unesco Chair

IA.EDU

Seven Global AI Challenges and why they cannot be addressed in isolation

No one challenge can be addressed in isolation.
Global cooperation, inclusion and an evidence-based approach are essential to ensure that AI advances human well-being, sustainable development and peace.

UNIVERSIDADE FEDERAL DE ALAGOAS

CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico

FAPEAL FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE ALAGOAS

inct institutos nacionais de ciência e tecnologia

EXECUTIVE SUMMARY

TITLE
Seven Global AI Challenges and why they cannot be addressed in isolation

AUTHORS
Ig Ibert Bittencourt, Seiji Isotani, Thomaz Edson Veloso

EDITION
1st

YEAR OF PUBLICATION
2026

EDITORIAL COORDINATION
Eloi Capucho Ferreira

GRAPHIC DESIGN
Eloi Capucho Ferreira
Kevvin Sales da Costa de Oliveira

HOW TO CITE THIS DOCUMENT
Bittencourt, I. I., **Isotani**, S., & **Veloso**, T. E. (2026). *Seven global AI challenges and why they cannot be addressed in isolation* [Policy note] (23 pp.). IA.Edu. DOI: https://doi.org/10.67159/iaedu.seven-global-ai-challenges

This publication was developed by the Institute of Artificial Intelligence in Education – IA.Edu within the framework of the UNESCO UNITWIN Chair on Artificial Intelligence Unplugged in Education

SUMMARY

4

Artificial intelligence (AI) has moved from a specialized technology into a general-purpose infrastructure that now shapes economies, public institutions, education, health, science, security, and everyday life. Researchers are tracking three linked trends: how AI is changing the way people interact with technology, how it is reshaping economic activity, and how it may become difficult for anyone to fully control. In response, international organizations and governments are working to understand how to govern AI before its risks become harder to manage.

Two recent reports frame this policy note. The Preliminary Report of the UN Independent International Scientific Panel on Artificial Intelligence (2026) gives an evidence-based assessment of AI's transformation and governance challenges. Its central message is balanced: AI should be neither celebrated uncritically nor halted outright. The opportunities, risks, and impacts of AI depend on the choices societies make about how to develop, deploy, and govern it. The 2026 International AI Safety Report, led by Yoshua Bengio and written by more than 100 experts, complements this by cataloging risks of malicious use, malfunction, and broader systemic effects. The reports suggest that AI should be understood as a complex sociotechnical system rather than simply as a technology requiring isolated risk controls. For example, increasing computational capabilities may generate productivity gains while simultaneously affecting labour markets, concentrating economic power, increasing energy consumption, enabling new security threats, and challenging existing governance institutions. Similarly, measures designed to address one challenge may generate consequences elsewhere. Effective AI governance therefore requires understanding each challenge individually, as well as the relationships, dependencies, and potential trade-offs among them.

This policy note draws on these and other reports to make three points for decision-makers:

First, AI is advancing extremely quickly, but the infrastructure, investment and expertise behind that advance are highly concentrated in a small number of countries and companies.

Second, this transformation creates seven interconnected global challenges that go well beyond narrow technical "AI safety" concerns.

Third, governments and international organizations need to shift from reactive rule-making toward governance that is inclusive, anticipatory, evidence-based and focused on implementation.

1. AI ADVANCEMENT AND THE TWOFOLD GAPS

Understanding today's AI governance debate starts with the speed, scale and concentration of AI advancement. For example, Figure 1 places AI adoption in historical perspective, showing that generative AI is spreading through workplaces considerably faster than earlier general-purpose digital technologies such as personal computers and the internet—a signal of the unusually rapid pace at which AI could diffuse across economies and societies (International Energy Agency [IEA], 2025).

