Artificial Intelligence Needs Environmental Ethics

Seth D. Baum and Andrea Owe
Global Catastrophic Risk Institute
http://gcrinstitute.org

Ethics, Policy & Environment, 26(1):139-143, 2023, DOI 10.1080/21550085.2022.2076538
This version 10 July 2023.

Abstract
The growing importance of artificial intelligence (AI) has prompted work on the ethical issues raised by AI technology. We argue that environmental ethicists can make important contributions to AI ethics, and that doing so would be a worthy activity for environmental ethicists. Environmental ethicists can highlight the environmental dimensions of AI, such as its energy footprint and its potential application for environmental protection. Additionally, environmental ethicists can assess the ethics of novel situations that could be enabled by AI, such as computer-based artificial life and artificial ecosystems.

Keywords: ethics; artificial intelligence; environmental ethics

Since around 2012, there has been a 'deep learning revolution' in artificial intelligence (AI) that has brought AI to the forefront of many sectors of human activity. As new AI technology has spread, the field of AI ethics has emerged alongside it to address the many important ethical issues raised by its design and application. Although much good work has been done, it often falls short in ways that are familiar to environmental ethicists. It is our view that work on AI would greatly benefit from the contributions of environmental ethicists, and furthermore that contributing to work on AI would be a worthy activity for environmental ethicists.

Broadly speaking, we see two different roles for environmental ethics to play in AI. One is to highlight the environmental dimensions of AI, such as the technology's significant energy footprint or its potential application for environmental protection. These sorts of issues already get some attention, but not as much attention as they deserve. The other is to apply environmental ethics perspectives to novel situations that could be made possible by AI, such as computer-based artificial life and artificial ecosystems. Environmental ethicists are well-placed to help answer many of the distinctive ethical questions raised by these situations.

As a simple example, consider the ethics of autonomous vehicles (i.e., self-driving cars, trucks, buses, etc.). Some commercially available vehicles already involve a degree of autonomy, such as the ability to autonomously stay centered within a highway lane or parallel park. The technology is improving, though it remains uncertain if or when fully autonomous vehicles will become widely available. Meanwhile, a sizable literature has arisen to address the ethics of 'trolley problems' related to how autonomous vehicles should act in rare cases when they have to decide whether to crash into X or Y for some X and Y (Goodall, 2019). Far less attention has been paid to the—arguably much more important—ethical issues raised by the environmental impacts of autonomous vehicles. Autonomous vehicles could bring environmental benefits, for example by optimizing trips for energy efficiency, as well as environmental harms, for example by increasing the use of private automobiles at the expense of public transit, walking, bicycling, etc. (Taiebat et al., 2018). An environmental ethics of autonomous vehicles could insist that robust attention be paid to their environmental implications.

A subtler example involves algorithmic bias. Deep learning and related techniques are essentially tools for identifying statistical patterns in large and complex data sets. As a consequence, they tend to

reproduce biases in the underlying data. Such 'algorithmic biases' have caused some substantial controversies, such as when racially biased AI was used as evidence in criminal court proceedings (Angwin et al., 2016). Although algorithmic bias is a major topic of study within AI ethics, the focus has almost exclusively been on biases within human populations. Human-specific biases are undoubtedly important, but they are not the only type of bias. The field of ecolinguistics shows that human language is often biased in its treatment of 'nature' and nonhumans, for example in the use of the word 'nature' to refer only to things unaffected by humans, as if humans were not also part of nature (Fill & Penz, 2018; Stibbe, 2021). More generally, as anthropocentrism is prevalent in many modern cultures, there is reason to suspect that there are important environmental and nonhuman algorithmic biases, that have thus far gotten zero attention. A study of environmental and nonhuman algorithmic biases would be a worthy research project to pursue.

Then there is the environmental impact of the AI itself. Deep learning requires AI to undergo an extensive training process. For example, an image recognition AI may need to process millions of images (such as in the 'ImageNet' database) to be trained to distinguish different types of images. This training process requires large amounts of computing power, which in turn requires large amounts of energy. One study finds that training a single AI model can produce greenhouse gas emissions comparable to the total lifetime emissions of one or several cars (Strubbel et al., 2019). Extending the cutting-edge AI systems largely depends on the use of larger and larger amounts of computing power. This raises the question of when AI development is worth the energy consumption it requires. This question can be answered using environmental ethics concepts such as the ecological footprint and the social cost of carbon.

