Human Technology Institute

UTS

United Nations High Level Advisory Body on Artificial Intelligence

Submission

Human Technology Institute

September 2023

About the Human Technology Institute

The Human Technology Institute (HTI) is building a future that applies human values to new technology. HTI embodies the strategic vision of the University of Technology Sydney (UTS) to be a leading public university of technology, recognised for its global impact specifically in the responsible development, use and regulation of technology.

HTI is an authoritative voice in Australia and internationally on human-centred technology. HTI works with communities and organisations to develop skills, tools and policy that ensure new and emerging technologies are safe, fair and inclusive and do not replicate and entrench existing inequalities.

The work of HTI is informed by a multi-disciplinary approach with expertise in data science, law and governance, policy, and human rights.

Acknowledgement of Country

UTS acknowledges the Gadigal people of the Eora Nation, the Boorooberongal people of the Dharug Nation, the Bidiagal people and the Gamaygal people upon whose ancestral lands our university stands. We would also like to pay respect to the Elders both past and present, acknowledging them as the traditional custodians of knowledge for these lands.

Authors:

Professor Nicholas Davis, Co-director, Human Technology Institute

Dr Kate Bower, Fellow, Human Technology Institute

Dr Kobi Leins, Visiting Senior Research Fellow, King's College, London

To discuss this submission, please contact us at hti@uts.edu.au.

Introduction

HTI welcomes the opportunity to provide input to the UN High-Level Advisory Body on Artificial Intelligence (the Body).

Artificial intelligence (AI) systems are already ubiquitous in global digital ecosystems, yet their benefits are not equally shared, nor are their harms equally distributed. Virtually every company is now a tech company, with multiple reliances on contracted and third-party suppliers using data to nudge, influence, shift decision-making, and target advertising, to name a few examples.

Despite AI's ubiquity, countries' access to and regulation, governance, and consideration of the risks of AI systems differ vastly. This raises issues of interoperability and broader concerns about societal impacts that remain inconsistently acknowledged or addressed.

While some jurisdictions, such as the European Union and China, have moved relatively rapidly on the discussion and proposal of regulatory frameworks, standards development, and organisational governance approaches, most countries and regions are only just now starting to explore this area.

This combination of rapid technological rollout and uneven progress in terms of regulatory response makes serious and thoughtful efforts towards and research around practical forms of global AI governance both urgent and complex.

HTI believes that there is a clear need for far higher levels of coordination and a commitment to the interoperability of AI governance at the global level, including among both UN member states and multinational corporations across all sectors, including the prohibition of certain uses that undermine human rights or expedite the demise of democracy. Recent events have raised awareness across the public, industry and government stakeholders with regard to both the lack of specific governance and technical understanding required to use these tools safely and responsibly.

Motivations for global governance of AI

AI requires governance because AI and automation, if used correctly, can be exceptionally powerful. The decisions made about how data is collected, curated and connected have profound impacts on individuals' and communities' lives. The systems themselves are also propelling regional power and new political fault lines.

At the macro-level, fit-for-purpose governance approaches can ensure that rules, incentives, standards, and collective investments harness the scalable capabilities of AI systems to support the maintenance and expansion of critical, life-affirming public goods. International norms will be critical to ensuring that AI applications contribute positively to common objectives, help significantly accelerate progress towards the Sustainable Development Goals (SDGs) and human development more broadly. For example, AI applications offer new and innovative opportunities to effectively track and ensure compliance with the International Sustainability Standards Board's environmental compliance requirements.

AI governance also plays a crucial role in safeguarding human rights and meeting the SDG goals, including the protection of the environment, one of the greatest challenges we have seen in our lifetime. Robust and effective governance mechanisms at the global level must ensure that the benefits of AI are democratized, reaching all countries, not just those with the resources to fund AI system development, and share resources equitably and sustainably.

All UN member states stand to gain immensely from a globally coordinated AI approach. For industries, a standardized framework of AI governance will create more predictable market conditions and clearer paths to innovation. For individual organisations, a unified global AI governance framework reduces the complexity of navigating different regional regulations, allowing for more straightforward global expansion and collaboration. Consistent considerations of water and energy use equally being shared will help to avoid future conflict.

Global patterns around AI governance

Multiple 'fragmented'[1] approaches exist, with differing levels of efficacy.[2] To date, most international and national governance efforts around AI – including most of the efforts referred to above – have been primarily oriented towards establishing principles to guide AI development and use. The global bias towards principle-based regulation – as opposed to rules-based regulatory models – has been influenced by a range of factors, including the immense power the technology companies hold and their lack of appetite for regulation.

Despite best efforts from some researchers, there is very little to no empirical evidence of such principles and guidelines having impact on the actual development or use of AI systems at the level of governments or private sector organisations.[3] HTI's extensive research, as evidenced in our report the State of AI Governance in Australia[4] confirms this – interviews and workshops with over 300 organisations across Australia indicated that top-down, principle-based approaches are insufficient to the challenge of systematically ensuring that AI systems are accurate, accountable, fair and fit-for-purpose.

Of course, some national and subnational governance efforts have ventured beyond mere principles or ethical frameworks. Examples include the EU's draft AI Act, Canada's proposed Artificial Intelligence and Data Act (AIDA), China's rule-based governance for specific AI systems, and California's Bolstering Online Transparency Act[5], and the NSW Government's AI Assurance Framework in Australia. The rarity of these efforts underscore the need for more concrete, actionable governance mechanisms in the AI domain.

