Conceptual Foundations of Sustainability. A Sustainability Perspective on Artificial Intelligence: Extended Abstract

Larissa Bolte[1]

[1] *Institute for Science and Ethics, University of Bonn, Bonner Talweg 57, 53113 Bonn, Germany*

1. Abstract

This project investigates the conceptual foundations of 'sustainability' with the goal of assessing approaches to the ethics of Artificial Intelligence (AI) under the lens of that notion. In a previous paper, my co-authors, Tijs Vandemeulebroucke and Aimee van Wynsberghe, and I have suspected that in order to do justice to the normative demands of sustainability, the way in which we conceive of AI ethics, AI regulation, and ultimately AI as a technology has to be adjusted [1].

The study of 'Sustainable AI', i.e. of AI applications for sustainability and of the sustainability of AI itself [2], is currently in its infancy. First publications in the field point to significant environmental and social costs attached to the widespread adoption of AI technologies [3][4][5]. And yet, comprehensive frameworks for how these costs can be identified, assessed, and evaluated are largely missing.

At the same time, a particular approach to AI policy crystallises – there is a tendency in AI Ethics Guideline documents to focus on technical fixes for isolated artefacts, deterministically construed, that lie in the responsibility of expert technicians – an approach my colleagues and I have dubbed an "ethics of carefulness" [1]. For the most part, they do not consider broader societal transformations, the embeddedness of AI technologies in social and ecological structures, or the possibility of not developing a particular AI application at all.

By contrast, in the context of discourse on AI and sustainability, AI has increasingly been conceptualised not as an artefact, but rather as infrastructure. This includes consideration of the hardware infrastructure that is necessary to run AI algorithms [6][7], the fact that AI underpins and upholds infrastructures [8][9], and that the interplay of AI algorithms and their environment also constitutes an infrastructure in its own right [1][9]. Indeed, it has been argued that conceptualising and assessing the sustainability of AI requires considering AI artefacts not in isolation, but rather in their embeddedness in the broader ecological and socio-technical systems that surround, enable, and constitute them [1][10].

It seems that sustainability is simply not 'happening' at the artefact level. This may explain why social and ecological costs of AI, costs related to sustainability, are often described as "hidden"[2][8]: Through the lens of an ethics of carefulness, they are invisible.

My research contributes a thorough examination of the normative demands inherent to the sustainability perspective. These demands require modelling AI not as a particular artefact, but rather as a socio-technical system embedded in social, environmental, and economic structures. The normative demands of sustainability would thus require a different ontology for AI than the one that is predominantly found in AI policy documents.

FOIS 2023 Early Career Symposium (ECS), held at FOIS 2023, co-located with 9th Joint Ontology Workshops (JOWO 2023), 19-20 July, 2023, Sherbrooke, Québec, Canada

bolte@iwe.uni-bonn.de (L. Bolte)

0000-0003-4989-9558 (L. Bolte)

2. Motivation

The notion 'sustainability' has risen to extraordinary relevance in the face of the current climate crisis. Policy makers on all levels of government, businesses, research institutions, NGOs, and individuals alike have made 'sustainability' their guiding concern. This trend now also extends to current debates on digitalisation and, more specifically, Artificial Intelligence (AI).
First pointers to an environmental cost attached to AI applications have been provided by papers assessing the energy consumption and associated greenhouse gas emissions produced by training, tuning, and using AI systems. According to first estimates, the carbon emissions produced by training, and even more so by tuning, just one Natural Language Processing model may be considerable [3][11].

And yet, the environmental impact of AI extends beyond carbon emissions produced by the energy consumption of algorithms in development and use. For AI systems to run, they require instantiation in hardware and an industrial infrastructure to supply, maintain and replace this hardware, the environmental impact of which is of yet to be fully assessed. Given that AI is in the process of forming vital infrastructures that will shape our societies for decades to come [9], suitable frameworks to steer this development into a sustainable direction are now timelier than ever and direly needed. It is thus essential to understand what 'Sustainable AI' entails conceptually, i.e. what empirical data is needed to assess the sustainability of AI, what normative demands are supported by the data, and how we ought to conceptualise AI as a technology in light of sustainability concerns.

In the AI ethics context, only few researchers have made first attempts at adapting the sustainability notion for their purposes [1][2][5][6][7][9][10], and, ultimately, no comprehensive sustainability framework has to date been proposed for AI ethics. It stands to reason that a thorough examination of the sustainability concept within and outside of its employ in AI ethics discourse will yield insights that will prove fruitful for anyone working on Sustainable AI from a research or a regulation perspective.

3. Research Questions

3.1. Overarching Research Question

How can 'sustainability', construed as a theoretical lens, inform the way we conceive of 'Sustainable AI' as a new paradigm for AI ethics?

3.2. Phase 1 Research Questions

What are the central characteristics of 'sustainability'?
What are the normative demands implied by or inherent to 'sustainability'?
What ontology is required so 'sustainability' demands can be modelled?

3.3. Phase 2 Research Questions

What is the current state of AI ethics?
How does the sustainability framework developed in Research Phase 1 apply?
How can this framework inform or reform debates in AI ethics?

