Journal of Ethics and Emerging Technologies

IEET

Book Review

An Introduction to Ethics in Robotics and AI (Five Years Later)

Philip Højme (Hoejme)[1,*]

[1] Center for Bioethics and Biolaw, Department of Ethics, Faculty of Philosophy, Warsaw University; p.hoejme@uw.edu.pl
* p.hoejme@uw.edu.pl

Abstract: New developments in robotics and AI necessitate revisions to introductions. This review critically assesses *An Introduction to Ethics in Robotics and AI* by Bartneck et al. (2021). The review both criticises central shortcomings in the book and engages with specific parts that, in light of recent developments, ought to be revised. The review concludes that the book provides an apt introduction for laypeople, students, and scholars seeking a concise overview of the ethics of robotics and AI.

Keywords: AI; Robotics; Ethics; Neurorights; Healthcare; Brain-computer interfaces; Autonomy; Privacy; Mental integrity.

Citation: Højme, Philip. 2026. An Introduction to Ethics in Robotics and AI (Five Years Later). *Journal of Ethics and Emerging Technologies* 36: 1. https://doi.org/10.55613/jeet.v36i1.229

Received: 9/04/2026
Accepted: 11/04/2026
Published: 19/04/2026

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

1. Introduction

An Introduction to Ethics in Robotics and AI (Bartneck et al. 2021), part of the *SpringerBriefs in Ethics* series, concisely describes central ethical issues about robotics and AI at the time of publication (2021). The review takes into account the time lapse between the book's publication date and the review's date (2026). It is, nevertheless, important to *emphasise that*, because of the accelerated development in robotics and AI, *the book would benefit from a revision*. However, despite rapid development in the relevant fields, the book nevertheless provides a solid introduction aimed at familiarising the general public, university students, policy-makers, and professionals with the broader topics of Ethics in Robotics and AI.

The book is written by Christoph Bartneck (University of Canterbury, NZ), Christoph Lütge (Technical University of Munich, DE), Alan R. Wagner (Penn. State, US), and Sean Welsh (PhD in Philosophy from the UC, NZ), who, together, have expertise spanning Human-computer interactions, design, engineering, ethics, law, philosophy, and social issues.

The purpose of the book, to provide a general introduction to ethics in robotics and AI, is clearly enhanced by its structure. The book contains 12 chapters, each covering a specific area. Following a short introduction to the authors and the book's structure, chapters two and three are especially important because they provide concise introductions to AI and Ethics, thereby giving novice readers the background knowledge needed to engage fully with the following chapters. The remaining chapters, four to eleven, contain text that, to varying degrees, engage with theoretical and practical issues related to robotics and AI. Some, e.g., chapters four, five, and seven, specifically deal with trust, fairness, responsibility, and psychology in relation to AI systems. While later chapters, e.g., ten and eleven, concern themselves with self-driving cars, the design of safety systems, and autonomous weapons systems. Thus, while some chapters lean towards theory, e.g., chapters six and eight on Risks in the Business of AI and Privacy Issues in AI, others, e.g., chapters nine, ten, and eleven, engage with specific areas and cases where ethical issues related to robotics and AI are already emerging. The twelfth and last chapter concludes the book with a brief

discussion of strategic considerations for current (2021) and future developments in the ethics of robotics and AI.

Hence, the authors live up to their promises, as the book concisely describes various cases and accounts of more theoretical ethical issues that must be addressed, or at the very least considered, by the public, policy-makers, business professionals, and others interested in these topics.

2. Omissions and gaps: on the debates on the ethics of restoration and enhancement

In this section, I want to highlight omissions and gaps that the book would have benefited from including.

Within the field of bioethics, discussions on restoration versus enhancement are nothing new. In fact, such discussions precede this book, and thus cast doubt on the authors' statement that "[n]o ethical dilemma ensues when the use of robotic replacement parts is for restorative purposes" (Bartneck et al. 2021, 71). On the one hand, the literature on prosthetic/robotic devices is flush with ethical concerns related to benefit and harm (Glannon, 2016) and dignity (Jotterand, 2010)—these are but two specific discussions in which ethical issues concerning robotic prostheses arise. On the other hand, the broader discussion on restoration and enhancement is laden with ethical concerns: "[r]obotic prostheses raise ethical and legal issues because they further problematise the distinction between therapy and enhancement" (Leenes et al., 2017). As such, the authors' statement that there are no ethical dilemmas related to restorative purposes seems odd, and this claim ought to have been substantiated. However, even with the background provided, the claim is essentially flawed and problematic, given that it contradicts common bioethical discussion on the ethics of robotics and prostheses.