The opportunity for the UN High Level Advisory Body on AI

The complexity of changing the regulation, culture and literacy around the rapidly-changing, poorly-defined and commercially lucrative technology that is AI will require a multi-pronged approach. This is where the Body can play a critical role.

1) The Body can collect and coordinate initiatives to map where existing national and regional governance exists and can be analysed, reproduced, and shared, to avoid duplicating efforts and contribute to internationally-consistent approaches to AI governance.

2) The Body can engage independent global interdisciplinary experts (much like the expert working groups adjacent to the Chemical Weapons or Biological Weapons Conventions) to provide content and be canaries in the coalmine of rapidly generated research across multiple data science fields.

3) The Body can advocate for best practice and capacity building around AI risk identification, including setting "red lines" for peace and security-threatening uses of AI.

---

[1] See Lewin Schmitt (2022), "Mapping global AI Governance: a nascent regime in a fragmented landscape", *AI and Ethics* (Article 2022) 2:303-314, https://doi.org/10.1007/s43681-021-00083-y; also Anna Jobin, Marcelo Ienca, & Effy Vayena "The Global Landscape of AI Ethics Guidelines, Nature Machine Intelligence, (Article 2019) 1: 389-399 https://doi.org/10.1038/s42256-019-0088-2.

[2] State-led initiatives include AI-specific programs initiated by the G7, G20, Council of Europe, the Group of Governmental Experts on Lethal Autonomous Weapons Systems (GGE), the AI Partnership for Defence, the OECD's AI Policy Observatory and of course a range of governance efforts the United Nations, including the ITU and UNESCO. Hybrid, multistakeholder and non-state efforts include extensive efforts by technical standards bodies, such as IEEE, and ISO/IEC, the Global Partnership on AI, the Future of Life Institute, the AI Now Institute, the World Economic Forum's Centre for the Fourth Industrial Revolution, the Ada Lovelace Institute, and the Partnership on AI to Benefit People and Society.

[3] See for example A. McNamara, J. Smith, and E. Murphy-Hill (2018), "Does ACM's code of ethics change ethical decision making in software development?" in Proc. of the 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE). ACM, 2018, pp. 729–733, and Munn, Luke. (2022). The uselessness of AI ethics. AI and Ethics. 3. 10.1007/s43681-022-00209-w.

[4] Lauren Solomon and Davis, Nicholas (2023). *The State of AI Governance in Australia,* Human Technology Institute, Sep 2023

[5] In addition, it should be noted that, in every jurisdiction and in a wide number of international treaties and agreements, there exist a wide range of laws and regulations of general application which have both national and global relevance to the use and impacts of AI systems. These include privacy regimes that govern the collection, use and disclosure of personal data, such as the EU's General Data Protection Regulation, as well as product liability, intellectual property, trade and environmental laws.

4) The Body can support the integration (in appropriate forms of regulation) of international AI standards that allow states and companies to put globally-agreed AI principles into effective practice.

5) The Body can influence thought leadership. Such thought-leadership could partner with or take inspiration from the *Global AI Observatory* (GAIO) proposed by Carnegie Council for Ethics in International Affairs[6], which proposes standardised reporting of governance efforts in the form of an annual state of AI report, along with the creation and maintenance of registries that support the responsible governance of AI.

6) The Body can ensure best practice internally by the UN in its operations, ensuring the safety and security of its staff by using secure interoperable systems across its operations.

7) The Body can create a regular meeting place for experts with concerns to meet, not as their state representatives, but, similar to Pugwash, as technical experts who can provide insights and be the canaries in the coalmine. This group should also advise other UN bodies for whom this expertise is relevant, creating a global web of prevention.

8) The Body can encourage and support educational institutions to provide effective training at the intersection of AI and global public goods, such as the SDG goals. For example, this could include the signing of an 'oath' and having a list of certified AI practitioners who operate in compliance with human rights and the SDG goals.

9) The Body can play a critical role in championing the development of technical solutions and methods within AI models and systems that help organisations realise "responsible AI by design".

10) The Body could be given and execute on a mandate to design and propose a new, dedicated UN entity focused on supporting AI governance across the UN system and member states, taking care to ensure legitimacy through an independent expert group.

Key capabilities and functions

To meet the challenges of adequate representation of member bodies across the UN landscape and provide the normative legitimacy required, the Body will require a mix of technical, advocacy, policy research and dialogue functions.

This work will require sizeable funding and support to effectively undertake a systemic change of how data and AI are contemplated, used, governed, and applied, so buy-in and participation by technical experts will be key.

Conclusion

In Australia, HTI has led in the drafting of facial recognition draft regulation, consumer and human rights impacts of AI. At the global level, HTI and its partners have provided extensive input to the new AI Management Standard. We therefore welcome the opportunity to continue to engage and contribute on the global stage through the Body.

As Secretary-General noted in the Roadmap for Digital Cooperation, "there is a gap in international coordination, collaboration and governance".[7] HTI is pleased to work with the UN High-Level Advisory Body on AI to find innovative and effective ways to close this gap.

---

[6] Carnegie Council for Ethics in International Affairs, *A Framework for the International Governance of AI - Artificial Intelligence and Equality Initiative,* (Article, July 2023) <https://www.carnegiecouncil.org/media/article/a-framework-for-the-international-governance-of-ai>
[7] United Nations, *Report of the Secretary-General Roadmap for Digital Cooperation*, p18.