

Making AI work for everyone: join the global conversation

Position Paper 2024-01

February 2024

Wendy Hall, Ben Hawes

The United Nations Artificial Intelligence Advisory Body has asked for feedback on its report, [Governing AI for Humanity](#).

Recently, many experts in artificial intelligence have been arguing in public about how powerful AI could become, and when. Even seasoned observers have been surprised by how widely their opinions have diverged. However, nearly all experts agree that, in time, applications of AI will impact most industry sectors and areas of life, with many potential benefits and harms.

Antonio Guterres, Secretary-General of the United Nations made this one of the central themes of his [recent speech](#) to the World Economic Forum in Davos.

“We need governments urgently to work with tech companies on risk management frameworks for current AI development; and on monitoring and mitigating future harms. And we need a systematic effort to increase access to AI so that developing economies can benefit from its enormous potential. We need to bridge the digital divide instead of deepening it.”

If AI is going to affect all of, this isn't just a matter for governments and tech companies. Governing AI is everyone's business.

So there are good reasons for everyone to become better informed, and take part in the debates. You can support and take part in this action right now by reading and responding to the recent report by the United Nations Artificial Intelligence Advisory Body, which came out just before Christmas.

I will explain the background and purpose of this work, the process, and why you might want to contribute to it, using some of the key statements in the report.

Why should the United Nations get involved in governing AI?

“We believe that global AI global governance is essential to reap the significant opportunities and navigate the risks that this technology presents for every state, community, and individual today. And for the generations to come.” *Governing AI for Humanity*

The UN sees the need for this action because impacts from AI will be felt in every nation. Not every nation has the capability to develop ways to manage AI, and many of the potential benefits will not be realised at scale if each country develops its own rules and approaches independently. Establishing common principles and practices should improve the balance between benefits and risks for citizens of every country.

The UN already champions international cooperation to protect people's rights everywhere and to pursue common goals together, notably the Sustainable Development Goals. It's right that the UN should support work towards common principles for using AI around the world.

“Global governance with equal participation of all member states is needed to make resources accessible, make representation and oversight mechanisms broadly inclusive, ensure accountability for harms, and ensure that geopolitical competition does not drive irresponsible AI or inhibit responsible governance.” *Governing AI for Humanity*

Our work recognises that technology applications are often developed to suit specific and sometimes narrow commercial purposes. In AI, much of the leading capability and many of the critical assets are within a small number of companies whose activity is concentrated in a limited number of locations. Without appropriate governance, the power offered by AI may not be used in the interests of the majority of people in the world. Making sure that does not happen means agreeing common principles for governing AI that will work in practice.

The report highlights the stark inequality in access to digital technologies, and the potential for this to become worse as AI is developed only for some people and places.

“The AI opportunity arrives at a difficult time, especially for the Global South. An “AI divide” lurks within a larger digital and developmental divide. According to ITU estimates for 2023, more than 2.6 billion people still lack access to the Internet. The basic foundations of a digital economy — broadband access, affordable devices and data, digital literacy, electricity that is reliable and affordable are not there.” *Governing AI for Humanity*

Additionally, as well as minimising harms and reducing digital inequality, we need international collaboration to maximise the use of AI to positive purposes, including healthcare, education and sustainable growth. In the UK recently there have been debates about the balance between addressing current harms and longer term risks. In all that discussion of minimising harms, we are in danger of missing the other, equally critical principle: developing and applying AI for the benefit of humanity.

What is the United Nations Artificial Intelligence AI Advisory Body?

The Advisory Body is a group of experts who have been invited by the United Nations to develop common international approaches to governing AI. The Body has 38 [members](#) from governments, private sector, civil society, and academia, balanced

across gender, age, geographic representation, and area of expertise related to the risks and applications of artificial intelligence.

Members have extensive experience not only in developing technology but in applying it for social, economic and environmental benefits. While most members work in leading institutions, and some of them in several, on this Body they serve in their personal capacity.

What really sets us apart from the other valuable international initiatives is our membership, which includes Russia and China, and our determination to ensure that the Global South is not excluded from the development and use of AI. And we interpret that commitment broadly: as well as recognising the needs of poorer countries with comparatively fewer digital resources and capabilities, we are concerned with ensuring that poorer and relatively excluded people and groups in all countries are not disadvantaged as AI develops.