4. Objectives

4.1. Objectives of Phase 1 (Developing a Theoretical Sustainability Framework)

- Scoping out the conceptual space of 'sustainability':

First, sustainability conceptualisations in the literature will be identified, grouped, contrasted, and contextualised, with special focus on what aims, norms, goods, etc. sustainability theorists posit and on how the world must be construed from a sustainability perspective. A conceptual overview will be created.

- Developing a theoretical framework:
  A theoretical framework is an analytical structure that is used for interpretation and assessment. My project asks how AI ethics approaches can be interpreted and assessed from a sustainability perspective, i.e. whether and how AI ethics approaches are capable of answering to sustainability concerns. A sustainability framework will thus have to identify these concerns. Furthermore, normative concepts cannot be understood without the ontology on which they rely. The theoretical framework I develop will thus also have to map out what model of the world 'sustainability' requires, i.e. what aspects of a situation it picks out.

4.2. Objectives of Phase 2 (AI Ethics from a Sustainability Perspective):

- Identifying broader movements or paradigms in AI ethics:
  Before the sustainability framework developed in Research Phase 1 can be applied to the AI ethics context, the state of the latter must first be determined. Instead of giving a comprehensive overview of singular issues, broader movements in AI ethics will be identified.
- Revising our conceptualisation of AI as a technology from a sustainability perspective:
  Ordinary conceptions of AI as a technology will be assessed and revised in light of sustainability concerns.

5. Research Methodology

I work from a critical theory perspective and follow what Sally Haslanger calls a "revisionary project" [12]: Such a project approaches the definition of concepts from a pragmatic needs-based perspective. Revisionary projects amend concepts to turn them into effective tools to achieve legitimate purposes. They ask: What iteration of this concept would serve our cognitive or practical purposes best? Haslanger contrasts this kind of epistemic project with conceptual projects, which explore and articulate the nuances of ordinary concepts, and with descriptive projects, which study the extension of a concept to refine it.

In the context of the concepts of race and gender, in light of which Haslanger makes this distinction, a descriptive project could investigate whether there are social kinds that are tracked by our uses of race and gender vocabulary. A conceptual project would explore and articulate our notions of race and gender as they are used. A revisionary project, however, asks how we should use the concepts of race and gender if we want to achieve our goal of, for example, properly addressing racial and sexual injustices.

In the context of my project, I ask: What are our practical purposes when engaging in sustainability discourse? How ought we revise our conception of what AI is and how it interacts with the world from a sustainability perspective?

One objective in joining the ECS at FOIS 2023 has been to explore methodologies for investigating implicit ontological commitments in sustainability conceptions as well as for how to deduce suitable ontologies from normative demands.

6. Research Results to Date

A first paper with the outlook that sustainability may require a systems ontology [1].

Acknowledgements

Funding for this research was provided by the Alexander von Humboldt Foundation in the framework of the Alexander von Humboldt Professorship for Artificial Intelligence endowed by the Federal Ministry and Research to Prof. Dr. Aimee van Wynsberghe.

References

[1] Bolte, Larissa, Tijs Vandemeulebroucke and Aimee van Wynsberghe. "From an Ethics of Carefulness to an Ethics of Desirability: Going Beyond Current Ethics Approaches to Sustainable AI." Sustainability 14.8 (2022), 4472.
[2] van Wynsberghe, Aimee. "Sustainable AI: AI for sustainability and the sustainability of AI." AI and Ethics 1.3 (2021), 213-218.
[3] E. Strubell, A. Ganesh, A. McCallum, Energy and Policy Considerations for Deep Learning in NLP, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Florence, Italy, 2019, pp. 3645-3650. doi: 10.18653/v1/P19-1355.
[4] Brevini, Benedetta. "Black boxes, not green: Mythologizing artificial intelligence and omitting the environment." Big Data & Society 7.2 (2020), 2053951720935141.
[5] M. Coeckelbergh, Green leviathan or the poetics of political liberty: Navigating freedom in the age of climate change and artificial intelligence, Routledge, Abingdon and New York, 2021.
[6] K. Crawford, The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence, Yale University Press, New Haven, 2021.
[7] M. M. D. Bolger, A. Tofighi-Niaki, L. Seelman, 'Green Mining' is A Myth: The Case for Cutting EU Resource Consumption, joint report by the European Environmental Bureau and Friends of the Earth Europe, 2021. URL: https://eeb.org/library/green-mining-is-a-myth/.
[8] P. Dauvergne, AI in the Wild: Sustainability in the Age of Artificial Intelligence, The MIT Press, Cambridge, MA, 2020.
[9] Robbins, Scott and Aimee van Wynsberghe. "Our New Artificial Intelligence Infrastructure: Becoming Locked into an Unsustainable Future." Sustainability 14.8 (2022), 4829.
[10] Sætra, Henrik Skaug. "AI in context and the sustainable development goals: Factoring in the unsustainability of the sociotechnical system. " Sustainability 13.4 (2021), 1738.
[11] Lacoste, Alexandre, Alexandra Luccioni, Victor Schmidt and Thomas Dandres. "Quantifying the Carbon Emissions of Machine Learning." Arxiv preprint (2019). URL: https://arxiv.org/abs/1910.09700.
[12] Haslanger, Sally. "Gender and Race: (What) Are They? (What) Do We Want Them To Be? ." Noûs 34.1 (2000), 31-55.