Toward a code of conduct for technology ethics practitioners

Vermaas, Pieter; Ammon, Sabine; Mehnert, Wenzel

DOI
10.1080/23299460.2024.2440958

Publication date
2025

Document Version
Final published version

Published in
Journal of Responsible Innovation

Routledge
Taylor & Francis Group

Journal of Responsible Innovation

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjri20

Toward a code of conduct for technology ethics practitioners

Pieter Vermaas, Sabine Ammon & Wenzel Mehnert

To cite this article: Pieter Vermaas, Sabine Ammon & Wenzel Mehnert (2025) Toward a code of conduct for technology ethics practitioners, Journal of Responsible Innovation, 12:1, 2440958, DOI: 10.1080/23299460.2024.2440958

To link to this article: https://doi.org/10.1080/23299460.2024.2440958

Published online: 20 Dec 2024.

JOURNAL OF RESPONSIBLE INNOVATION
2025, VOL. 12, NO. 1, 2440958
https://doi.org/10.1080/23299460.2024.2440958

Routledge
Taylor & Francis Group

DISCUSSION

OPEN ACCESS Check for updates

Toward a code of conduct for technology ethics practitioners

Pieter Vermaas [a], Sabine Ammon[b] and Wenzel Mehnert[b]

[a]Philosophy Department, TU Delft, Delft, the Netherlands; [b]Berlin Ethics Lab, TU Berlin, Berlin, Germany

ABSTRACT

This paper explores possibilities for creating a code of conduct for practitioners working in technology ethics. The number of technology ethics practitioners is growing when including ethicists embedded in research projects on technology, members of research ethics committees who assess the consequences of technological research, ethicists advising companies, and facilitators in moral/societal exploration through workshops, games, and brainstorm sessions. And, it is not just assessment what technology ethics practitioners do, but increasingly also the guidance of technology development through processes like responsible research and innovation, ethics by design, and design for values. Our exploratory is initial, and this paper may be seen as a position paper. It focusses on preliminary issues such as identifying the types of ethics practitioners the code can be for, the roles the code can play for these practitioners, charting controversies it should address, and the (institutional) arrangements needed for making a code effective.

ARTICLE HISTORY
Received 13 March 2024
Accepted 5 December 2024

KEYWORDS
Ethics; code of conduct; technology ethics; practitioners; integrating ethics

Introduction

This paper explores possibilities for creating a code of conduct for practitioners working in technology ethics. The integration of ethics into the development of technologies is meant to transform the way technology is developed. But it is also changing ethics. Ethics is not anymore confined to the academic schools of philosophy, but ethicists work with or in laboratories, R&D institutes, and technological companies. And researchers and engineers working in technology development who may never have entered those schools can become active and productive ethicists. The number of technology ethics practitioners is moreover growing when including ethicists embedded in research projects on technology, members of research ethics committees who assess the consequences of technological research, ethicists advising companies, and facilitators in moral/societal exploration through workshops, games, and brainstorm sessions. And, it is not just assessment of what technology ethics practitioners do, but increasingly also the guidance of technology development through processes like responsible research and innovation, ethics by design, and design for values.

CONTACT Pieter Vermaas p.e.vermaas@tudelft.nl TU Delft, section Ethics and Philosophy of Technology, Jaffalaan 5, 2628BX Delft, the Netherlands

We welcome these developments and aim with our exploration to support them by articulating and strengthening technology ethics through a code of conduct. If ethics remains to be seen as an academic endeavor, standards for academic research provide some guard rails for researchers in technology ethics. As soon as it is seen as a profession in which practitioners give assessments, advice, and guidance, more support is needed. Practitioners as well as their clients can then use standards of expertise that describe obligations towards clients and societies, set norms for professional conduct, and help in case of conflict. We believe that a professional code of conduct can deliver those standards.

We are not the first to observe that ethics itself may be ready for normative guidance. Already Hansson (2017) argued for an ethics of ethics, for academic research and for ethics consultancy, ending with a proposal to create ethical codes for subdisciplines of ethics. With our exploration, we want to reaffirm the observation and proposal. As researchers, we are ourselves working in technology ethics, in academic research,[1] in tool development for ethical analysis and guidance,[2] and in projects on technology development.[3] We partake in this professionalization of technology ethics and acknowledge that it needs normative guidance over and above standards for academic research.