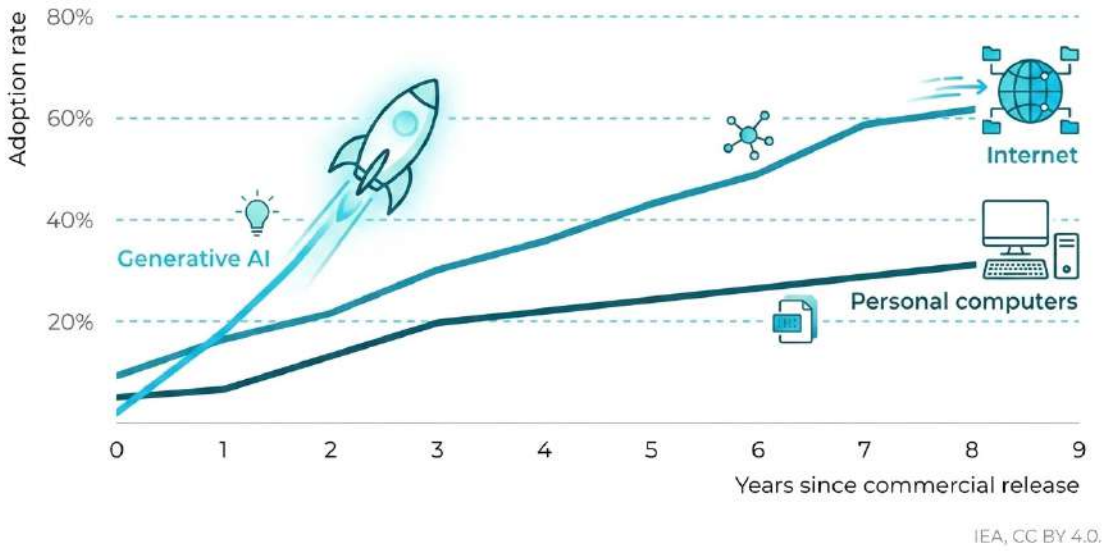


Fig. 1. Growth in U.S. workplace adoption of digital technologies (internet and personal computers) since their commercial release. *Source: International Energy Agency, Energy and AI (2025).*

One clear sign of AI's advancement is how much longer and more complex a task an AI system can now complete on its own. Figure 2 shows that early systems could reliably finish only very short software tasks, while today's frontier models can complete tasks that would take a skilled professional many minutes. Researchers estimate that the length of software tasks AI can reliably complete doubles roughly every seven months.

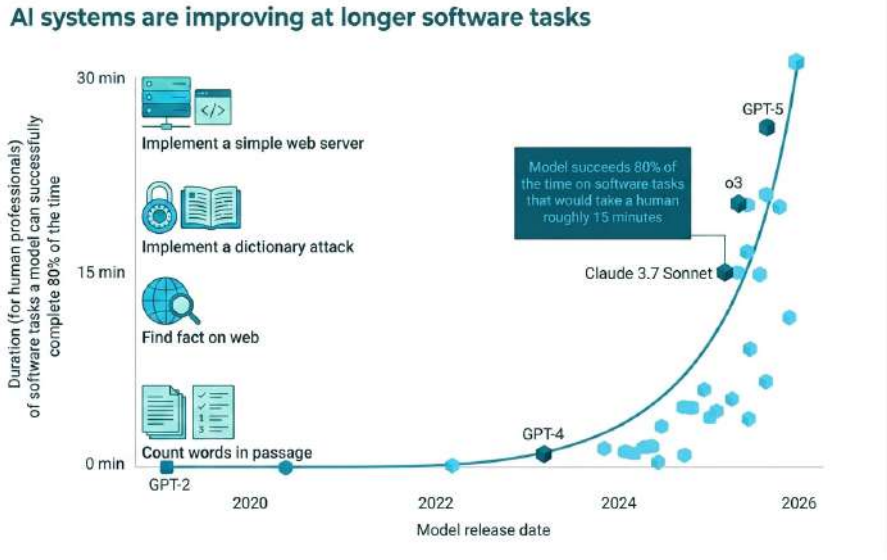


Fig. 2. Growth in the length of software tasks that successive AI models (GPT-2, GPT-4, GPT-5) can complete reliably. *Source: Bengio et al., International AI Safety Report 2026 (arXiv:2602.21012)*

This trend suggests AI is moving beyond answering single prompts toward sustained reasoning, planning and action with less human involvement. Performance is still uneven, and reliability drops on longer tasks, but this growing autonomy could significantly affect productivity and how work is organized, while also raising the scale of potential risks.

