



The case for global governance of AI: arguments, counter-arguments, and challenges ahead

Mark Coeckelbergh¹

Received: 16 November 2023 / Accepted: 10 April 2024
© The Author(s) 2024

Abstract

It is increasingly recognized that as artificial intelligence becomes more powerful and pervasive in society and creates risks and ethical issues that cross borders, a global approach is needed for the governance of these risks. But why, exactly, do we need this and what does that mean? In this Open Forum paper, author argues for global governance of AI for moral reasons but also outlines the governance challenges that this project raises.

Keywords Artificial intelligence · Global governance · Hobbes · Democracy · AI ethics

Recently there have been more calls for a global approach to the governance of AI across international organizations, industry, and academia. The UN's Secretary-General and his Envoy on Technology, for example, have called for globally coordinated AI governance as 'the only way to harness AI for humanity while addressing its risks and uncertainties'.¹ Earlier a Resolution adopted by the UN's General Assembly called for improving digital cooperation and deliberation using the UN as a platform for stakeholders,² thus preparing work on global governance. In September, the G20 leaders called in New Delhi for global governance for AI to harness AI for 'Good and for All'.³ OpenAI CEO Sam Altman called for coordinated international regulation of generative AI.⁴

And while still relatively rare, several academics have discussed how to achieve global governance of AI, often calling for new policies and new institutions (Erman and Furendal 2022; Dafoe 2018) and recognizing existing and emerging initiatives and regimes (Schmitt 2022; Butcher and Beridze 2019; Veale et al. 2023), also from non-governmental and non-profit directions. For example, next to the AI for Good summits⁵ that have discussed how AI can contribute to solving global, the Institute for Electrical and Electronics Engineers (IEEE) has its Global Initiative on Ethics of Autonomous and Intelligent Systems⁶ and in May 2021, the International Congress for the Governance of AI (ICGAI) held its first conference in Prague.⁷

¹ <https://www.un.org/techenvoy/ai-advisory-body>.

² <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N20/248/80/PDF/N2024880.pdf?OpenElement>.

³ <https://economictimes.indiatimes.com/tech/technology/g20-leaders-call-for-global-governance-for-ai-inclusive-digital-public-infra-for-service-delivery/articleshow/103537219.cms?from=mdr>.

⁴ <https://edition.cnn.com/2023/06/09/tech/korea-altman-chatgpt-ai-regulation-intl-hnk/index.html>.

⁵ <https://aiforgood.itu.int/>.

⁶ <https://standards.ieee.org/industry-connections/ec/autonomous-systems/>.

⁷ <https://www.icgai.org/>.

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

¹ Department of Philosophy, University of Vienna, Vienna, Austria

But why, exactly, is global governance needed, and what form can and should it take?

The main argument for the global governance of AI, which is also applicable to digital technologies in general, is essentially a moral one: as AI technologies become increasingly powerful and influential, we have the moral responsibility to ensure that it benefits humanity as a whole and that we deal with the global risks and the ethical and societal issues that arise from the technology, including privacy issues, security and military uses, bias and fairness, responsibility attribution, transparency, job displacement, safety, manipulation, and AI's environmental impact. Since the effects of AI cross borders, so the argument continues, global cooperation and global governance are the only means to fully and effectively exercise that moral responsibility and ensure responsible innovation and use of technology to increase the well-being for all and preserve peace; national regulation is not sufficient.

Some might add that the alternative to global governance is a race to the bottom: a kind of Hobbesian situation in which nations engage in a competitive race without heeding ethical standards, safety, and accountability, resulting in widespread injustice and inequality, displacement, security problems, power concentration, and perhaps even totalitarianism. Just as Hobbes thought that individuals left to themselves and not ruled by a state authority would render the life of individuals nasty, brutish, and short, one could argue that nation states left without global governance would result in a global disastrous situation where only some nations and their citizens benefit from the technology and others suffer. A global authority that reigns in the power of the individual nation states could solve this situation. A similar Hobbesian argument can and has been made regarding the climate crisis and other global challenges (Saetra 2022).

The Hobbesian for of the global governance of AI argument is not absolutely necessary, at least not in that form. Without world government, one could argue, the situation might not be as bleak as sketched here. There is already regulation at national and even supranational level. The EU, for example, will implement its AI Act, Biden recently issued an Executive Order to create A.I. safeguards,⁸ and China has published rules for generative AI.⁹ However, while this objection defuses the specific Hobbesian view, it does not undermine the general moral argument for global governance of AI: with national regulation in place in some countries, the world might get less nasty for some (e.g., for EU citizens), but such islands of regulation do not benefit

those who do not have the luck to live in these parts of the world. In other words, even without a race to the bottom everywhere and for everyone, the general argument still holds. For sake of justice, equality, and inclusion, we need a global governance framework, regardless of national regulation.