The idea of separating restorative from enhancement, common in medical ethics (Levy, 2007), has nevertheless been problematised. Laura Colleton (2008) has criticised this division based, in part, on issues related to treatment—i.e., depending on our definition of medically necessary treatment (restoration of normal functions) and enhancement (augmentation beyond normal levels of function), specific treatments become available for healthcare (private or public) coverage, while others are not covered. Moreover, the idea of a 'normal' body has been criticised by feminist and disability studies (Karpin & Mykitiuk, 2008).

The book would have benefited from clearly addressing contemporary debates over restoration versus enhancement, as well as broader discussions of normal or species-typical bodies. While this might have been beyond the book's initial scope, the lack of engagement with, at the time, ongoing discussions means the book fails to account for the full scope of the scholarly discourse. Moreover, by claiming that restoration is ethically unproblematic, the book simply states a fact that seems more likely to demarcate the scope of a specific section than to be thoroughly researched.

3. Novel developments: suggestions for a new edition

Among the parts that would benefit from revision or expansion, the following seems most pertinent: 1) Chapter 7.2 *Persuasive AI*, 2) Chapter 9.1, 9.2, and 9.3 concerning robots and AI in connection to healthcare, 3) Chapter 10 *Autonomous Vehicles*, 4) Chapter 11 *Military Uses of AI*, and 5) Chapter 12 *Ethics in AI and Robotics* concerning labour. This list is in no way meant to be exhaustive, but rather limited to specific parts where later developments warrant special attention.

1) The book contains a short, one-page entry on the persuasive power of AI and robots that clearly outlines central issues related to autonomy diminishment if people are confronted with a super robotic salesperson that can utilise persuasive superpowers (Bartneck et al., 2021, 58). However, such issues have been central to broader discussions related to 'Neurorights' (Ienca & Andorno, 2017; Yuste et al., 2017) since 2017. Neurorights has been defined in various ways, most of which overlap in their concern with issues such as cognitive liberty, privacy, mental integrity, and psychological continuity (agency and identity). Robot/AI persuasion touches on all of these. As such, the brief discussion of persuasive AI would greatly benefit from being expanded in relation to these rights.

2) The discussion on healthcare, robotics, and AI only mentions brain-computer interfaces (BCI) in passing (Bartneck et al., 2021, 73). However, this topic has been researched by various authors (Lucivero & Tamburrini, 2008; Ienca & Haselager, 2016; Gilbert et al., 2019; Ligthart, 2020) who have examined ethical and technical issues related to BCI technologies and how these implants may impact our mental integrity and cognitive liberty. Thus, the book would benefit from expanding the BCI paragraph, potentially elevating it to a standalone section, including both previous (before 2021) and later research on BCI technologies.

3) Recent developments in AI have led to rapid developments in fleet learning, leading to machine learning and AI becoming the standard solution for the safety of self-driving vehicles. Thus, a revised version of the book should expand on this by highlighting how the industry has moved from treating machine learning and AI as part of safety protocols to treating them as central to the safety of self-driving vehicles, by, e.g., looking at the ISO PAS 8800 framework (Held et al., 2024).

4) Since 2021, the war in Ukraine has seen an unprecedented increase in the development of semi-autonomous weapons systems. As such, a revised version of this book ought to include references to such cases (Mukhina, 2026) as examples of the proliferation of autonomous weapons systems and the mechanisation of war (robotics and AI).

5) The references to AI's impact on labour could be expanded in light of the past five years' development. When the book was written in 2021, data on the impact of AI and robotics on the labour market were, presumably, more limited. Nowadays, more data is available, and a revised version of the book could be more decisive in its descriptions of the effects of AI and advanced robotics on the labour market, especially regarding algorithmic bias (e.g., in hiring) and management decisions.