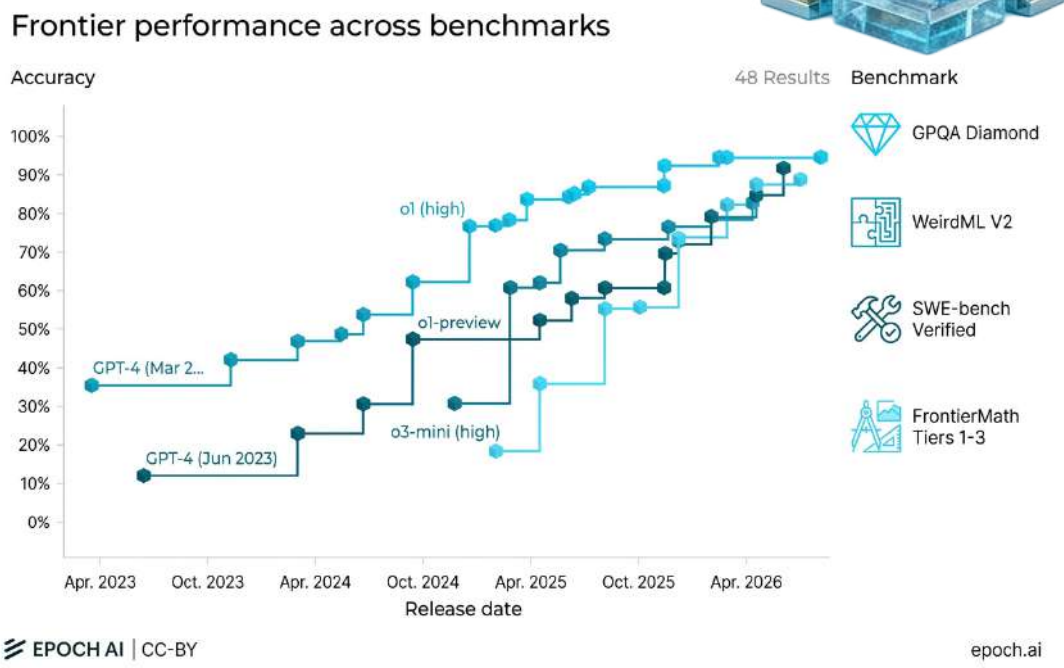


Fig. 3. AI model performance accuracy across several benchmarks, illustrating uneven capability across task types.
Source: epoch.ai

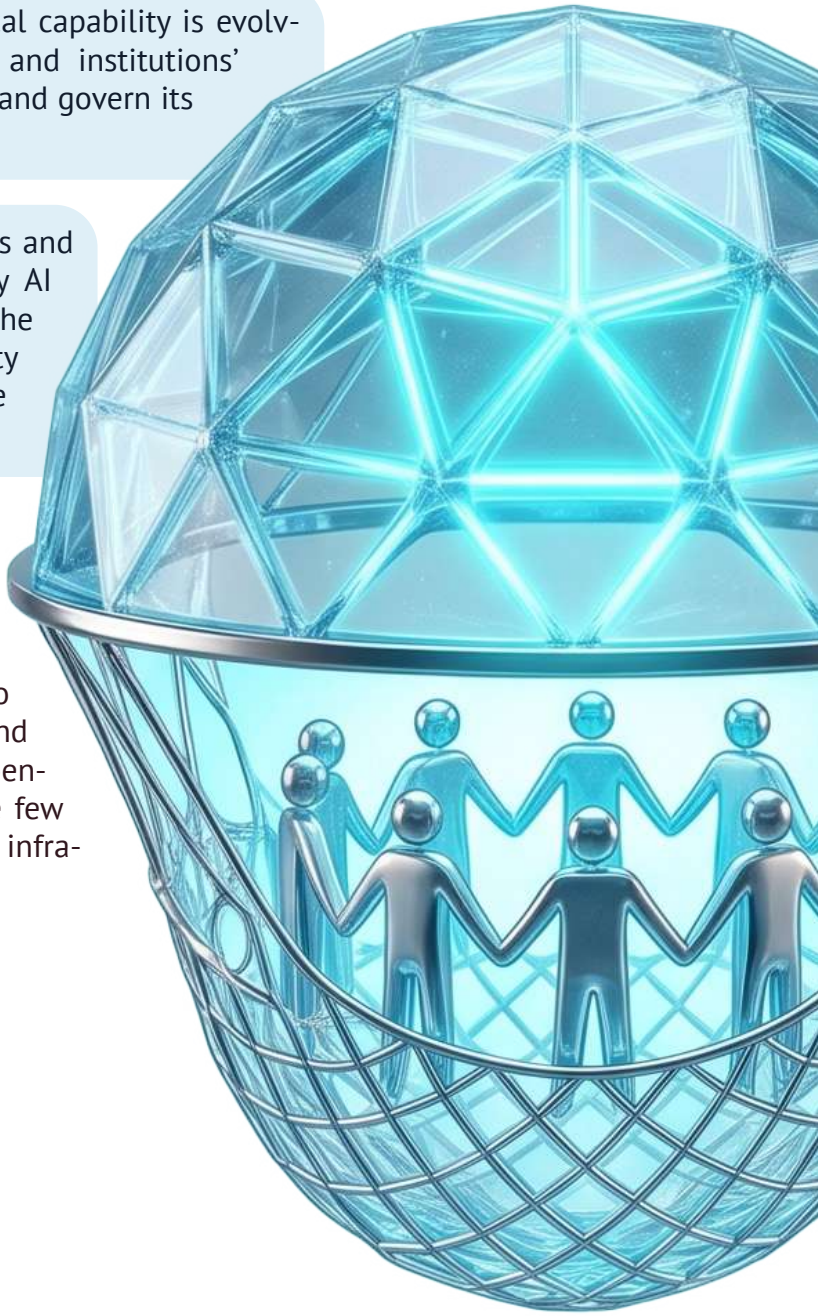
The Epoch Capability Index shows a related structural shift, where AI infrastructure has expanded very rapidly, and commercial revenue has only recently begun to catch up. Both infrastructure and revenue remain concentrated among a small number of companies. For this reason, today's AI transformation is not just a story of better algorithms, but it also rests on a connected supply chain of advanced models, specialized computing infrastructure, data centers, semiconductors, cloud platforms, large datasets, scientific talent, and capital. Countries and institutions without access to these resources risk becoming consumers of AI rather than developers of it or participants in how it is governed.