We are not working in a wholly new and empty space. The Body will work with existing processes and platforms, and regulatory and governance structures. It will work through existing bodies where possible, including existing UN structures.

“Governance is not an end but a means, a set of mechanisms intended to exercise control or direction of something that has the potential for good or ill.”
Governing AI for Humanity

We will work through a series of stages of reports and consultation, to reach conclusions on how to govern AI which can attract broad international support and influence how the technology is used globally.

International context: regulating and governing AI

Our work for the UN is clearly not the only activity in oversight of AI.

There have been many recent developments in regulation. The [EU AI Act](#) has been passed, though there are still many details of implementation to be worked out. The EU has a record of influencing other countries, whose companies want to sell into EU markets. The tendency for industry sectors to accept EU standards as the basis for global standards is known as the Brussels effect. Many people working in AI outside the EU will be following the development of the details of implementation of the EU AI Act.

The US is working through a major package of proposals, last October’s [Executive Order](#) on the Safe, Secure, and Trustworthy Development and Use of AI. China is developing regulation of supplication areas including AI-generated video and generative AI.

The UK government has just published its [response](#) to the [public consultation](#) on its “pro-innovation” approach to regulating AI. The Response keeps the same course, holding back from new regulation of AI in the medium term, but with additional recognition of risks and noting that legislation will be needed in time. This approach puts a lot responsibility onto established sector regulators, and they are asked to

publish plans in the coming months setting out how they will respond, supported by increased coordination and cross-sector guidance.

The Response also sets out “the case for a set of targeted, binding requirements on developers of highly capable general-purpose AI models in the future to ensure that powerful, sophisticated AI develops in a way which is safe.”

A UK Parliamentary Committee has also published a very considered [report](#) on large language models and generative AI.

However, the United Nations Body’s report is not about the detail of regulatory approaches, but at a higher level about governing AI. Most of the members of the Body are engaged in some way with one or more of the various government actions in process around the world. But we are doing something different and additional here. We aim to promote productive collaboration on how AI can support and promote existing shared global objectives on human rights and the Sustainable Development Goals.

And that it is important: rights in relation to AI need to be meaningful and coherent with shared human priorities, not separate to those.

“We look forward to engaging with diverse stakeholders as we answer more fully the questions identified in this interim report, in support of the ongoing efforts of the United Nations on digital cooperation and on social progress and better standards of life in larger freedom.” Governing AI for Humanity

What is in this report?

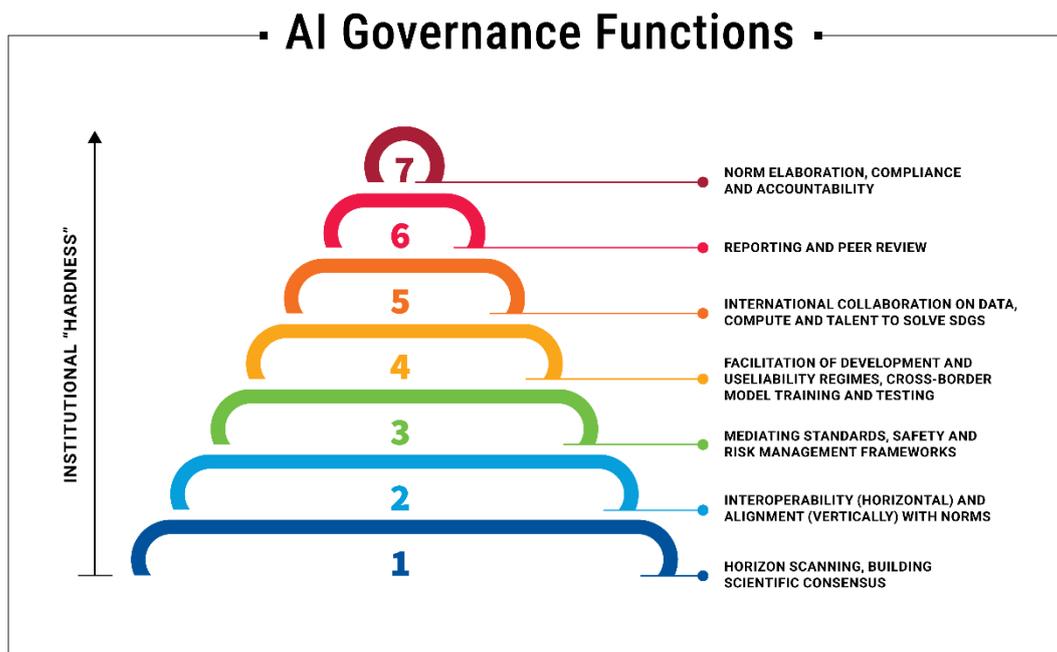
This first public report is a statement of where our discussions have reached, and we are putting this out in public for everyone to respond to. It proposes core principles and functions for any global AI governance framework, and invites challenges, related evidence and additional proposals, in order to refine and agree the developed principles and functions.

