

Digital Humanism and Global Issues in Artificial Intelligence Ethics



Guglielmo Tamburrini

Abstract In the fight against pandemics and climate crisis, the zero hunger challenge, the preservation of international peace and stability, and the protection of democratic participation in political decision-making, AI has increasing – and often double-edged – roles to play in connection with ethical issues having a genuinely global dimension. The governance of AI ambivalence in these contexts looms large on both the AI ethics and digital humanism agendas.

1 Introduction

Global ethical issues concern humankind as a whole and each member of the human species irrespective of her or his position, functions, and origin. Prominent issues of this sort include the fight against pandemics and climate crisis, the zero hunger challenge, the preservation of international peace and stability, and the protection of democracy and citizen participation in political decision-making. What role is AI playing – with its increasingly pervasive technologies and systems – and will be likely to play in connection with these global ethical challenges?

The COVID-19 pandemic has raised distinctive challenges of human health and well-being protection across the planet, which are inextricably related to worldwide issues of economic resilience and protection of education, work, and social life participation rights. Artificial Intelligence (AI) has the potential to become a major technological tool to meet pandemic outbursts and the attending ethical issues. Indeed, infection spreading data and machine learning (ML) technologies pave the way to computational models for predicting diffusion patterns, identifying and assessing the effectiveness of pharmacological, social, and environmental measures, up to and including the monitoring of wildlife ecological niches, whose preservation is so important to restrain frequent contacts with wild animal species and related virus spillovers. Similarly, AI affords technological tools to optimize food

G. Tamburrini (✉)
Università di Napoli Federico II, Naples, Italy
e-mail: guglielmo.tamburrini@unina.it

production and distribution so as to fight famines and move toward the zero hunger goal in the UN sustainable development agenda.

Failures to use effective AI technologies to fight pandemics and world hunger may qualify as morally significant omissions. Along with these omissions, another source of moral fault may emerge from the ethically *ambivalent* roles that AI is actively assuming in the context of other global challenges. On the one hand, AI models may contribute to identify energy consumption patterns and corresponding climate warming mitigation measures. On the other hand, AI model training and related big data management produce a considerable carbon footprint. Similarly, AI military applications may improve Communications, Command, and Control (C3) networks and enhance both precision and effectiveness of weapons systems, leading to a reduction of military and civilian victims in warfare situations. And yet, the ongoing AI arms race may increase the tempo of conflicts beyond meaningful human control and lower the threshold to start conflicts, thereby threatening international peace and stability. Just as importantly, AI systems may help one in retrieving the diversified political information which is needed to exercise responsible democratic citizenship. However, in both authoritarian and democratic countries, AI systems have been already used to curtail freedom and participation in political decision-making.

As exemplary cases of AI playing ambivalent roles in global ethical issues, I will focus here on the climate crisis and the preservation of global peace and international stability. Universal human values and needs that are prized by digital humanism play a crucial role in the governance of such AI ambivalence.

2 Artificial Intelligence Ethics and the Climate Crisis

AI models are well-suited to identify and monitor energy consumption patterns, in addition to suggest policy measures for curbing carbon emissions in transportation, energy, and other production sectors characterized by high carbon footprints. AI potential contribution to climate warming mitigation actions is extensively illustrated by the Climate Change AI Group (Rolnick et al. 2019) and advocated in multiple current initiatives of AI research and industry (<https://aiforgood.itu.int>). AI is presented there as a new technological opportunity to promote both intergenerational and intragenerational justice and to enact human responsibilities toward other living entities. However, exactly the same ethical values and responsibilities impel AI communities to look closely into the backyard of their own carbon footprint. The more optimistic forecasts suggest that the carbon footprint of the entire digital technology sector, including AI, will remain stable between now and 2050 (Blair 2020). But even this optimistic outlook is no reason for inaction. Indeed, if other production sectors will reduce their carbon footprint in accordance with the Paris Agreement goals, the proportion of global carbon emissions taking their origin in the ICT sector will considerably increase over the same temporal interval.