The exploratory is initial, and this paper may be seen as a position paper. It focusses on preliminary issues such as identifying the types of ethics practitioners the code can be for, the roles the code can play for these practitioners, charting controversies it should address, and the (institutional) arrangements needed for making a code effective.

Technology ethics practitioners

We start by characterizing the types of work technology ethics practitioners have taken up. In general, these practitioners are working on a project basis in research and development processes, often temporarily limited. This research and development typically takes place in research institutions, in administration, and in industry. The technologies the practitioners are working on can be of all readiness levels, leading to a wide spectrum of types of work environments, from foundational research on emerging technologies to bringing mature technological applications to market.

A code of conduct could be of use to all these technology ethics practitioners, and for further articulating this scope, we consider three paradigm cases for capturing three different key tasks an ethicist can have, namely being a partner, a facilitator, or a consultant.

(1) As a partner, the ethicist is embedded in the team for the duration of the project. The ethicist contributes their knowledge and methodological competencies to the joint project and ethics expertise stands on equal footing with the expertise of the other team members.
(2) As a facilitator, the ethicist joins a project team for an ethical intervention or a series of interventions but is not themself part of the project team. E.g. by conducting workshops with a company or organization, the ethicist supports the project team in identifying and addressing ethical issues of the ongoing project.
(3) As a consultant, the ethicist advises on or evaluates ethical issues as an external expert. The ethicist is not directly involved in the research and development process and is not part of the project team.

The roles of a code of conduct

We take a code of conduct as

> an agreement among members of a professional association, umbrella group, or a single organization, in which they agree to act in a certain way. Such codes are typically developed in professions or trades that are not regulated by a governmental institution.[4]

This agreement includes actions of how members act towards each other and the association, as well as how they act towards external stakeholders.[5]

As said, when practitioners working in technology ethics are taken as academics, there is community agreement as captured by codes for good scientific practices. Yet, it seems that codes of conduct for academia have proved to be insufficient in avoiding a number of more controversial practices in technology ethics, such as ethics washing (see the next section). Hence, there is a need for a more targeted agreement about the actions of technology ethics practitioners. And, acknowledgment that such practitioners can also work outside of academia – and when discerning the three tasks ethics practitioners can have, it makes more sense to do so – gives further support for aiming at a code of conduct separate from those for academic researchers. A code of conduct for technology ethics practitioners can furthermore play a useful role in agreement about actions toward external stakeholders such as clients and societies.

The first role we can imagine is that a code of conduct can express the values we as technology ethics practitioners hold. This is not a metaethics question (which would introduce all academic discussions in ethics into discussions about a code, and obviously derail any attempt to arrive at consensus) but one about how we as professionals want to operate, from honesty to transparency, a commitment to keeping up our knowledge bases, and helping clients with taking into account the interests of society.

A second role a code of conduct could play concerns our relationship with clients. When entering a supporting or commercial relationship as embedded ethicists in technological research or as advisers to companies, it should be clear to all parties concerned what the mutual responsibilities are. Clients may expect that the ethics practitioners have state of the art knowledge, but should also accept that this knowledge may lead to results and advice that can counter the direct goals of the clients. Technology ethics practitioners also have responsibilities towards other stakeholders and towards society and may conduct research for these stakeholders and for society. And, it should be clear between the client and ethics practitioner how results and advice are implemented.

This leads us to a third role a code of conduct can play. For arriving at an effective and constructive profession, we do not want the work of technology ethics practitioners to end up in the metaphorical drawer, nor that we become window-dressers or ethics washers. Controversial practices and scandals are often the basis for arriving at codes of conduct (e.g. Vrielink and van Montfort 2010) and our field of technology ethics has those as well, as we discuss in the next section. A code of conduct can lay down what we see as good technology ethics practices, give standards and role models for technology ethics practitioners, introduce measures to defend well-acting practitioners and avoid controversial ones, and support that the results of technology ethics practices become advice clients act on.

