes such technologies from time to time, for example when he notes that "[. . .] if cars are equipped with safety features modeled on male drivers, this might systematically disadvantage female drivers, who might then enjoy less protection of their safety in crashes" (p. 89). This extends, of course, to the standards with which we test for sufficient safety (Linder and Svensson 2019) which is an interesting topic of its own. Discussion of a greater variety of technologies and their social setting might have been beneficial for an introductory book on technology ethics. Even focusing only on emerging technologies could have broadened the discussion: think of gene-editing, as an example. However, as potentially revolutionary current digital and AI-driven technologies are not covered by older introductory sources, a focus on these does provide a distinguishing feature of Nyholm's book.

Nyholm's book exemplarily shows how technology ethics can address difficult issues in a way that enhances reflexivity and thus provides normative orientation.

With regard to Nyholm's focus on AI-driven technologies there is one odd detail: Nyholm seems to equate narrow AI with weak AI, and general AI (AGI) with strong AI (p. 81). While this is a quite common understanding in public discourse and academic literature, it would make sense to differentiate from a philosophical and systematic perspective. If we say that an AI system can operate successfully not only with respect to a very specific task (narrow AI) but with regard to a very wide range of tasks (AGI), then this is different than saying that an AI system has some cognitive capacities like consciousness or sentience (strong AI), or that it does not have these capacities (weak AI) (for the origin of the latter distinction see Searle 1980). There should be conceptual space for accepting an AI system as AGI while refraining from accepting it as strong AI, with all the possible ethical implications.

Despite the issues that I have highlighted in critical spirit in this review, Nyholm's introduction to the ethics of technology is engaging throughout, and exemplarily shows how technology ethics can address difficult issues in a way that enhances reflexivity and thus provides normative orientation. It is essential reading for newcomers to the field and will be interesting also for established scholars. Moreover, in its style and approach it is a great complement to other introductory sources. The book also comes with a special treat: There is a mini podcast series in which John Danaher and Sven Nyholm intriguingly discuss the chapters of the book in detail (Danaher and Nyholm n.d.), and within the book the interested reader will find many links to podcast episodes or videos that deepen the topic at hand.

References

Beauchamp, Tom; Childress, James (2019): Principles of biomedical ethics. New York: Oxford University Press

Brandstedt, Eric; Busch, Henner; Lycke, Ellen; Ramasar, Vasna (2024): Winds of change: An engaged ethics approach to energy justice. In: Energy Research & Social Science 110 , p. 103427. https://doi.org/10.1016/j.erss.2024.103427

Danaher, John; Nyholm, Sven (n.d.): Technology ethics podcast. Online available at https://technologyethicspod.wordpress.com/, last accessed on 12.09.2024.

Doorn, Neelke; Taebi, Behnam (2018): Rawls's wide reflective equilibrium as a method for engaged interdisciplinary collaboration. Potentials and limitations for the context of technological risks. In: Science, Technology, & Human Values 43 (3), pp. 487–517. https://doi.org/10.1177/0162243917723153

Linder, Astrid; Svensson, Mats (2019): Road safety. The average male as a norm in vehicle occupant crash safety assessment. In: Interdisciplinary Science Reviews, 44 (2), pp. 140–153. https://doi.org/10.1080/03080188.2019.1603870

Nussbaum, Martha (2007): The capabilities approach and ethical cosmopolitanism: a response to Noah Feldman. In: The Yale Law Journal 117, available online at https://www.yalelawjournal.org/forum/the-capabilities-approach-and-ethical-cosmopolitanism-a-response-to-noah-feldman, last accessed on 21.11.2024.

van de Poel, Ibo; Zwart, Sjoerd (2010): Reflective equilibrium in R & D Networks. In: Science, Technology, & Human Values 35 (2), pp. 174–199. https://doi.org/10.1177/0162243909340272

Poznic, Michael (2024): Analytische Technikphilosophie. In: Mathias Gutmann, Klaus Wiegerling and Benjamin Rathgeber (eds.): Handbuch Technikphilosophie. Berlin: Metzler, S. 111-120.

Schmidt, Michael (2024): Defining the method of reflective equilibrium. In: Synthese 203 (5), p. 170. https://doi.org/10.1007/s11229-024-04571-6

Searle, John (1980): Minds, brains, and programs. In: Behavioral and Brain Sciences 3 (3), pp. 417–424. https://doi.org/10.1017/S0140525X00005756