



AI for climate: freedom, justice, and other ethical and political challenges

Mark Coeckelbergh¹

Received: 27 August 2020 / Accepted: 31 August 2020 / Published online: 6 October 2020
© Springer Nature Switzerland AG 2020

Abstract

Artificial intelligence can and should help to build a greener, more sustainable world and to deal with climate change. But these opportunities also raise ethical and political issues that need to be addressed if this project is to be successful. For example, the use of AI and the required data centers may involve high energy consumption, vulnerability to climate change and impact of climate measures differ across the globe and raise issues of justice, and when dealing with climate change in a way that influences or governs human behavior, there are trade-offs between effective measures that mitigate climate change and respecting human freedom. AI may also contribute to increasing humanity's hyper agency in relation to the planet, thus adding to what is known as the problem of the "Anthropocene". This article outlines and discusses these issues, with a focus on problems concerning freedom and justice at a global level, and calls for responsible use of AI for climate in the light of these challenges.

Keywords Artificial intelligence · Environment · Climate change · Ethics · Freedom · Nudging · Global justice · Anthropocene

1 Introduction

Today, climate change is widely regarded as one of the biggest and most threatening global challenges facing humanity—if not the most threatening. Global temperatures keep rising, there are more extreme weather events, and many species face extinction. Many believe that a response is urgently required or have already started to implement measures. Technologies such as artificial intelligence (AI) can be part of such measures: they can help to mitigate climate change and, more generally, help us to deal with a wide range of environmental issues.

There are various ways in which AI, especially but not exclusively in the form of machine learning applications, can be used for dealing with climate change. For example, AI can help to gather and process data on temperature change and carbon emissions, predict weather events and climate, show the effects of extreme weather, improve predictions of

how much energy we need and manage energy consumption (e.g., by means of smart grids), process data on endangered species, transform transportation in a way that leads to less carbon emissions and more efficient energy management and routing (car traffic, shipping, etc.), track deforestation and carbon emissions by industry, monitor ocean ecosystems, predict droughts and enable precision agriculture, contribute to smart recycling, assist carbon capture and geoengineering, and nudge consumers to behave in more climate friendly ways and create more awareness about the environmental and climate impact of their behavior (for an overview of what machine learning can do see for example [1]). Governments are interested in these applications, but also tech companies such as Microsoft, Amazon, and Google have started to invest in programs that develop AI applications to fight climate change.¹

Yet, like all technologies, AI also creates new problems. Next to technical challenges, the use of AI in general raises ethical issues, for example, threats to privacy and data protection, responsibility attribution, explainability, fairness. Ethical principles for AI such as beneficence, justice, and

¹ All major tech companies have sustainability and climate plans. However, they have also been criticized by environmentalist organizations such as Greenpeace for pursuing AI applications for oil and gas companies.

Invited contribution to the inaugural issue of "AI and Ethics".

✉ Mark Coeckelbergh
mark.coeckelbergh@univie.ac.at

¹ University of Vienna, Wien, Austria

explicability have been proposed and discussed by academics (e.g., [2–4]), policy advice bodies such as the High-Level Expert Group on AI set up by the European Commission [5] and professional societies such as IEEE, which took a global initiative on ethics of autonomous and intelligent systems.² These ethical issues need to be dealt with in all uses of AI, including the use of AI to improve the climate situation. For example, many reports rightly emphasize that it is important that humans can (still) take responsibility for automated systems, and there is currently a lot of attention to how machine learning may lead to, or increase, biased outcomes for specific individuals and groups—an issue that is important in ethics of machines in general (see for example [6]).

However, even if these issues were addressed, it is also important to recognize and address the issue that AI itself, next to creating the mentioned opportunities, can be problematic with regard to its impact on the environment and on climate change. While the environment is sometimes mentioned in policy documents and academic papers, it has received relatively little attention in ethics of artificial intelligence, and much more needs to be said about this topic. Nevertheless, there are important issues. For example, computer data centers use a lot of electricity, which is often produced in a way that creates emissions that warm the planet. Computers may be produced in a way that creates more emissions. And AI technology may also be sold to oil and gas industry to help extract more fossil fuels. But not only industry or regulators carry responsibility: as long as consumers frequently buy new electronic gadgets and use cars that run on oil, for example, these markets and economies will continue to exist in their current form.

The ethically and politically responsible thing to do then is to call for and develop green and climate friendly (uses of) AI, which does not only render our existing technologies more efficient and have them developed within an effective (green) regulatory framework, but also changes the way we live and—ultimately—transforms our entire economy and society. However, attention to the ethical issues does not solve all problems for at least two reasons, which both concern the global, planetary level.