Controversies

The third role a code of conduct can play includes helping practitioners navigate through the controversies and pitfalls that have emerged around the practice of technology ethics. These controversies are manifold and cover, among others, topics like checkbox ethics, ethics washing and the instrumentalization of ethics experts.

Checkbox ethics refers to a method in which the potential ethical implications of new technologies are evaluated according to given and fixed ethical principles and rules (e.g. fairness, privacy, or other values). It aims at translating philosophical principles into engineering practice, thus helping developers to apply guidelines or regulations during the development process. Although the approach has clear benefits (e.g. reducing complexity), evaluation of these approaches has shown that the use of checklists does not necessarily lead to its envisioned impact. Instead, if not accompanied by prior training, it creates just another burden for developers (Schiff et al. 2021) and can promote a 'checkbox culture' (Balayn et al. 2023) where pre-established categories are blindly tick-boxed without reflecting the specific sociotechnical complexity of developing and deploying emerging technologies. This standardized approach disregards the sociotechnical complexity of developing and deploying an emerging technology (Kiran, Oudshoorn, and Verbeek 2015).

Particularly in the private sector, ethics can be seen as a mere communications strategy to cover-up or façade unethical behavior (Bietti 2020). This is called 'ethics washing' and means pretending to engage in ethical considerations or activities with the only purpose of improving how a company or application is perceived but not creating actual changes or actions; it means being loud about ethics but without taking serious actions to implement meaningful change (Buedo and Waligora 2022). Ethics washing is strongly connected to the self-regulatory efforts in private companies. In situations where governmental regulations are missing, companies develop or use predefined guidelines presenting oneself as contributing towards the common good (Wagner 2018). However, as the companies can decide on the guidelines, they are allowed to select the principles that limit one's action as little as possible while still claiming to act ethically (van Maanen 2022).

The controversies in the field are manifold and the examples given are just a selection. Our aim is not to discuss the controversies in detail, but to bring the field further by offering guidance on how to navigate the struggles that the practice of integrating ethics brings along. As argued several times (e.g. Kiran, Oudshoorn, and Verbeek 2015),[6] integrating ethics is not done by using guidelines or checklists. One approach is to collaborate with technology ethics experts that help to identify issues (e.g. through their own expertise or by bringing in external knowledge). Based on our experience, this collaboration can have different forms and that also brings along different controversies.

Working as partners in research and development projects ideally means that ethicists meet developers on an equal footing and implies the right to intervene and take action when necessary. Otherwise, ethicists become toothless tigers, writing reports to fulfill compliance regulations or the requirements of the research proposal. Based on our own experience, this can lead to unbalanced power dynamics in which the ethicist becomes a mere fig leaf to cover for the ethics of the project in question, leading in

the worst case to ethics washing. A code of conduct can help to clarify an ethicist's responsibility and ability to act as a partner within a research project and also how they should behave if recommendations are ignored or the development cannot be further supported.

Another issue arises when working as an external facilitator in development contexts, supporting development teams to identify and address ethical issues of the ongoing project. In this role, the ethicist has to navigate between the interests of (potentially) affected parties and the interests of the client (i.e. the development team). Working in these relations raises the question of what responsibility ethicists have towards which stakeholders. Being financially dependent on the client can create power dynamics, which runs the danger of creating biased results just to confirm the results expected. Here, a code of conduct can help to state the independence of the ethicist.

Last but not least, as evaluators, ethicists run the danger of following predefined guidelines or checkboxes to ensure compliance with pre-defined regulations or (inter)national laws. Furthermore, understanding ethics as mere evaluation frames the practice as yet another burden to be followed. The negative connotation overshadows the innovative potential and the original purpose of responsible innovation. A code of conduct allows to creation a professional self-understanding of the discipline that emphasizes the societal benefits of innovation processes instead of framing ethics as yet another activity that restrains science and technology.

Toward a code of conduct

We are as a community in a good position to do the work needed for writing a code of conduct for ourselves. Within technology ethics, there is enough knowledge and experience with creating codes of conduct, in part because some of us took part in the formulation of such codes for other fields, such as engineering, research, health care, etc. Also, there is ample research on codes of conduct to arrive at an effective one (e.g. (Carson, Baetz, and McGill 2008); (Vrielink and van Montfort 2010)).