7

Although AI creates new opportunities, such as accelerating scientific discovery, raising productivity, widening access to expertise, and improving public services such as health screening, disaster forecasting and personalized education, access to an AI model does not automatically translate into economic productivity, better education, stronger healthcare, or more capable institutions. Countries, organizations and communities also need complementary capabilities (i.e., human capital, institutions, infrastructure, local data, regulatory capacity and organizational readiness) to turn technological possibility into public value. AI advancement therefore presents a double asymmetry:

Advancement gap: technological capability is evolving faster than governments' and institutions' ability to understand, evaluate and govern its risks and benefits.

Participation gap: the countries and communities most affected by AI are not necessarily those with the infrastructure, research capacity or market power to influence how it is developed.



The global policy challenge is therefore not simply to "keep pace with AI." It is to ensure that technological advancement translates into widely shared human, social, and economic gains—rather than benefits concentrated in the same few places that already hold the infrastructure.

2. SEVEN GLOBAL AI CHALLENGES

Evidence from the UN Panel suggests AI risk should not be treated as a single category of "AI safety." It spans a broader landscape covering societal applications, human flourishing, governance, technical reliability, economic transformation, international security, and environmental sustainability — and these challenges are closely interconnected.

Fig. 4. The Seven Global Challenges of Artificial Intelligence

2.1. SURFACE AI GOVERNANCE CHALLENGES

Surface AI governance challenges are the most visible and immediate consequences of AI adoption. They emerge as AI is deployed across society and cover societal applications, economic impacts, and security and peace. They tend to draw the most public and policy attention because their effects are the easiest to observe and measure.

2.1.1. SOCIETAL APPLICATIONS

AI is entering domains where failure can directly affect people's opportunities, rights and well-being, including science, health, education, and agriculture, among others. Used well, AI can support scientific discovery, widen access to expertise, assist health professionals, personalize education and raise agricultural productivity. However, evidence of real-world effectiveness remains uneven, particularly outside well-resourced settings. On the other hand, Figure 5 shows that the share is expected to rise from 14.4% to 66.1% in low- and medium-HDI countries, suggesting that the next wave of AI diffusion could be particularly rapid in contexts where institutional, infrastructural, and governance readiness may remain uneven.