First, some of the mentioned climate beneficial uses of AI create additional political problems. The mentioned general ethical issues are anyway already partly political (consider justice and stereotyping, which is currently hotly debated), but when we consider the proposed solutions for climate change, there are additional political challenges. For example, when AI is used to change behavior, there are also trade-offs with freedom, and the question of justice is not just a matter of potentially disadvantaging specific individuals and

groups within a society, but also needs to be asked again given global and generational differences in vulnerability to climate change and in the impact of measures taken to mitigate climate change.

Second, the very idea of managing the planet with the help of science and technology—by means of geoengineering but also by all the measures proposed—can also be seen as problematic from the angle of the so-called problem of the “Anthropocene”. Is increasing our agency with regard to the planet necessarily a good thing? Is it part of the solution or part of the problem?

Let me first elaborate how and why the use of AI may *not* necessarily be good for climate change and then further describe and discuss the two mentioned additional problem areas.

2 Why AI can be bad for climate change: the (ir)responsible use of energy and materials

Machine learning AI needs a lot of data, and data processing and data storage use energy, which has impact on the environment and climate. Some types of computing uses more energy than others. For example, training of neural networks, used for machine learning (in particular so-called deep learning), consumes substantial amounts of energy as compared to, say, running a word processing program or even simply running the AI. For example, according to a much-cited study, the process of training a single natural language processing (NLP) model can lead to emissions of nearly 300,000 kg of carbon dioxide equivalent, which is five times the amount produced by an average car over its lifetime [7]. Not all models are as large as those used for (this kind of) NLP; energy use will be less with small models. And often models are not trained from scratch. Nevertheless, there will always be a significant environmental impact in terms of electricity use.

Companies such as Google, Amazon, and Microsoft have started to invest in renewable energy and AI is used to increase energy efficiency, which is good since many AI applications use cloud services from such companies, but it is questionable if such investments will be enough to offset the environmental and climate footprint of these technologies in general and at a global level. Moreover, the production of electronic devices requires not only a lot of energy but also intensive use and extraction of raw materials such as nickel and cobalt, next to the use of plastics for the devices and their packaging. These uses of energy and material resources tend to be invisible to individual users of AI; but that does not mean they do not happen. In the meantime, AI applications using deep learning increases, and so does the required computation and energy consumption. Unless this

² <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>.

problem is effectively dealt with, the use of AI for environmental and climate purposes remains a double-edged sword.

What could be done to render AI more environmentally responsible and climate friendly? For a start, it would be good to increase awareness about ethics among users of AI and data scientists, and support more research on methods to make the energy and materials ecosystem around AI more visible. More generally, those working with AI need to be made more aware of the consequences of their computing for the world outside the computer lab. This is slowly but surely starting now, but often efforts in this direction (e.g., in higher education) are not yet much institutionalized and often depend on individual initiative. In the next stage, when one tries to reduce energy use in practice (in academia but also in industry), there might be trade-offs between, for example, accuracy of the technology and energy use. There is no magic formula to deal with such “micro” ethical challenges. But it would already be a huge improvement if AI and data science practitioners recognize energy use as a relevant ethical value and—ideally—as one metric of success. At the very least, it should be a requirement for developers to track energy use. For example, Anthony et al. [8] have proposed that the energy and carbon footprint of model development and training in deep learning is reported alongside (the usual) performance metrics.

However, if AI was doing well according to all the ethical and environmental principles mentioned so far, there would be still two additional political problems (or rather clusters of problems).

3 Political problems concerning freedom: nudging or Green Leviathan?

The first cluster of political problems is created by at least two options we have when using AI for steering human behavior, both of which threaten the important ethical and political principle of human freedom.

The first option is to influence human behavior towards a more climate-friendly direction by means of AI. In particular, there is the option to “nudge” people to use less energy, produce less waste, not use a car, and so on. Nudging is not coercing people to change, but rather changing what Sunstein and Thaler [9] called their “choice architecture”. One pushes someone in a particular direction by changing the decision environment. For example, a supermarket could be designed in such a way that products that have a smaller carbon footprint (and are hence more climate friendly) are presented in prominent places. The idea is that freedom is not taken away from people, but at the same time, their proneness to make biased decisions is exploited—albeit in this case for good, environmental purposes. This is a form of paternalism, but according to

Thaler and Sunstein, it is a libertarian form of paternalism since—according to them—freedom is preserved.

