

AI and the G20: Striking a balance between innovation and governance

Grzegorz Stal
University of Sheffield

Executive Summary

The economic benefits of Artificial Intelligence (AI) are no longer speculative; they are both measurable and transformative. According to [Goldman Sachs Research](#), AI could drive a 7% boost in global GDP, lift productivity growth by 1.5% over a decade, and create considerable positive social impact. Representing nearly 85% of global GDP, the G20 has both the influence and capacity to lead in this transformative era. Reflecting on its pivotal role in addressing the [2007-2008 global financial crisis](#), the G20 has demonstrated its ability to provide decisive leadership during periods of economic transformation. This history uniquely positions it to guide the integration of AI into global economies and establish a robust framework of global governance.

AI at Previous G20 Summits

Alongside the rapid [expansion of machine learning over the last decade](#), G20 leaders have increasingly recognised the transformative potential of artificial intelligence and digital technology in driving economic growth as well as the risks associated with it. The [2019 Osaka Summit](#) marked the G20's [first focused discussion](#) on artificial intelligence, where leaders highlighted both its economic potential and the need for responsible governance among the nations. Prime Minister Shinzo Abe introduced the pivotal concept of the 'Osaka Track' including the '[Data Free Flow with Trust](#)' (DFFT) promoting cross-border data sharing with strong privacy standards. The summit also adopted the non-binding G20 [AI Principles](#), based on OECD guidelines, advocating for human-centred, trustworthy AI development.

Nonetheless, the Osaka Track and its associated principles were not without controversy. Countries like India and South Africa [boycotted the framework](#), concerned over its implications for digital sovereignty and equitable development. Consecutive summits in [Riyadh, Rome, and Bali](#) did not bring any groundbreaking changes in terms of AI policy development. They reinforced existing commitments rather than concrete-focused advancements. The [New Delhi 2023 summit](#) brought a fresh wind of change. Prime Minister Modi urged the G20 to prioritise 'Harnessing Artificial Intelligence Responsibly for Good and for All' by reaffirming the 2019 G20 AI Principles, adopting a pro-innovation regulatory approach, and promoting responsible AI for achieving the Sustainable Development Goals. The G20 in Rio builds on New Delhi's commitments, slowly translating them into actionable steps for global AI governance.

Global Perspectives on AI

The Global North faces an unprecedented [demographic challenge](#) as ageing populations and workforce shortages threaten to strain economic growth and social welfare systems. In the US, the economy requires an additional 4.6 million workers annually to maintain current levels of supply, demand, and population balance, equating to 2% of the US population. Similarly, Germany needs to

find 1.6 million workers (3% of its population) to sustain its economic equilibrium. Incorporating AI and robotics into healthcare can help address workforce shortages and improve patient outcomes. As an example, solutions such as [IBM's Watson](#) for Health leverages cognitive computing to process vast amounts of medical data far more efficiently than humans, enabling precise and swift diagnoses and thus alleviating pressure on overburdened healthcare systems while supporting economic sustainability amidst workforce shortages.

While productivity is a cornerstone of economic prosperity, regions like Europe and countries like the UK have faced significant challenges in this area over the past decade. The [UK's productivity](#) has been notably sluggish since the 2008 financial crisis, with growth rates lagging behind that of other nations such as China and the US. Similarly, Europe has experienced a [decline in productivity growth](#), raising concerns about sustaining welfare systems and economic competitiveness. Particularly in [service economies](#), AI-powered tools such as chatbots and automation can bring benefits with streamlined operations and enhanced customer engagement, presenting itself as a potential solution to the Western 'productivity puzzle' by reducing the present inefficiencies and driving innovation.

All of the mentioned benefits and challenges ultimately pave the way for Osaka's summit vision for Society 5.0, where technology seamlessly integrates with daily life to enhance the quality of living and societal outcomes. The Rio Summit took this vision to the next level, placing a bold focus on actionable strategies to integrate advanced technologies into daily life. G20 countries of Global North can serve as the drivers in the pro-innovatory governance with technological advancement balanced alongside fairness and safety building on earlier established AI Principles, leading to [AI democratisation](#). This ultimately sets the stage for Rio's drive for the establishment of AI guidelines in labour markets across the globe under next year's South Africa's G20 Presidency.