Sometimes the argument is made that AI will accelerate and that we need global governance given the risks of AGI (Artificial General Artificial Intelligence)—intelligence comparable to human intelligence—or superintelligence. It is argued that AGI might be in charge of global governance or may lead to (other) global existential risks. Sam Altman and Geoffrey Hinton, for instance, hold this view.¹⁰ Mitigating such risks, including risk of extinction from AI, is then a reason for global governance. While neither the acceleration thesis nor this view concerning the existential risks of AGI are shared by everyone in the scientific community, they have received increasing attention and are currently influencing AI policy—not only in the US but also in the EU, for example. I am very concerned about this development, if only since it contributes to increased power of people like Altman: they do not only create the problem but also claim to sell the solution, which gives them a unique undemocratic position of power. However, regardless of one's view on these matters, it is important to see that the world governance of AI argument does not depend on it. Just as a specific Hobbesian version is not necessary, a specific AGI version of the argument is also not necessary for it to work. Even without the supposed risks that might be created by AGI (if such a thing would ever exist), there are sufficient risks left and there is sufficient moral reason to mitigate them. Not believing in the possibility of AGI or in the acceleration thesis is not an excuse to reject global governance of AI.

A more challenging range of counter-arguments, however, has to do with the precise form global governance of AI can and should take. These counter-arguments point to important challenges for those who support this project and wish to implement it, and deserve careful consideration.

A first objection is that global governance is undemocratic. Here the assumption is that global governance means establishing a world government and that a world government is necessarily undemocratic. But these assumptions do not hold. Global governance can in principle be organized in a (more) democratic way, for instance, more democratic than currently the UN works, and there is no obvious reason why world governance should be organized along the lines of the nation state (or any particular nation state for that matter). If we can find a way to do this differently but still establishing sufficient authority then let us do that. In the history of politics and political theory, it has always been a

⁸ <https://www.nytimes.com/2023/10/30/us/politics/biden-ai-regulation.html>.

⁹ <https://edition.cnn.com/2023/07/14/tech/china-ai-regulation-intl-hnk/index.html>.

¹⁰ <https://www.theguardian.com/technology/2023/jun/07/what-should-the-limits-be-the-father-of-chatgpt-on-whether-ai-will-save-humanity-or-destroy-it>.

challenge to combine legitimacy and authority; this is not different in this case. Supporters of global governance of AI, therefore, can (and do) argue that they want a multi-stakeholder approach and want inclusivity and participation not only in terms of AI ethics but also when it comes to the global governance process. For example, the UN has recently established a multistakeholder advisory body on AI.¹¹ While this is arguably not democratic enough since it is composed of a rather selective membership, there is a growing awareness of the need for inclusivity and democratisation. Moreover, global agencies and (other) authorities are just one form global governance can take; there are also councils, international agreements, and other instruments of global governance. That being said, how to organize global governance remains a challenge and requires much more research and innovation efforts. Unfortunately, usually the degree and pace of institutional and political innovation does not match the speed of technological development. This needs to change. Institutions needed to be created that can respond faster to technological developments.

Another objection is that global governance of AI is unrealistic and too idealistic: that nation states are not, and will not be, willing to give up national sovereignty and delegate power to a global governance entity or framework, and that even if they would do so, it would be difficult to enforce anything since they would anyway do what they want. This objection can have two faces: a normative and a descriptive one. If the point is that we should not delegate this to supranational governance then one can reply with the moral imperative that we should do something about the risks and ethical problems; in other words, one can reiterate the main argument. If the point is that, as a matter-of-fact, nations are not and would not be willing to do this; one could point to existing global governance in other technological areas such as aviation and nuclear technology, and point to current and emerging initiatives that get the support of nation states. For example, those who argue for global regulation of AI often refer to the current nuclear governance model. Altman has used the analogy and UN Secretary Antonio Guterres has proposed the establishment of an international AI agency akin to the International Atomic Energy Agency.¹² While there are good reasons to be sceptical about the comparison between AI and nuclear weapons (Does AI pose existential risk similar to nuclear weapons, if it poses an existential risk at all? Does this distract us from real and known risks? And are nuclear weapons not easier to control given that they need specific resources?¹³), the example shows that it is not

only desirable but also possible to reach agreements about global regulation of technology. The UN's history when it comes to nuclear, aviation, and indeed climate change (Guterres also referred to the IPCC) shows that it is perfectly possible to come to new rules, treaties, and agencies at a global level in response to global threats.