Since behavioral change is key to improving our environmental and climate predicament as societies and as humanity, climate nudging may be perceived as an attractive option. If we all live in a more climate-friendly way, then this would significantly help to mitigate climate change and other environmental problems. Nudging provides crutches that support what we all should want. However, nudging threatens human freedom in the following way: while it preserves freedom of choice, it fails to respect human autonomy and rationality, since it subconsciously influences people’s choices and behavior. It claims to know what your better, rational self wants and should want, and then manipulates you, but does not involve your capacity for autonomous and rational decision making in the process; instead, it bypasses it.

The challenge for policy makers is then to decide whether giving up respect for autonomy and rationality of citizens and consumers is a price a society should be prepared to pay for the expected environmental and climate benefits, and if there is an alternative given the urgency of responding to climate change. We see that, today, companies and (to some extent) governments already use nudging, sometimes with the help of AI. For example, if you buy something on Amazon, the algorithm suggests some other products you may want to buy; this can function as a nudge. However, the fact that it is done does not render it ethically and politically right. One could argue that in a liberal democracy the covert manipulation of citizen’s choices and behavior has no place. But what if the alternative—trying to convince people about doing something about climate change by means of argument—fails? This remains a tricky problem, but one that has to be dealt with if AI for climate is to succeed.

The second option is to use AI to (help) govern humanity. Here, the idea is that if the current political situation continues, with a serious lack of climate governance at a planetary level, this is likely to end in planetary disaster. Current (mainly national) institutions seem inadequate to deal with climate change. This ‘institutional inadequacy’ [10] is problematic since without global governance there is so-called ‘free riding’. As long as there is no framework for effective collective decision making at the level of humanity and the planet, national states just do what they want, and this is often not helping climate change mitigation. Moreover, human intelligence alone may not be sufficient to deal with the situation, given the complexity of the problem. In order to fill these gaps, one could install a green government helped by AI or have AI take over. AI could then ensure that humanity is governed in such a way that climate goals are reached. It would regulate countries and individuals based on a policy created by the data it gathers and the data analysis it conducts.

In this option, freedom is seriously threatened once again, but in this case through straightforward coercion. The justification provided is that this is a necessary evil: it is bad, but the only way humanity (and other species) can survive, the only way to save the planet. The reasoning is similar to another, famous argument in political philosophy: that made by the seventeenth century political philosopher Thomas Hobbes. Hobbes [11] argued that a so-called “state of nature”, where there is no governance, necessarily leads to a chaotic and brutal, indeed violent condition. The only way to avoid this, then, is the necessary evil of what he called a “Leviathan”: a ruler who acts in an authoritarian way but preserves the peace. Similarly, one could call for a “Green Leviathan,” which would ensure that climate change is sufficiently governed and mitigated at a global level. But this would come at the price of a loss of freedom.

However, luckily the dilemma presented by this argument is false. It is not necessary to choose between absolute laissez-faire and a global authoritarian government. As some nation states show (for example in Europe), it is possible to put environmental and climate regulation in place which restricts freedom *to some extent* (for the purpose of improving the climate situation) but still leaves *enough* freedom. There is a middle way. However, defining what this “to some extent”, “enough”, and “middle” is, is of course a huge challenge in a democratic society, and even more so at the global level, given differences in political culture and values.

In practice, it is to be expected that governments and parliaments—if they will want to do something about the climate problem at all—will use a mix of nudging and regulation. But these are the underlying political–philosophical challenges regarding freedom they will have to deal with. And it remains problematic if only some nation states take climate action, whereas others continue to act as free riders. This is neither effective (with regard to addressing the climate problem) nor fair.

Fairness brings us to a second cluster of problems, which concern justice, and in particular justice as fairness. Justice was already mentioned as one of the ethical principles that has been proposed for AI, but this time we consider the global perspective and frame it as a political problem.

4 Political problems concerning global and intergenerational justice

Not everyone and every society and community on this planet is equally vulnerable to climate change: some are more vulnerable than others. For example, those living in areas vulnerable to flooding (e.g., a specific Pacific island population) or those living in regions with long droughts are more at risk, and elderly people suffer more from heat. There are also effects such as migration and economic

destabilization. When AI gives us more predictive knowledge, these vulnerabilities become more visible. Knowledge may also not be shared between nation states, which can be seen as unfair. Furthermore, one generation may suffer the consequences of climate change, whereas another (earlier one) caused it. If the former has to pay the costs of dealing with climate change, is this fair? Moreover, the impact of climate change measures may differ: the measures may be some harder than others (e.g., those whose economy is developing), and some may benefit some more than others (e.g., people who are already advantaged in many ways). This raises questions concerning justice as fairness at a global level, and they are relevant to the use of AI for climate. If we use AI for dealing with climate change, we better think twice about specific measures in terms of their consequences for global and intergenerational justice. As a COMEST report already put it a decade ago, but in a way that is unfortunately still very relevant to today’s challenges:

‘Failure to act could have catastrophic implications, but responses to climate change that are not thought through carefully, with ethical implications in mind, have the potential to devastate entire communities, create new paradigms of inequity and maldistribution, and render even more vulnerable those peoples who have already found themselves uprooted by other man-made political and ideological struggles.’ [12]

For AI for climate, this means that interventions should not only be ethical in the senses outlined above, but also need to be assessed on the basis of their effects on different communities, different generations, and different parts of the world in the light of political principles such as justice as fairness. For example, if we need not only more efficient technology (e.g., AI that is more energy efficient) but also a change in lifestyle (perhaps nudged by AI), then the question is who should change lifestyle in order to save whom? For example, it may be that in many countries in the West, climate change poses no immediate danger to young people now living in areas that will probably not be flooded. Yet one could argue that they have an ethical and political duty to take measures and change lifestyle to help those who live in areas vulnerable to climate change, the elderly, and the next generations. AI nudging could be part of these measures, and one could then discuss whether it is fair that they bear the costs in terms of freedom to enable the survival of others. Consider also geoengineering: if geoengineering benefits specific countries who employ these technologies (e.g., rich countries) but does not benefit, or even through their unintended effects harm, other countries and parts of the world, then this could also be seen as unfair.

Note also that when we consider the global perspective, both AI and climate change may be perceived as non-priorities by some people who have to deal with challenges such

as poverty, lack of clean water, or malaria. While climate change is certainly an urgent and global problem, it may be a matter of justice to negotiate the distribution of political attention and resources between those who can afford to think about, and invest in, AI for climate (often living in affluent countries), and those who have other urgent and certainly more immediate and visible concerns, needs, and interests which also need ethical and political attention. Without taking such a wider global political perspective and without addressing these matters of global justice, the discussion about AI for climate may well be perceived as a neo-colonial hobby.

5 Hyper agency in the Anthropocene

Finally, the use of AI for dealing with our climate predicament, while well meant, may well exacerbate the problem, or at least one dimension of the problem, in the following sense. It could be argued that one reason (perhaps a “deeper” reason) why we find ourselves in this climate change situation is related to our modern desire to control over everything and everyone by means of science and technology. This has resulted in a planetary condition that has been aptly called the “Anthropocene” [13]: human agency on earth—including especially technological agency—has increased to such an extent that humanity has become a geological force. Climate change can then be interpreted as the outcome of, perhaps the pinnacle of, this strong grip we have gained on the planet: it is the result and manifestation of this hyper-agency that has put nature and the planet under our full control and that has pervaded the earth and its ecosystems to such an extent that even the climate is now the result of our agency.

Now if this is indeed our predicament, then the use of AI to deal with climate change is pouring gasoline on the fire, since it is yet another expression of our technological will to power (to use a famous Nietzschean term), another effort to increase our grip on the earth. Instead of letting go, we use AI to turn the whole planet and all its beings into what we could call with Heidegger [14] a ‘standing reserve’ of data. It could be argued that such a ‘datafication’, not only of our worldview [15] but in the end also of the world itself, can only lead to more problems rather than less, since we keep our problematic mental habit of wanting to control and do not see that it might be sometimes good to let go, to *not* control. Instead of increasing and improving planetary management with the help of AI, it might be better to let go: it might be better for planet and people to loosen our grip on the earth and the climate, rather than frantically implementing all the technology we have, in a desperate attempt to save what we do not even fully understand. Paradoxically, then,

solving the problem of climate change might require that we put limits to our technological solutionism.

If this makes sense (and more discussion is needed), then this is a difficult road to go for tech people (and in fact for anyone living in modernity), who often have a solutionist attitude. This is at least partly due to differences in education. Whereas most humanities students are introduced to the sensitivities of, for example, tragic world views and (other) literature that suggests that we accept the limitations of the human condition even as we struggle with it or even rebel against it, scientists and technology researchers are trained to solve problems and persistently look for technological solutions; and both worlds seldom meet. But this need not be the end of the story. There is an increasing need for, and interest in, interdisciplinary ethics education that brings together both worlds, and we can create opportunities to familiarize people with each other’s perspectives and continue this discussion.