AI is not uniformly bad for the environment. To the contrary, it can be quite good. AI is a general-purpose technology with a wide range of applications. As a general rule, anything that can be analyzed using complex statistical pattern recognition is a potential AI application. (Future AI technology may be even more widely applicable; more on that below.) Existing AI applications include the use of image analysis to monitor endangered species, geographic information systems analysis to optimize mosquito control to jointly minimize dengue fever and environmental disruption, and oceanographic analysis to optimize oceanic plastic waste cleanup.[1] A wide variety of AI applications to address climate change have been proposed (Rolnick et al., 2019). These and other examples demonstrate some of the positive environmental potential of AI, and further demonstrate that AI communities already include many people who care about the environment. In short, an environmental ethics of AI does not need to start from scratch.

Our own experience as environmental ethicists working on AI has been encouraging. We have found that environmental issues tend to be neglected in AI ethics. More precisely, our work has documented that work on AI ethics has been overwhelmingly anthropocentric (Owe & Baum, 2021). This work on AI ethics has played an important role in shifting the emphasis of AI development and use from the interests of private AI developers to the interests of the broader public. (As an aside, the role of profit-seeking corporations in AI development is another familiar theme for veterans of environmental issues.) The work nonetheless falls short by neglecting the interests of nonhumans.

Fortunately, we believe the human-centrism of AI ethics is mainly an oversight rather than a rejection of the importance of environmental issues or nonhumans. Indeed, other people involved in AI ethics have generally welcomed our contributions. We believe that contributions from other environmental ethicists would likewise be well received, especially if the contributions are presented as friendly, constructive contributions from a different disciplinary perspective rather than as hostile, harsh criticism. We believe it can be appropriate to be harshly critical, but only in response to those

1 These examples refer to Wild Me, the World Mosquito Program, and The Ocean Cleanup, all of which are projects of the Microsoft "AI for Earth" program, https://www.microsoft.com/en-us/ai/ai-for-earth.

who actively oppose the project of including consideration of environmental issues and nonhumans in AI ethics.

The distinction between environmental issues and nonhumans highlights another area in which environmental ethicists could contribute to work on AI. A major theme in environmental ethics is that 'the environment', or certain aspects of it, may be of intrinsic moral value, meaning that it may be valuable for its own sake and not just valuable for its relations to humans. Environmental ethicists have debated the intrinsic value of sentient nonhuman animals, living organisms, species, biodiversity, ecosystems, natural landscapes, and more (Rolston, 1988; Curry, 2011). Each of these entities may have artificial counterparts. Sentient AI and computer-based artificial life might or might not yet be possible, but both nonetheless have been the subject of serious scholarly attention (Aguilar et al., 2014; Oizumi et al., 2014). More and better analysis of the ethical implications is needed. The ethics of these artificial entities may resemble that of their 'natural' counterparts, though perhaps with some important differences. For example, the ethics of biodiversity sometimes considers the natural evolutionary processes that produce biological species (Rolston, 1988). 'Artificial biodiversity', if such a thing is possible, would emerge from different processes. Unpacking these sorts of nuances would be a worthy research direction.

The above pertains to how humans should value AIs and similar computer systems. A separate but also important question is on the ethical principles that humans should build into AI systems. The term 'machine ethics' is often used in this context to refer to the ethics held by the machines themselves. AI is a distinctive class of technology in that it can contain representations of moral values. There is a sense in which computers 'observe', 'think', 'decide', and 'act'. Indeed, 'sense/think/act' is a foundational paradigm for robotics. According to what ethical principles should AI or robotic systems base their actions? One influential view in machine ethics is that AI systems should be 'human-compatible' or 'aligned' with human values and/or interests (Russell, 2019). This human-centrism may be problematic, but alternatives may also be hard to implement. Work on this issue of aligning AI systems with the interests and values of also nonhumans would greatly benefit from collaborations between environmental ethicists and computer scientists.