Guiding Principles

The interim report identifies the following principles that should guide the formation of new global AI governance institutions:

- **Inclusivity:** all citizens, including those in the Global South, should be able to access and meaningfully use AI tools.
- **Public interest:** governance should go beyond the do no harm principle and define a broader accountability framework for companies that build, deploy and control AI, as well as downstream users.
- **Centrality of data governance:** AI governance cannot be divorced from the governance of data and the promotion of data commons.

- **Universal, networked and multistakeholder:** AI governance should prioritize universal buy-in by countries and stakeholders. It should leverage existing institutions through a networked approach.
- **International Law:** AI governance needs to be anchored in the UN Charter, International Human Rights Law, and the Sustainable Development Goals.



While the report sets out principles, we want to bring in everyone’s views on what will work in practice.

“AI governance must deliver tangible benefits and safeguards to people and societies. An effective global governance framework must bridge the gap between principles and practical impact.” Governing AI for Humanity

More detail and the process for responding are [here](#). The AI Advisory Body and its members will engage with all stakeholders over the coming months. They invite individuals, groups, and organizations to submit contributions by 31 March 2024.

Members of the Body will engage with governments, the private sector, civil society, and research and technical communities. At the same time, they will proceed with research, including on methodologies for risk assessment and how governance could function effectively across countries and sectors. They will look in particular at themes that are already widely recognised as critical, including Open Source, AI and the financial sector, standard setting, intellectual property, human rights, and the future of work.

What happens next?

When the Body has received and worked through responses to this report, the process will move on to design of the institutions. The principles would guide the formation of new global governance institutions. The functions would be the actions such institutions would need to perform, and should “add value, fill identified gaps, and enable interoperable action at regional, national, industry, and community levels.”

The idea is that form should follow function: when principles and functions are settled, there will be further consultation on what institutional structures would most effectively put those into practice.

The final report of the Body will be released by 31 August 2024, ahead of the [Summit of the Future](#) in September. Depending on the input we receive and the options that emerge, we hope to be able to provide detailed recommendations on the functions, form, and timelines for a new international agency for the governance of artificial intelligence.

I hope that people and organisations will take this opportunity to convene events around this report, and send back summaries of discussions, with any suggestions that emerge.

About the Authors

- Professor Dame Wendy Hall, Regius Professor of Computer Science at the University of Southampton
- Ben Hawes, technology policy consultant and Associate Director at the Connected Places Catapult

About WSI

The Web Science Institute (WSI) draws together the University's world-class, interdisciplinary, sociotechnical expertise in Web Science, Data Science and Artificial Intelligence. We act as a focus for international esteem as we create new opportunities to bring faculties, schools, and disciplines together to leverage the unique role of online technologies in tackling global challenges, including the challenges posed by society's use of those technologies themselves.

The WSI was established to study the evolution of the Web and society but has evolved into an institute that specialises in the sociotechnical study of the evolution of digital technologies and society in general, focussing currently on, but not restricted to, the new discipline of Human-Centred Artificial Intelligence (HCAI) as well as Web Science.

Copyright © the authors 2024

The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the University of Southampton, the Web Science Institute or its Board of Directors.



possibly change

This work is licensed under a Creative Commons Attribution — Non-commercial — No Derivatives Licence. To view this licence, visit (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) For reuse or distribution, please include this copyright notice.

Web Science Institute, Building 32, Highfield Campus, University of Southampton, SO17 1BJ
wsis@soton.ac.uk (DOI: 10.5258/SOTON/WSI-WP008)