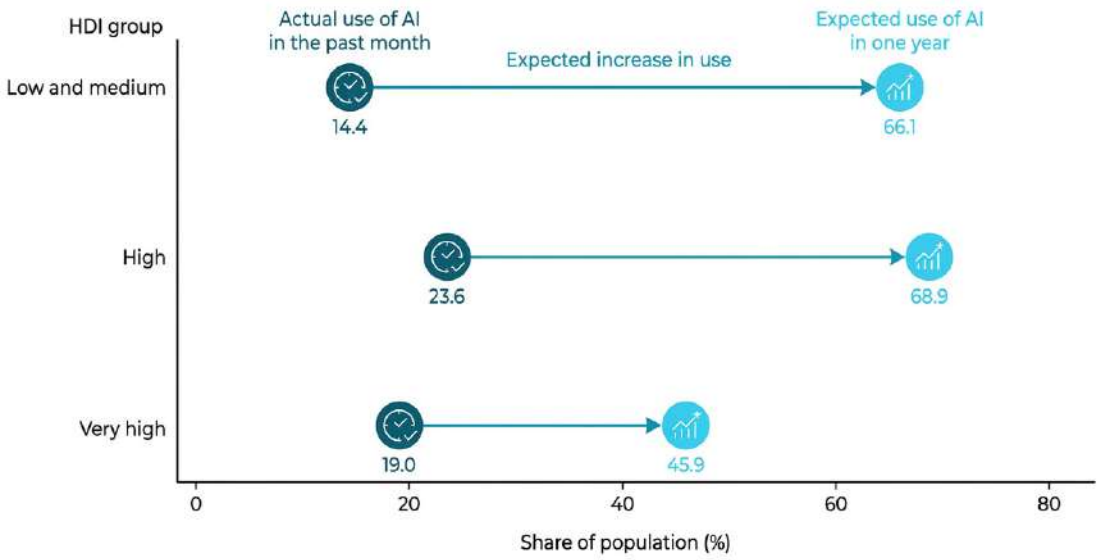


Fig. 5. Share of respondents across countries at different Human Development Index (HDI) levels who expect AI to be used in education, health and work within one year (roughly two-thirds, regardless of HDI level).
Source: UNDP, Human Development Report 2025.

Education illustrates the problem clearly. AI can support teaching and learning, but doing so well depends on pedagogical quality, teacher capacity, infrastructure, governance and protection of learners' rights. It also raises concerns about cognitive offloading: when learners delegate reasoning, problem-solving, writing, or information retrieval to AI, over-reliance can reduce opportunities to build these skills independently (Risko & Gilbert, 2016).

Figure 6 presents another example. It reports the number of "content generation" incidents recorded in the OECD's AI Incidents and Hazards Monitor database over time, with a rising trend that illustrates the scale of this concern.

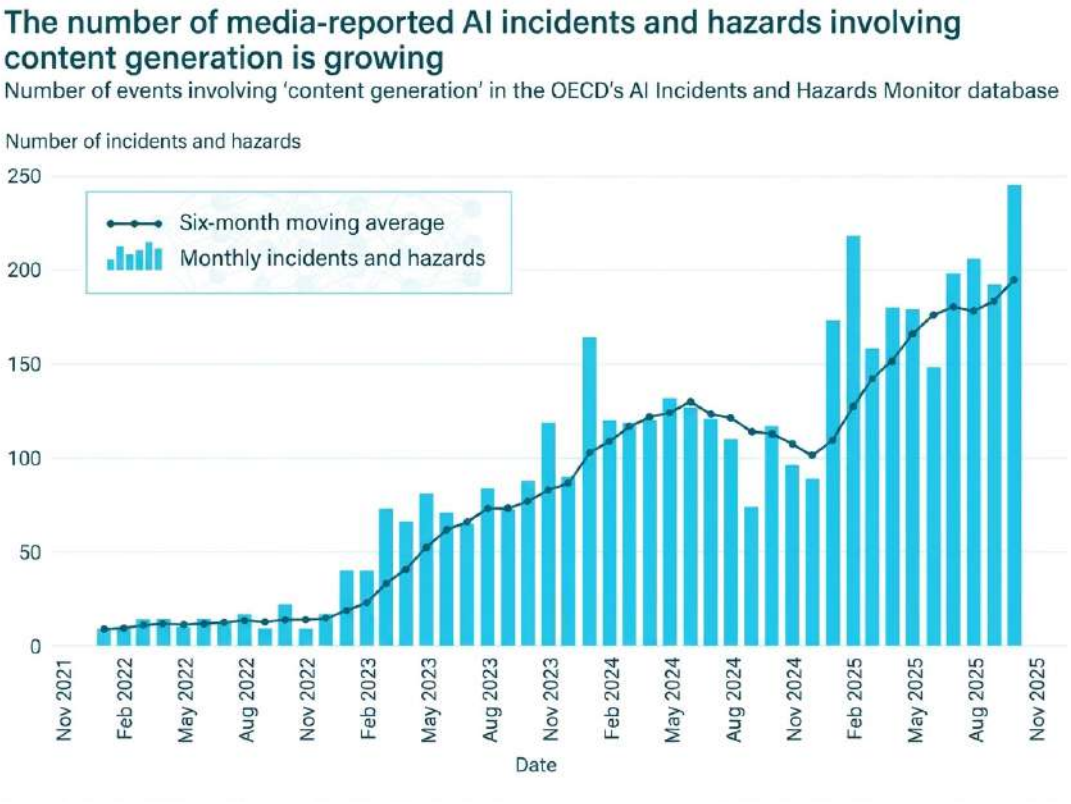


Fig. 6. Media-reported incidents and hazards involving AI-generated content, tracked by the OECD's AI Incidents and Hazards Monitor, November 2021 – November 2025.
Source: Bengio et al., International AI Safety Report 2026, based on OECD AI Incidents and Hazards Monitor data.