6 Conclusion: who should deal with its ethical and political challenges and how?

To conclude, what I called “AI for climate” is an excellent idea and is rightly applauded. We should use artificial intelligence for dealing with environmental and climate problems. But in this article, I have argued that this project can only be successful if it sufficiently and adequately deals with some important ethical and political issues: issues raised by AI in general, but also a number of specific issues that are highly relevant in the case of AI for climate and that have a global and planetary dimension: political problems concerning freedom and justice, and the challenge of using AI given the problem of (hyper)agency in the Anthropocene.

The next question (which already figured in the justice section) is then who should deal with these problems. If and to the extent that we all contribute(d) to climate change, we are all responsible for the future of our communities, societies, and the planet. To the extent that we can do something as individuals, for example by changing our lifestyle and by playing a (more) ethical role in the development or policy of AI, we should exercise that responsibility. However, some have more impact on climate than others and should carry more responsibility, and many of the indicated challenges are also political in nature and need to be addressed at a societal level, local and global. In a democratic context, that means that more public discussion about these issues (and about who should deal with them) is needed. As the many papers and documents on AI ethics show, we have lists of ethical and political principles and values. But these do not offer an a priori, “correct” answer to the difficult questions posed here. What ethical and political principles such as freedom and justice as fairness mean needs to be discussed in

particular contexts with the relevant people (stakeholders) and with regard to particular uses of AI.

That being said, those who develop and use AI have a special (in the sense of “specific”) responsibility. To make sure that AI leads to a greener and more climate friendly world is definitely also the responsibility of computer scientists, engineers, designers, managers, investors, and others involved in, managing, and promoting, AI and data science practices. Yet they can only exercise this responsibility if others (e.g., humanities people but also, for example, climate scientists, and society in general) support them: for example, if we transform technology education and training in radically interdisciplinary and transdisciplinary directions, and integrate ethical and political considerations into development, design, management and investment practices. Institutionally, we also need more permanent interfaces between, on the one hand, technology development (in industry, academia, etc.) and, on the other hand, political and societal discussions. In the light of fast technological developments such as those in AI and data science and the imminent local and global risks related to climate change, bridges between different worlds are more important and more urgent than ever. I hope that this new, much needed journal may contribute to building these bridges.

Compliance with ethical standards

Conflict of interest The author declares that he has no conflict of interest.

References

- Rolnick D, et al (2019) Tackling climate change with machine learning. <https://arxiv.org/abs/1906.05433>. Accessed 24 August 2020
- Floridi, L., et al.: AI4People—an ethical framework for a good ai society: opportunities, risks, principles, and recommendations. *Minds Mach* **28**, 689–707 (2018)
- Dignum, V.: Responsible artificial intelligence: how to develop and use ai in a responsible way. Springer, Cham (2019)
- Coeckelbergh, M.: AI ethics. MIT Press, Cambridge (2020)
- HLEG (High-Level Expert Group on Artificial Intelligence): Ethics guidelines for trustworthy AI. European Commission, Brussels (2019)
- Benjamin, R.: Race after technology. Polity, Cambridge (2019)
- Strubell E, Ganesh A, McCallum A (2019) Energy and policy considerations for deep learning in NLP. <https://arxiv.org/abs/1906.02243>. Accessed 24 August 2020
- Anthony LF, Kanding B, Selvan R (2020) Carbontracker: tracking and predicting the carbon footprint of training deep learning models. ICML workshop on “challenges in deploying and monitoring machine learning systems”. <https://arxiv.org/abs/1910.09700>. Accessed 24 August 2020
- Sunstein, C., Thaler, R.: Nudge—improving decisions about health, wealth, and happiness. Yale University Press, New Haven (2008)
- Gardiner, S.M.: A perfect moral storm: climate change, intergenerational ethics and the problem of moral corruption. *Environ Values* **15**, 397–413 (2006)
- Hobbes, T.: Leviathan. Oxford University Press, Oxford (1996)
- COMEST (World Commission on the Ethics of Scientific Knowledge and Technology): The ethical implications of global climate change. UNESCO, Paris (2010)
- Crutzen, P.J.: The “Anthropocene”. In: Ehlers, E., Krafft, T. (eds.) *Earth System science in the anthropocene*, pp. 13–18. Springer, Berlin (2006)
- Heidegger M (1977) The question concerning technology. In: Heidegger M, *The question concerning technology and other essays*. Lovitt W (trans). Harper & Row, New York
- Romele, A.: The datafication of the worldview. *AI Soc.* (2020). <https://doi.org/10.1007/s00146-020-00989-x>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.