Meanwhile, in the Global South, AI holds the transformative potential for its [urgent developmental needs](#) such as job creation, industrial development, and long-term economic growth, with a focus on more practical applications. [Latin America](#) particularly Lula's AI plan for Brazil committing approximately US\$4 billion to benefit from AI in fields such as agriculture. As highlighted in the [IEEE review](#), integrating AI into sectors like farming, where technologies such as satellite imagery and machine learning are revolutionising operations by enabling precise weed identification, yield prediction, and resource optimisation,

boosting both productivity and sustainability, exemplifies the possibilities.

In Africa, AI offers transformative potential for addressing the urgent developmental challenges of the region, particularly in education. [The African Union's report](#) underscores the critical role AI can play in bridging educational gaps, fostering digital literacy, and preparing the continent's youth for the AI-driven global economy. Initiatives like [Rwanda's Smart Education Project](#) expand internet access and upskill the workforce, while AI-powered tools automate grading and enhance teaching, moving towards more inclusive and digital education across the continent.

The Global South as a whole has an opportunity in AI which promises prosperity and expansion of the global digital economy. While acknowledging the digital divides within and between countries, southern G20 members can lead in a call for the promotion of inclusive international cooperation and capacity building for developing countries in this domain. Echoing President Xi's words in Rio that AI should not be a ['game of rich countries and the wealthy.'](#) the countries of the Global South can advocate for equitable access to AI's benefits, emphasising its transformative potential for addressing their developmental priorities. By pushing towards more inclusive frameworks, capacity building, and fair governance, developing countries can position themselves as key players in shaping AI's

global landscape, ensuring that its promise as mentioned by Lula is ['for the Good of All'](#).

The Regulation Dilemma

Visions of robot rebellion and human extinction, reminiscent of Kubrick's *2001: A Space Odyssey*, are largely exaggerated. While such fears tend to capture our imagination, the real need for AI regulation lies in addressing more immediate and [mundane concerns](#). Existing AI systems raise pressing issues around bias, privacy, and intellectual property rights, the challenges that demand urgent attention to ensure ethical and responsible deployment of the technology in a human-centric manner protecting people's rights and safety. One of the pressing priorities of G20 is to guide a direction protecting the public from those associated risks.

Bias in AI systems is a significant concern, as algorithms often reflect and amplify societal prejudices present in their training data. [Amazon](#) developed a machine learning tool to review job applications. Still, it exhibited gender bias by prioritising male candidates due to skewed training data. Similarly, [facial recognition technologies](#) have demonstrated higher error rates for women, particularly those with darker skin tones, leading to potential misidentification and discrimination. These examples highlight the necessity for vigilant oversight and the implementation of strategies to mitigate gender bias in AI,

ensuring equitable and fair outcomes across all applications.

Another challenge stems from intellectual property issues, adding further complexity to the AI landscape. Generative AI systems, which create new content based on existing data, blur the lines of authorship and ownership. Questions arise regarding who owns the rights to AI-generated works: the creator of the AI, the entity using it, or the individuals whose data may have been used in training. This uncertainty poses significant legal and ethical challenges for industries reliant on creativity, such as art, music, and publishing. Moreover, the emergence of deepfakes — realistic but fabricated images, audio, or video further exacerbate these concerns. Cases such as the [deepfakes involving Joe Biden](#) highlight how such technologies not only erode trust in digital media but also complicate accountability, as tracing the origin and intent of these creations becomes increasingly difficult.

On the other hand, over-regulation could pose a significant threat to the pro-innovation regulatory/governance fueling AI's transformative potential which the G20 strives for. Excessive rules and bureaucratic barriers may deter investment, slow the pace of technological advancement, and limit the ability of smaller firms or developing nations to compete in the AI space. Some argue that AI's real danger lies in its role as a ['digital bureaucrat'](#) by subtly amplifying its

influence within systems like banking or social media rather than rebelling outright. For AI Policy '[boring is good](#)' technocratic, measured approaches are preferable to reactionary or overly stringent regulations.

Often the only result of overregulation is a complex and often-conflicting regulatory regime that accomplishes little more than disadvantaging small and medium-sized enterprises (SMEs) while benefiting tech giants like Google, Meta or Amazon, which can better absorb compliance costs. Looking no further than the EU's 'revolutionary' AI act, the full implementation of which may [cost](#) developers €193,000-330,000 upfront plus €71,400 in yearly maintenance costs. While these rules aim to ensure 'algorithmic fairness', such well-intentioned measures risk stifling innovation with proven life-saving potential, weakening national economies, and handing a competitive advantage to countries with less restrictive frameworks.