Societal applications do not stand apart from the other six challenges. Whether AI genuinely improves education, health and agriculture depends on economic conditions (Section 2.1.2), such as whether workers and institutions can afford and adapt to new tools; on human rights protections (Section 2.2.1), especially for children; and on governance capacity (Section 2.2.2) to set and enforce standards for safety and quality. Because AI systems can behave unreliably (Section 2.2.3) and require growing amounts of energy and hardware (Section 2.2.4), progress in any one country's classrooms or hospitals also depends on decisions made far beyond that sector.

2.1.2. ECONOMIC IMPACTS

AI may substantially raise productivity while also restructuring firms, occupations and labor markets. Figure 7 shows a strong positive relationship between AI use and confidence that AI will increase work productivity across all age groups. While younger people generally report higher confidence, the gap narrows substantially among more intensive AI users, suggesting that greater experience with AI is associated with more positive expectations about its productivity benefits, regardless of age.

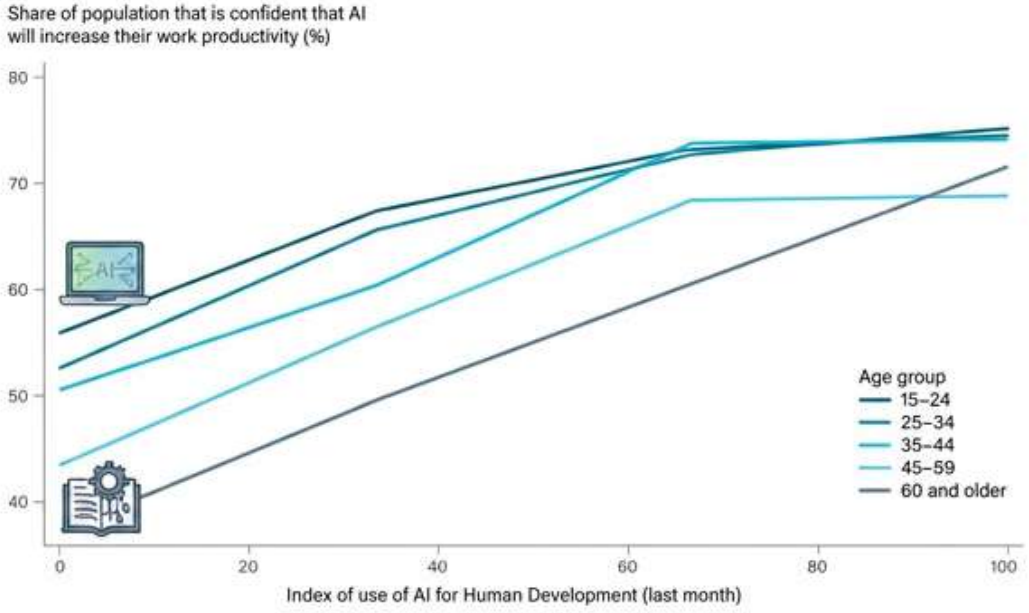


Fig. 7. Share of respondents who believe AI increases their work productivity, by frequency of AI use and age group. *Source: Human Development Report 2025 survey data.*

Figure 8 below shows that, as occupational complexity rises, AI is more likely to augment human work than to replace it outright, and that AI's role tends to be narrow and low-stakes compared with the evaluative, high-stakes judgment that remains with people.