Within the widely differentiated ICT sector, extensive discussion is under way about the energy consumption of some non-AI software, like blockchain and other cryptocurrency software, which are estimated to consume amounts of energy exceeding the energy need of countries like Ukraine or Sweden (<https://cbeci.org/cbeci/comparisons/>). In contrast with this, it is still unclear which fraction of the ICT sector energy consumption can be specifically attributed to AI in general or to machine learning or other prominent research and commercial subfields in particular. Available data are mostly anecdotal. It was estimated that GPT-2 and GPT-3 – successful natural language processing (NLP) models for written text production developed by ML techniques – were trained by means of huge amounts of textual data and gave rise to a carbon footprint comparable to that of five average cars throughout their lifecycle (Strubell et al. 2019). More systematic assessment efforts are clearly needed.

Considering the increasingly pervasive impact of AI technologies, the *White Paper on AI* released in 2020 by the European Commission recommends addressing the carbon footprint of AI systems across their lifecycle and supply chain: “Given the increasing importance of AI, the environmental impact of AI systems needs to be duly considered throughout their lifecycle and across the entire supply chain, e.g., as regards resource usage for the training of algorithms and the storage of data” (EU 2020, p. 2). However, one should carefully note that developing suitable metrics and models for estimating the AI carbon footprint at large is a challenging and elusive problem. To begin with, it is difficult to precisely circumscribe AI within the broader ICT sector. Moreover, a sufficiently realistic assessment requires one to consider wider interaction layers between AI technologies and society, including AI-induced changes in work, leisure, and consumption patterns. These wider interaction layers have proven difficult to encompass and measure in the case of various other technologies and systems.

Without belittling the importance and the difficulty of achieving a sufficiently realistic evaluation, what is already known about the lifecycle of both exemplary AI systems like GPT-2 and GPT-3 and the supply chain of big data for ML suffices to spur a set of interrelated policy questions. Should one set quantitative limits to energy consumption for AI model training? How are AI carbon quotas, if any, to be identified at national and international levels? How to distribute equitable shares of AI limited resources to business, research, and public administration? Who should be in charge of deciding which data for AI training to collect, preserve, and eventually get rid of for the sake of environmental protection? (Lucivero 2019). Only by addressing these issues of environmental justice and sustainability can AI be made fully compatible with the permanence on our planet of human life and the unique moral agency that comes with it, grounding human dignity and the attending responsibilities that our species has toward all living entities (Jonas 1979).

3 Ethics and the Artificial Intelligence Arms Race

The protection of both human life and dignity has been playing a crucial role in the ethical and legal debate about autonomous weapons systems (AWS), that is, weapons systems that are capable of selecting and attacking military objectives without requiring any human intervention after their activation. The wide spectrum of positions emerging in this debate has invariably acknowledged as a serious possibility the occurrence of AWS suppressing human lives in violation of International Humanitarian Law (IHL) (Amoroso and Tamburrini 2020). Indeed, AI perceptual systems, developed by machine learning and paving the way to more advanced AWS, were found by adversarial testing to incur into unexpected and counter-intuitive errors that human operators would easily detect and avoid. Notable in the AWS debate context is the case of a school bus taken for an ostrich (Szegedy et al. 2014). Since properly used school buses and their passengers are protected by IHL distinction and proportionality principles, the example naturally suggests the following question: Who will be held responsible for unexpected and difficult to predict AWS acts that one would regard as war crimes, had they been committed by a human being?

The use of AWS has been additionally claimed to entail a violation of human dignity (Amoroso and Tamburrini 2020, p. 5). Robert Sparrow aptly summarized this view, pointing out that the decision to take another person's life must be compatible with the acknowledgement of the personhood of those with whom we interact in warfare. Therefore, "when AWS decide to launch an attack, the relevant interpersonal relationship is missing," and the human dignity of the potential victims is not recognized. "Indeed, in some fundamental sense there is no one who decide whether the target of the attack should live or die" (Sparrow 2016, pp. 106–7).

These various concerns about IHL and human dignity respect have been upheld since 2013 by the international Campaign to Stop Killer Robots in advocacy for a ban on lethal AWS. The Campaign has also extensively warned that AWS may raise special threats to international peace. The latter is a fundamental precondition for the flourishing of human life that any sensible construal of humanism as a doctrine and movement – including digital humanism – is bound to recognize as a highly prized value. AWS threaten peace by making wars easier to wage on account of reduced numbers of involved soldiers, by laying conditions for unpredictable runaway interactions between AWS on the battlefield, and by accelerating the pace of war beyond human cognitive and sensory-motor abilities.