As AI continues to evolve, a central consideration for G20 nations is striking a balance between fostering innovation and ensuring robust ethical governance. Adopting a human-centred approach is crucial not only to unlock AI's economic potential but also to support the well-being of wider societies. [Collaborative models](#), where AI enhances rather than replaces human capabilities, represent the optimal path forward. To address pressing concerns like bias and privacy, advocating

for solutions such as [human-in-the-loop systems](#) can mitigate risks and improve decision-making in critical risk sectors like healthcare, public administration, and finance where humans undertake the decisions. By prioritising measured, technocratic regulation over reactionary policies, the G20 can guide the development of AI frameworks that encourage fairness, safeguard human rights, and stimulate inclusive growth based on adopted [OECD AI Principles](#), where AI remains a tool for global benefit rather than a source of international division or public harm.

G20 in Rio: A Bold Vision for AI

The G20 Summit in Rio marked a critical moment in the global conversation on artificial intelligence (AI), highlighting both its transformative potential and disruptive risks. Leaders reaffirmed AI's ability to drive economic growth, improve societal outcomes, and bridge global divides, while also recognising challenges like bias, privacy concerns, and the amplification of misinformation. A pro-innovation regulatory approach was emphasised, aiming to mitigate risks while fostering AI's transformative capabilities through human-centric and inclusive governance. Notably, beyond reaffirming the G20 AI Principles, the inclusion of UNESCO's Ethics of AI further strengthened the focus on ethical and responsible AI development.

Central to the discussions was the recognition of AI's role in reducing inequalities through digital public infrastructure and advancing societal goals, such as halving the gender digital divide by 2030. At the same time, leaders acknowledged the risks of AI-driven misinformation and hate speech, calling for transparency and accountability from digital platforms. Reiterating principles from past summits, the G20 reinforced the importance of cross-border data flows and the DFFT.

In the workplace, the G20 committed to developing guidelines for the safe and ethical integration of AI, stressing the importance of incorporating worker feedback and prioritising fairness. Initiatives to promote digital literacy and empower vulnerable communities underscored the G20's intent to address inequalities while leveraging AI's benefits. However, the big concern remains about the feasibility of these goals without binding agreements or clear implementation strategies, raising doubt about G20's ability as a driver of meaningful change within this area.

The establishment of a High-Level Initiative on Artificial Intelligence and Innovation under South Africa's presidency signals continued ambition, but whether this initiative will lead to actionable outcomes remains uncertain. Without binding agreements or detailed guidelines, the G20 risks remaining a platform for discussions

rather than a driver of meaningful change. Some critics argue that the G20 must become [leaner and meaner](#) to maintain its relevance, prioritising a streamlined agenda and tangible, measurable outcomes over broad, non-binding commitments. Nevertheless, despite these uncertainties, the G20 has taken an important step toward shaping a balanced global AI governance framework. For proponents of light-touch regulation, the Rio commitments strike a promising balance between fostering innovation and safeguarding societal interests.

Conclusion

AI should no longer be regarded merely as a technological advancement but as a transformative force for global economic progress, deeply aligned with the founding principles of the Group of Twenty, which seek to foster international economic cooperation and promote global stability. In future summits, the G20 must take the lead in shaping the future of artificial intelligence by building on successful initiatives like the [UK's AI Safety Summit](#) at Bletchley Park and upcoming [France's Summit for Action on Artificial Intelligence](#), while reinforcing frameworks such as the UNESCO Ethics of AI and the G20 AI Principles.

By establishing a unified, inclusive global framework for AI governance, the G20 can set the standard for ethical, responsible, and innovative AI development that

addresses societal challenges like inequality and privacy while unlocking economic growth. This leadership role will not only help bridge the digital divides between and within nations but also position AI as a tool for sustainable development and shared prosperity. With decisive, forward-looking action, the G20

has the opportunity to guide AI's evolution into a global force that benefits all.

Grzegorz Stal is a BA Business Management & Economics student in the Management School at the University of Sheffield.