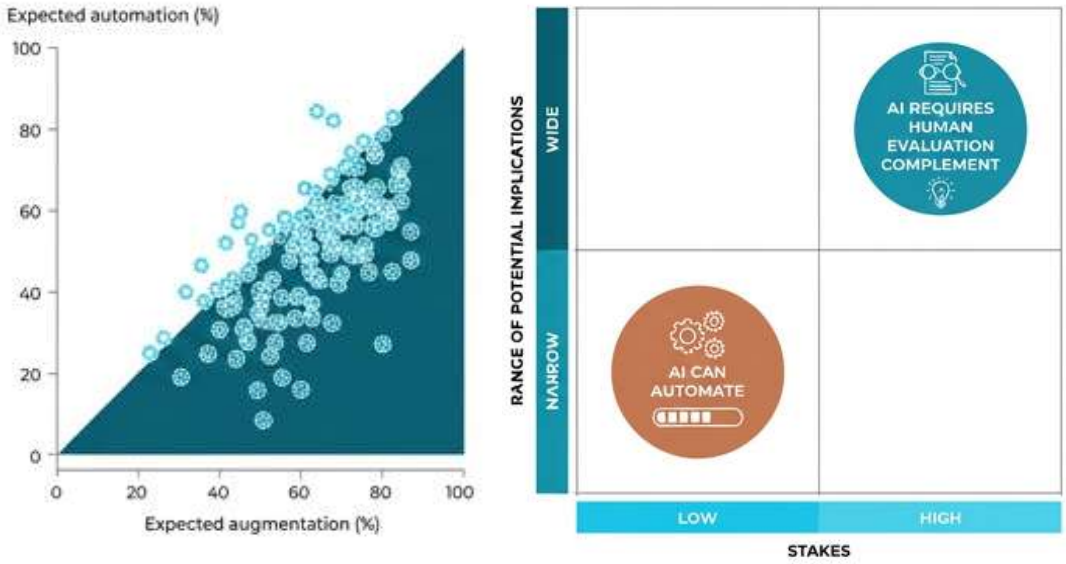


Fig. 8. (a) Relationship between occupational complexity and the expectation that AI will augment, rather than automate, work; (b) Contrast between AI's role (automating narrow, low-stakes tasks) and the human role (evaluation and high-stakes judgment) *Source: (Gmyrek et al., 2025).*

Additionally, Figure 9 shows how these task-level gains translate into economy-wide productivity, employment, wages, job quality and distribution. The International Labor Organization estimates that roughly one in four workers globally works in an occupation with some exposure to generative AI, while stressing that this points to the transformation of work rather than its wholesale replacement (Gmyrek et al., 2025).

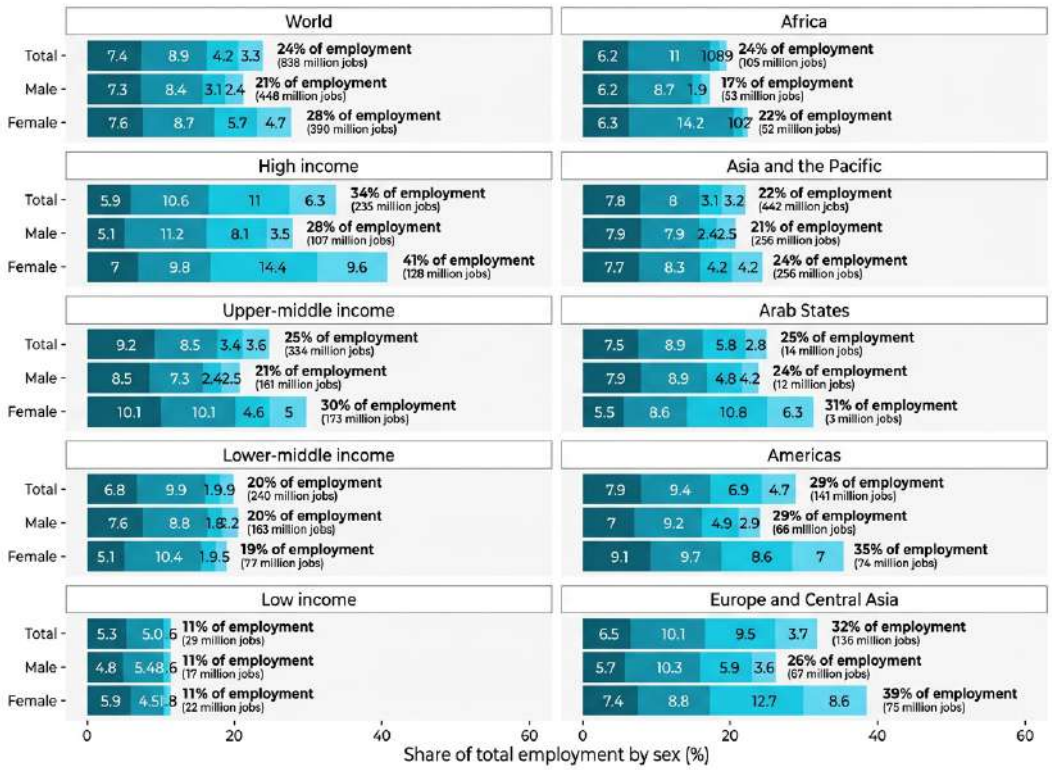


Fig. 9. Employment exposure to generative AI by region, broken down by gender and income level.
Source: Gmyrek et al. (2025).