AI may bring about threats to international peace and stability in the new cyberwarfare domain too. Indeed, AI learning systems are expected to become increasingly central there, not only for their potential to expand cyberdefence toolsets but also to launch more efficient cyberattacks (Christen et al. 2020, p. 4). Cyberattacks aimed at nuclear weapons command and control networks, at the hacking of nuclear weapons activation systems, or at generating false nuclear attack warnings raise special concerns. Accordingly, the confluence of AI cyberweapons with nuclear weapons intensifies that distinctive threat to the permanence on our

planet of human life and moral agency that physicists and other scientists have been publicly denouncing at least since the Russell-Einstein Manifesto in 1955.

From the development of AWS to AI systems for discovering software vulnerabilities and waging cyberconflicts, an AI arms race is well under its way. The weaponization of AI should be internationally regulated and the AI arms race properly bridled. Digital humanism, with its analyses and policies inspired by universal ethical values and the protection of human dignity, has a central role to play in this formidable endeavor.

4 Concluding Remarks

The AI ethics agenda has been mostly concerned with ethical issues arising in specific AI application domains. Familiar cases are issues arising in connection with loans, careers and job hiring automatic decisions, insurance premium evaluation, or parole-granting tribunal judgments. Selectively affecting designated groups of stakeholders, these may aptly be called *local* ethical issues. Here, the focus has been placed instead on AI ethics issues that are *global*, insofar as they impact on humankind and all members of the human species as such. The climate crisis and the AI arms race have been used as exemplary cases to illustrate both the difference between local and global ethical issues and the need for a proper governance of AI ethically ambivalent roles. Last but not least, it has been argued that the ethical governance of this ambivalence makes crucial appeal to universal human values that any doctrine or movement deserving the name of digital humanism must endorse and support in the context of the digital revolution.

References

- Amoroso D., Tamburrini G. (2020) 'Autonomous Weapons Systems and Meaningful Human Control: Ethical and Legal Issues', *Current Robotics Reports* 1(7), pp. 187–194. <https://doi.org/10.1007/s43154-020-00024-3>
- Blair G. S., (2020) 'A tale of two cities: reflections on digital technology and the natural environment', *Patterns* 1(5). [https://www.cell.com/patterns/fulltext/S2666-3899\(20\)30088-X](https://www.cell.com/patterns/fulltext/S2666-3899(20)30088-X)
- Christen M., Gordijn B., Loi M. (eds) (2020). *The Ethics of Cybersecurity*. Cham: Springer. <https://link.springer.com/book/10.1007%2F978-3-030-29053-5>
- European Commission (2020). *White paper on AI. A European approach to excellence and trust*, Bruxelles, 19 February 2020. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en
- Jonas H. (1979). *Das Prinzip Verantwortung. Versuch einer Ethik für die technologische Zivilisation*. Frankfurt am Main: Insel-Verlag.
- Lucivero, F. (2019) 'Big data, big waste? A reflection on the environmental sustainability of big data initiatives', *Science and Engineering Ethics*, 26, pp. 1009–30. <https://doi.org/10.1007/s11948-019-00171-7>

- Rolnick D. et al. (2019) 'Tackling Climate Change with Machine Learning', *arxiv.org.1906.05433*. <https://arxiv.org/abs/1906.05433>
- Sparrow R. (2016) 'Robots and Respect: Assessing the Case Against Autonomous Weapon Systems', *Ethics & International Affairs* 30(1), pp. 93-116.
- Strubell E., Ganesh A., McCallum A. (2019) 'Energy and Policy Considerations for Deep Learning in NLP', *arxiv.org.1906.02243*. <https://arxiv.org/abs/1906.02243>
- Szegedy Ch. et al. (2014) 'Intriguing properties of neural networks,' *arxiv.org.1312.6199v4*. <https://arxiv.org/abs/1312.6199v4>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