4. Summary

Despite the clear need for revisions, given the novel developments in robotics and AI, the book serves its purpose of introducing a general audience to the topic of ethics in robotics and AI. As such, the book is recommended for laypeople, students, and scholars seeking a concise introduction to these topics. However, the book is not without problematic claims, and it would have benefited from a thorough discussion of the debate over the ethics of restoration and enhancement in robotics and AI, including these perspectives.

Funding: This research was supported by the National Science Centre, Poland, grant no. 2023/50/E/HS1/00486.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

References

(Bartneck et al., 2021) Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). *An Introduction to Ethics in Robotics and AI*. Springer International Publishing. https://doi.org/10.1007/978-3-030-51110-4

(Colleton, 2008) Colleton, L. (2008). The Elusive Line Between Enhancement and Therapy and Its Effects on Health Care in the U.S. *Journal of Evolution and Technology*, 18(1), 70–78.

(Gilbert et al., 2019) Gilbert, F., Cook, M., O'Brien, T., & Illes, J. (2019). Embodiment and Estrangement: Results from a First-in-Human 'Intelligent BCI' Trial. *Science and Engineering Ethics*, 25(1), 83–96. https://doi.org/10.1007/s11948-017-0001-5

(Glannon, 2016) Glannon, W. (2016). Ethical issues in neuroprosthetics. *Journal of Neural Engineering*, 13(2), 021002. https://doi.org/10.1088/1741-2560/13/2/021002

(Held et al., 2024) Held, S., Bossert, A., Lenzen, F., & Sutter, U. (2024). AI in Safety-critical Automotive Applications. *ATZ Worldwide*, 126(9), 38–41. https://doi.org/10.1007/s38311-024-1954-z

(Ienca & Andorno, 2017) Ienca, M., & Andorno, R. (2017). Towards new human rights in the age of neuroscience and neurotechnology. *Life Sciences, Society and Policy*, 13(1). https://doi.org/10.1186/s40504-017-0050-1

(Ienca & Haselager, 2016) Ienca, M., & Haselager, P. (2016). Hacking the brain: Brain–computer interfacing technology and the ethics of neurosecurity. *Ethics and Information Technology*, 18(2), 117–129. https://doi.org/10.1007/s10676-016-9398-9

(Jotterand, 2010) Jotterand, F. (2010). Human Dignity and Transhumanism: Do Anthro-Technological Devices Have Moral Status? *The American Journal of Bioethics*, 10(7), 45–52. https://doi.org/10.1080/15265161003728795

(Karpin & Mykitiuk, 2008) Karpin, I., & Mykitiuk, R. (2008). Going out on a Limb: Prosthetics, Normalcy and Disputing the Therapy/Enhancement Distinction. *Medical Law Review*, 16, 413–436.

(Leenes et al., 2017) Leenes, R., Palmerini, E., Koops, B.-J., Bertolini, A., Salvini, P., & Lucivero, F. (2017). Regulatory challenges of robotics: Some guidelines for addressing legal and ethical issues. *Law, Innovation and Technology*, 9(1), 1–44. https://doi.org/10.1080/17579961.2017.1304921

(Levy, 2007) Levy, N. (2007). *Neuroethics: Challenges for the 21st Century*. Cambridge University Press.

(Ligthart, 2020) Ligthart, S. (2020). Freedom of thought in Europe: Do advances in 'brain-reading' technology call for revision? *Journal of Law and the Biosciences*, 7(1), lsaa048. https://doi.org/10.1093/jlb/lsaa048

(Lucivero & Tamburrini, 2008) Lucivero, F., & Tamburrini, G. (2008). Ethical monitoring of brain-machine interfaces. *AI & SOCIETY*, 22(3), 449–460. https://doi.org/10.1007/s00146-007-0146-x

(Mukhina, 2026) Mukhina, O. (2026, January 28). For first time in history, Ukrainian combat AI-robot takes captive three Russian soldiers [*sic*]. *Euromaidan Press*. https://euromaidanpress.com/2026/01/28/ukrainian-combat-ai-robot-captures-three-russian-soldiers/

(Yuste et al., 2017) Yuste, R., Goering, S., Arcas, B. A. y, Bi, G., Carmena, J. M., Carter, A., Fins, J. J., Friesen, P., Gallant, J., Huggins, J. E., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., … Wolpaw, J. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*, 551(7679), 159–163. https://doi.org/10.1038/551159a