A third potential weakness of the argument concerns, surprisingly perhaps, its moral component. The argument seems to assume that we all agree on AI ethics. But, so this objection goes, apart from nations having different interests (a point that is somewhat covered in the previous paragraph), they might also have different values. Given cultural diversity across the world, so it is argued, it is unlikely that nations might agree on a global governance framework. In response, one may point again to the fact that this has so far not been a barrier for international cooperation and global governance. Consider for instance human rights frameworks and their supranational institutions at UN and EU level, which despite being subject to decades of philosophical criticism that stresses difference and diversity, have been at least partly successful as a form of global governance by focusing on what we have in common as humans. And currently there seems consensus rather than divergence within the AI ethics community. Even if there is valid criticism that points to the danger of neo-colonialism and hegemony, ethical frameworks in this area look surprisingly similar and seem to have found some kind of pool of shared values. Consider for example UNESCO's Recommendation on the Ethics of Artificial Intelligence, which lists a number of such values.¹⁴ Moreover, from a philosophical point of view, it can be argued—as is done in the case of human rights for example—that while it is important to respect diversity and difference, humans also share a lot of needs, interests, and values, regardless their differences in terms of citizenship, culture, and identity. In other words, it is both possible and desirable to establish a global ethics, including a global AI ethics. Yet the objection does help to create sensitivity and awareness of the importance of respect for diversity and in this context must be seen as a call for creating global governance of AI in a global-inclusive way—for example, in a way that includes the Global South—and in a way that avoids the instalment of (another?) unjust and hegemonic regime. Global governance of AI can only succeed if it has broad global support across cultures and continents and takes into account all these values and interests.

Finally, there might be the worry that global governance of AI might hinder technological innovation. For example, in the process towards the E.U.'s AI Act, OpenAI and other

¹¹ <https://www.un.org/techenvoy/content/artificial-intelligence>.

¹² <https://www.reuters.com/technology/un-chief-backs-idea-global-ai-watchdog-like-nuclear-agency-2023-06-12/>.

¹³ See, for example, this Chatham House discussion: <https://www.chathamhouse.org/2023/06/nuclear-governance-model-wont-work-ai>.

¹⁴ <https://www.unesco.org/en/articles/unescos-recommendation-ethics-artificial-intelligence-key-facts>.

big tech companies have expressed concerns about this¹⁵; similar concerns exist concerning global governance of AI. But this is a familiar discussion also at the national level, and is not as such a good objection to global governance. What I currently see is that the tech industry itself also calls for regulation of AI, both at national level and at global level. The argument, I guess, is that innovation can only succeed if there is a regulative framework that brings more certainty and stability in this turbulent policy area, and that makes sure that the technology can be used and developed in a safe and ethical way. It is in the long-term interest of innovation and business that there is a robust and integrated global governance framework. The extent and nature of that framework may be under discussion—as it should be—and that discussion may well have to include this concern about protecting innovation, but this can hardly be an argument against a global approach. At most, it signals that there are of course power interests at play here, also at the global level. Big tech companies risk to monopolize both the development and the regulation of AI, at least those AI systems that are currently most successful and pervasive. The global governance of AI project questions this monopoly and rightly asks these companies to share the responsibility for better AI and a better world with global frameworks and global institutions that represent and protect citizens and their communities and cultures. How they can and should do this is a huge challenge, but this problem should not justify halting efforts towards more global governance of AI.

In conclusion, here is a good argument for global governance of AI, based on moral reasons and aimed at avoiding a situation in which only some citizens and countries benefit from AI whereas others have to deal with most of the risks and ethical issues. Objections that the global governance of AI project would necessarily be undemocratic, unrealistic, not respecting diversity, and hindering innovation, can be countered. Nevertheless, these objections point to challenging issues that the UN and other actors in this global policy arena will have to deal with in the coming years when trying to build this global governance framework. More research in this area is urgently required to support these efforts.

Funding Open access funding provided by University of Vienna.

Declarations

Conflict of interest There are no conflicting interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Butcher J, Beridze I (2019) What is the state of artificial intelligence governance globally. *RUSI J* 164(5–6):88–96
- Dafae A (2018) AI governance: a research agenda. Paper available at <http://www.fhi.ox.ac.uk/wp-content/uploads/GovAI-Agenda.pdf>. Accessed 9 Apr 2024
- Erman E, Furendal M (2022) The global governance of artificial intelligence: some normative concerns. *Moral Philos Politics* 9(2):267–291
- Saetra HS (2022) A hobbesian argument for world government. *Philosophies* 7(3):66
- Schmitt L (2022) Mapping global AI governance: a nascent regime in a fragmented landscape. *AI Ethics* 2:303–314
- Veale M, Matus K, Gorwa R (2023) AI and global governance: modalities, rationales, tensions. *Annual Review of Law and Social Science* 19. Retrieved from <https://discovery.ucl.ac.uk/id/eprint/10171121/1/Veale%20Matus%20Gorwa%202023.pdf>. Accessed 9 Apr 2024

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

¹⁵ <https://www.irishtimes.com/business/innovation/2023/07/13/chris-horn-could-regulation-stifle-ai-innovation-in-europe/>.