One aspect of codes of conduct that may, however, need attention is finding the institutional structures for having an effective code. One of the lessons drawn in the abovementioned literature is that a code of conduct is better not used for a kind of 'meta ethics washing' of our field; a code of conduct for technology ethics practitioners should not become a document that can be waved for fending off external criticism. Rather a code should be supported by clear mechanics for measuring compliance of individual practitioners and, if needed, for intervening and sanctioning when internal or external stakeholders suspect practices that do not meet professional standards.

Again, writing down descriptions of the mechanisms for measuring and intervening is probably not the problem. More challenging is identifying the organization that has the right institutional position and structure to run these mechanisms. We have organizations that are involved in technology ethics and as such have the disciplinary authority to develop and 'own' a code of conduct for technology ethics practitioners. One can then think of the Society for Philosophy and Technology[7] and the Society for Applied Philosophy.[8] For endorsing the code an organization should however also have the institutional strength for doing assessments and acting on those assessments. The organization should mandate committees and have the legal basis for defining or receiving issues about

individual practitioners, conducting investigations, and intervening in the research, facilitation, or advising of those practitioners. Larger associations such as the American Philosophical Association[9] may have this institutional structure available, yet may also stand at more distance from technology ethics. Hence creating a code of conduct may include building up the needed institutional infrastructure in an organization.

Demarcation and further steps

In this paper, we argued for having a code of conduct for technology ethics practitioners. We discussed what types of technology ethics practitioners the code can be for, what roles the code can play for these practitioners, what controversies it should address, and what work may be involved in creating the (institutional) arrangements needed for the code.

A final issue may be one of demarcation. So far we have treated technology ethics as an emerging and relatively separate field. Yet when introducing this field we already mentioned adjacent fields such as ethics in academia and research ethics. More of such fields can be thought of, such as medical ethics, or consultancy proper. Therefore, when considering a code of conduct for technology ethics practitioners, some demarcation work is needed. Acknowledging adjacency may also help in making a useful choice for the institutional embedding of a code, such as associations for medical ethics and consultancy.

One demarcation seems more difficult to make and that is between technology ethics and research ethics. When it is accepted that ethics is to be integrated throughout the whole design and development process for technologies, then ethics integration starts with basic technology research, continues with research to prove feasibility, technology development, technology demonstration, system, and subsystem development, and ends at system tests, launch, and operation. Technology ethics practitioners may currently work regularly on later phases in this process. The earlier phases of basic technology research are however also relevant to technology ethics and can nowadays be reviewed in research ethics in specifically technological universities and institutes. Integration of ethics in the whole development process of technology may therefore imply that technology ethics and research ethics should not be demarcated, suggesting a future merger of their codes of conduct.

Technology ethics has become increasingly constructive by contributing to technology development. Writing a position paper on a code of conduct should in a similar vein be constructive. Hence, in addition to exploring the possibilities for such a code, we should follow up with actions toward creating the code. This position paper is a first step and a call for technology ethics practitioners to join this exploration. The next step is to bring together the main stakeholders ranging from the practitioners, the clients, and the representatives of associations who could endorse and host a code of conduct. An option for this is to meet at an international conference in the field, such as the biennial conference of the Society for Philosophy and Technology,[10] which we indeed should do.

Notes

1. E.g., at the Berlin Ethics Lab https://www.tu.berlin/en/philtech/institutions-and-services/berlin-ethics-lab, projects included the Cluster of Excellence Science of Intelligence

https://www.scienceofintelligence.de/, ROMI (Robotic support for routine tasks to strengthen cooperation in care facilities) https://www.romi-projekt.de/ and rokit (Research and expertise for public robotics) https://public-robots.de/.
2. E.g., the Horizon project TechEthos https://www.techethos.eu; and Van den Hoven, Vermaas, and van de Poel (2015).
3. E.g., the European Quantum Flagship project OpenSuperQplus100 https://opensuperqplus.eu.
4. Adopted from the definition of a code of practices in Anheier and List (2005, 57).
5. Carson, Baetz, and McGill (2008, 2).
6. See also recently in this journal (Elhadj et al. 2024; Herzog and Blank 2024).
7. https://www.spt.org
8. https://www.appliedphil.org
9. https://www.apaonline.org
10. https://www.spt.org/next-conference/

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Pieter Vermaas http://orcid.org/0000-0002-8175-9179

References

Anheier, H. K., and R. A. List. 2005. *A Dictionary of Civil Society, Philanthropy and the Non-Profit Sector*. First ed. New York: Routledge.