Finally, there is the ethics of the future. Within the study of AI, there is great debate over whether to focus on existing AI technology ('near-term AI') and the issues it raises, or to focus instead on potential future AI technology ('long-term AI') (Baum, 2018). Near-term AI issues are more urgent, tractable, empirically robust, and politically salient. Long-term AI issues are more speculative and uncertain, but potentially of a much larger scale. In the extreme case, long-term AI could include runaway 'superintelligent' AI capable of outsmarting humanity, taking over the world, and, depending on the AI's design, remaking the world in either incredibly good or incredibly bad ways. In more moderate scenarios, humans could remain in control, but the impacts to the world could still significantly exceed the quite considerable impacts of current AI technology. This sort of present-vs.-future debate should feel familiar to environmental ethicists.

If significantly more advanced forms of AI can be built in the future, the environmental implications could be profound. The word 'if' may be doing a lot of work here: there is no guarantee that AI technology will advance in any particular way. Nonetheless, if there is a nonzero chance that significantly more advanced forms of AI could be built, then it would be exceptionally significant. The environmental issues raised by advanced forms of AI are similar to those raised by other prospective technologies such as geoengineering, which, because of its potential to be either extremely beneficial or extremely harmful, raises questions about whether it should be built or even discussed (Biermann, 2021). Advanced forms of AI also raise the additional machine ethics question of what ethical principles should be built into them. This is essentially a question about what world—or even what universe—we want to create, given an extraordinary capacity to do so. We believe it is self-evident that

environmental ethics can and should strongly inform consideration of this question.

For all of these reasons and more, AI needs environmental ethics. To be clear, it will not be easy work. The intersection of moral philosophy and environmental studies is difficult in its own right. Mixing in some computer science only makes it more difficult. Yet it is precisely this sort of interdisciplinary, high-stakes, socially relevant research that environmental ethics excels at. Indeed, many issues in AI ethics may be relatively easy 'low-hanging fruit' for environmental ethicists to analyze. Regardless, the stakes involved in AI technology are sufficiently large that we believe it is work well worth pursuing, even if it may at times be difficult. The significance for the world and potentially even the universe is simply too large to ignore.

Acknowledgments

We thank Robert de Neufville for helpful feedback on an earlier draft. Any remaining errors are the authors' alone.

References

Aguilar, W., Santamaria-Bonfil, G., Froese, T., & Gershenson, C. (2014). The past, present, and future of artificial life. *Frontiers in Robots and AI*. Advance online publication. https://doi.org/10.3389/frobt.2014.00008

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *ProPublica*, 23 May. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Baum, S. D. (2018). Reconciliation between factions focused on near-term and long-term artificial intelligence. *AI & Society*, *33*(4), 565-572.

Biermann, F. (2021). It is dangerous to normalize solar geoengineering research. *Nature*, *595*(7865), 30.

Curry, P. (2011). *Ecological ethics: An introduction*. (2nd ed.). Polity Press.

Fill, A., & Penz, H. (eds.). (2018). *The Routledge handbook of ecolinguistics*. Routledge.

Goodall, N. (2019). More than trolleys: Plausible, ethically ambiguous scenarios likely to be encountered by automated vehicles. *Transfers, 9*(2), 45-58.

Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Computational Biology, 10*(5), e1003588. https://doi.org/10.1371/journal.pcbi.1003588

Owe, A., & Baum, S. D. (2021). Moral consideration of nonhumans in the ethics of artificial intelligence. *AI & Ethics*, *1*(4), 517-528.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., & Luccioni, A. (2019). Tackling climate change with machine learning. https://arxiv.org/abs/1906.05433

Rolston III, H. (1988). *Environmental ethics: Duties to and values in the natural world*. Temple University Press.

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Penguin.

Stibbe, A. (2021). *Ecolinguistics: Language, ecology and the stories we live by*. (2nd ed.). Routledge.

Strubell, E., Ganesh A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 3645–3650. Florence, Italy.

Taiebat, M., Brown, A. L., Safford, H. R., Qu, S., & Xu, M. (2018). A review on energy, environmental, and sustainability implications of connected and automated vehicles. *Environmental Science & Technology*, *52*(20), 11449-11465.