



Artificial intelligence is too important to be left either to a handful of companies or to a small number of powerful states, the writers say. They propose an international AI agency that is independent, science-based, globally representative, and capable of helping the world capture AI's benefits while reducing its risks and sharing its gains more fairly. PHOTO: ADOBE STOCK

We can't wait for an AI crisis before taking action

Technology is outpacing regulatory efforts. The UN meetings in Geneva this week are a chance to begin building a global agency for AI governance.

Simon Chesterman and Wendy Hall

On July 2, OpenAI's Sam Altman called for a "simple framework" for governing advanced artificial intelligence. His proposal is not nothing. A US-led forum involving governments, independent technical experts and others would be better than leaving every country to improvise alone.

But the more striking point in his op-ed in the *Financial Times* is what is missing. Just as the United Nations process on AI governance enters its most important phase yet, the UN does not appear at all.

This week in Geneva, scientists, governments, companies and civil society will gather for the first session of the UN Global Dialogue on AI Governance. The new International Scientific Panel on AI will present its preliminary report.

The Dialogue is not perfect. UN processes move slowly. Yet it has one quality no club of powerful states or companies can claim – every country has a seat at the table.

That matters because AI governance is not only a question for those who build powerful models. It is also a question for those whose economies, schools, hospitals, labour markets and political systems will be reshaped by them.

Altman writes easily of joining

G7 leaders and AI "peers" at their gathering in Evian, France. The image captures the central challenge of AI governance: public authority trying to catch up with private power. Frontier AI companies are now invited into rooms once reserved for states. They should be heard. But they should not set the terms on which humanity is governed.

We, the co-authors, were both involved in the UN AI Advisory Body that first argued, in 2023, that AI is too important to be left either to a handful of companies or to a small number of powerful states. Since then, the pace of technical development has accelerated, geopolitics have hardened, and the social consequences have grown more visible.

The risks are no longer speculative. The rise of agentic AI, emerging cybercapabilities, recent revelations of security vulnerabilities in systems worldwide, growing evidence of labour-market disruption, and a widening gap between countries that shape AI and those merely shaped by it all point to the same conclusion – the present arrangements are insufficient.

At the same time, AI could transform health, education, climate adaptation, agriculture and scientific discovery. The danger is not only harm, but that benefits will be captured by a few while risks are distributed to many.

The UN has made a serious start in addressing these challenges. The Global Dialogue, Scientific Panel and UN Office for Digital and Emerging Technologies deserve support. But they should be understood as the beginning of a process, not its destination.

Four problems highlight the need to be more ambitious.

The first is speed. AI capabilities are advancing faster than governments can legislate, regulators can staff up, or courts can assign responsibility. The world should not wait for an AI catastrophe before building the institutions needed to prevent one.

The second is fragmentation. National regulation is necessary but insufficient. The EU, the United States, China, the United Kingdom, Singapore, and others are each moving in different directions, yet the technologies, firms, models, data flows, and harms are all transnational. Most states lack the capacity to regulate frontier AI meaningfully; many will simply import rules, tools, and dependencies designed elsewhere, with no voice in their making.

The third is accountability. Many leading AI companies ask governments and societies to

trust them while resisting meaningful constraints, external scrutiny, or liability for harms. Voluntary commitments have a role, but they cannot substitute for public authority. A market dominated by a small number of firms cannot define the public interest on behalf of everyone else.

The fourth is equity. AI could transform health, education, climate adaptation, agriculture, and scientific discovery. But without deliberate governance, it will also deepen existing divides in compute, data, talent, infrastructure, and bargaining power. Global AI governance must therefore be about opportunity as well as risk – not only preventing catastrophe, but ensuring that the benefits are more widely shared.

THE CASE FOR AN INTERNATIONAL AI AGENCY

The next step should be a blueprint for an International Artificial Intelligence Agency: independent, science-based, globally representative, and capable of helping the world capture AI's benefits while reducing its risks and sharing its gains more fairly.

Such an agency would not be world government for AI. Nor would it license every model. Its work should be practical: maintaining a global evidence base; supporting common standards for safety testing, model evaluation, incident reporting, and monitoring; helping governments build regulatory capacity; and providing a trusted forum for sharing frontier-risk information.

And it would not start from scratch. The network of AI institutes, now the International Network for Advanced AI Measurement, Evaluation and Science, is already developing shared approaches to measurement and evaluation.

The UK's new Centre for AI Measurement points in the same direction. An international agency could give these efforts a global home and ensure that capacity is not confined to wealthy states.

This need not begin as a universal treaty body. Waiting for unanimity would mean paralysis. A coalition of committed states, scientific institutions, civil society actors and responsible companies could begin the work, building legitimacy over time.

This challenge is sometimes framed as an East-West problem, with China cast as the central obstacle. In fact, it is primarily a public-private problem: all states, regardless of geopolitical alignment, are grappling with technology developed in private hands with minimal public oversight.

The United States and China will be indispensable in the long run. But they should not be permitted to hold the entire process hostage at the outset. Middle powers and AI-capable states – among them the United Kingdom, France, Germany, Switzerland, Singapore, India, Japan, South Korea, and the Gulf states – have both the interest and the capacity to build momentum towards a durable institution.

Companies must be part of the conversation, but not owners of it. Civil society must be at the table, not invited only to endorse decisions already made. Scientists and engineers must provide the evidence base, but cannot be asked to resolve distributional and political questions alone.

The legitimacy of global AI governance will depend on whether it is genuinely inclusive, not merely multi-stakeholder in name.

Around the world, awareness of the risks of AI is growing. Young people worry about their jobs. Cybersecurity professionals are raising alarms. Students question what learning means in an age of AI. Governments are asking whether they can maintain public trust.

That concern is a resource. It needs to be channelled into institution-building before it dissipates – or before a crisis forces our hand.

The lesson of past technological revolutions is that it is almost always cheaper to build guard rails before disaster than to repair the damage afterwards. The first AI crisis may be a cyberattack, a financial shock, a labour-market rupture, a biological incident, or something we have not yet imagined. If we wait until then to build serious international governance, we will already have failed.

The meetings in Geneva this week are a beginning. They should mark the moment when global AI governance begins to acquire institutional form. A simple framework is welcome. But simple cannot mean selective.

If AI is to be governed for humanity, the United Nations must be part of the answer – and the world should begin building the institution that can turn principles into practice.

• Simon Chesterman is David Marshall Professor of Law at the National University of Singapore, where he is also AI governance and policy lead at the NUS AI Institute and dean of NUS College. Wendy Hall is Regius Professor of Computer Science at the University of Southampton, where she is also director of the Web Science Institute. She co-chaired the UK Government's 2017 AI Review.

The rise of agentic AI, emerging cybercapabilities, recent revelations of security vulnerabilities in systems worldwide, growing evidence of labour-market disruption, and a widening gap between countries that shape AI and those merely shaped by it all point to the same conclusion – the present arrangements are insufficient.