Balayn, A., M. Yurrita, J. Yang, and U. Gadiraju. 2023. ""☑ Fairness Toolkits, A Checkbox Culture?" On the Factors that Fragment Developer Practices in Handling Algorithmic Harms." In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 482–495. https://doi.org/10.1145/3600211.3604674.

Bietti, E. 2020. "From Ethics Washing to Ethics Bashing: A View on Tech Ethics From Within Moral Philosophy." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 210–219. https://doi.org/10.1145/3351095.3372860.

Buedo, P., and M. Waligora. 2022. "We Need to Eliminate Ethics-Washing." *Journal of Medical Ethics Blog*. https://blogs.bmj.com/medical-ethics/2022/06/16/we-need-to-eliminate-ethics-washing/.

Carson, A. S., M. Baetz, and S. McGill. 2008. *Codes of Conduct in the Private Sector: A Review of the Academic Literature from 1987 to 2007*. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=bb96edca909acd6ce2c1e813135e73aa42ec0005.

Elhadj, E., Z. Van Horenbeeck, E. Lievevrouw, and I. Van Hoyweghen. 2024. "Brokering Responsible Research and Innovation in in Silico Medicine." *Journal of Responsible Innovation* 11 (1): 2414484. https://doi.org/10.1080/23299460.2024.2414484.

Hansson, S. O. 2017. "The Ethics of Doing Ethics." *Science and Engineering Ethics* 23 (1): 105–120. https://doi.org/10.1007/s11948-016-9772-3.

Herzog, C., and S. Blank. 2024. "A Systemic Perspective on Bridging the Principles-to-Practice Gap in Creating Ethical Artificial Intelligence Solutions: A Critique of Dominant Narratives and Proposal for a Collaborative Way Forward." *Journal of Responsible Innovation* 11 (1): 2431350. https://doi.org/10.1080/23299460.2024.2431350.

Kiran, A. H., N. Oudshoorn, and P.-P. Verbeek. 2015. "Beyond Checklists: Toward an Ethical-Constructive Technology Assessment." *Journal of Responsible Innovation* 2 (1): 5–19. https://doi.org/10.1080/23299460.2014.992769.

Schiff, D., B. Rakova, A. Ayesh, A. Fanti, and M. Lennon. 2021. "Explaining the Principles to Practices Gap in AI." *IEEE Technology and Society Magazine* 40 (2): 81–94. https://doi.org/10.1109/MTS.2021.3056286.

Van den Hoven, J., P. E. Vermaas, and I. van de Poel, eds. 2015. *Handbook of Ethics, Values and Technological Design*. Dordrecht: Springer.

van Maanen, G. 2022. "AI Ethics, Ethics Washing, and the Need to Politicize Data Ethics." *Digital Society* 1 (2): 9. https://doi.org/10.1007/s44206-022-00013-3.

Vrielink, M. O., and C. van Montfort. 2010. "Codes of Ethics and Codes of Practice." In *International Encyclopedia of Civil Society*, edited by H. K. Anheier, and S. Toepler, 492–497. New York: Springer. https://doi.org/10.1007/978-0-387-93996-4_165.

Wagner, B. 2018. "Ethics as an Escape from Regulation. From "Ethics-Washing" to Ethics-Shopping?" In *BEING PROFILED:COGITAS ERGO SUM*, edited by E. Bayamlioglu, I. Baraliuc, L. A. W. Janssens, and M. Hildebrandt, 84–89. Amsterdam: Amsterdam University Press. https://doi.org/10.1515/9789048550180-016.