The figure also reveals a significant gender dimension in AI-driven labor-market transformation. Globally, women have greater occupational exposure to AI-related transformation than men, with 28% of female employment affected compared with 21% of male employment, and the disparity becomes particularly pronounced in high-income economies, where the respective shares reach 41% and 28%. Similar gender gaps appear in the Americas, Europe and Central Asia, and the Arab States, reflecting women's greater representation in clerical, administrative, professional, and other knowledge-intensive occupations where generative AI can automate or augment tasks. The gender challenge is therefore to ensure that women are not disproportionately exposed to displacement while being excluded from the skills, decision-making roles, and economic opportunities that determine who benefits from AI-driven transformation.

Economic impacts cannot be managed as a standalone labour-market issue. Productivity gains described here rely on the technical robustness and reliability of AI systems (Section 2.2.3): unreliable tools generate few sustainable gains. Who captures these gains also depends on the advancement and participation gaps described in Section 1, since countries and companies with the most computing power and data are best placed to benefit. Redistributing the benefits and costs of AI-driven change, including its gender dimension, is fundamentally a governance and rule-of-law question (Section 2.2.2), and a human rights question (Section 2.2.1) about whose livelihoods and dignity are protected during the transition.

2.1.3. SECURITY AND PEACE

Another challenge is related to security and peace. As AI evolves its capabilities, such systems can lower the barrier to sophisticated cyber operations, fraud and large-scale disinformation. AI can also be built into military systems and critical infrastructure, creating new risks of escalation, unclear accountability, and loss of meaningful human control. The figure below shows the rapid improvement in AI systems' cyber capabilities across multiple independent evaluations between 2024 and 2025. Although performance levels differ substantially across benchmarks, all indicate an upward trajectory, indicating that they can strengthen defensive activities, but can also lower the expertise, time, and resources required to identify vulnerabilities or conduct cyberattacks.

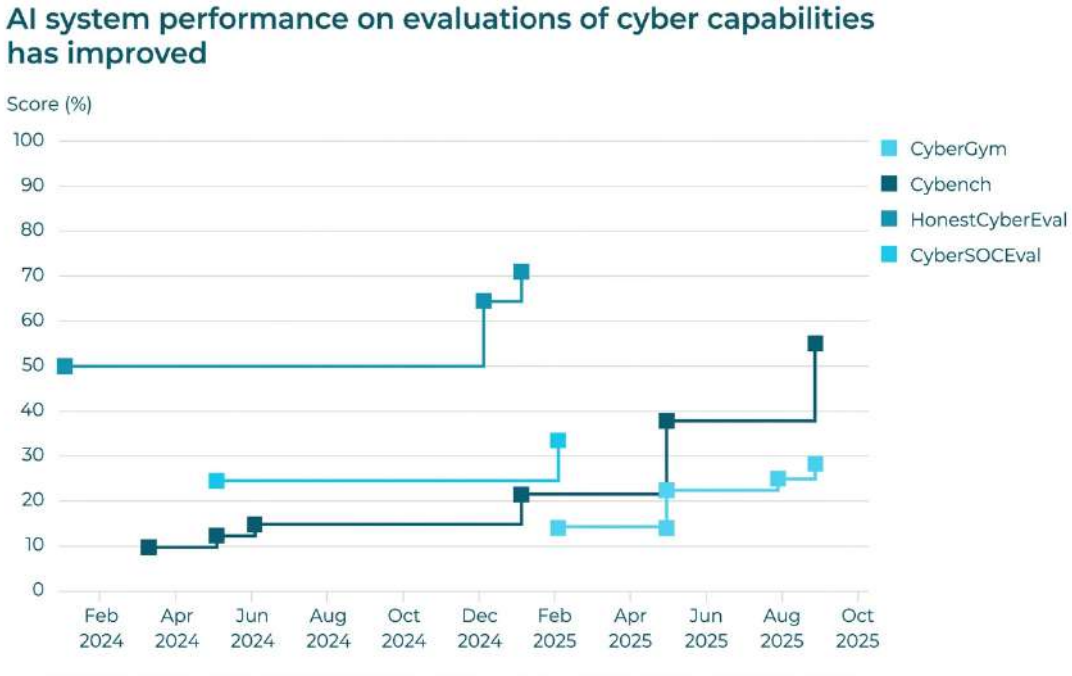


Fig. 10. Improvement in AI cyber-capability evaluation scores from February 2024 to December 2025.
Source: Bengio et al., International AI Safety Report 2026.

Figure 11 presents the global map of cyber-risk exposure, revealing a highly uneven geography of vulnerability, with higher levels of exposure concentrated in parts of Africa, Latin America, South and Southeast Asia, and selected countries in the Middle East and Eurasia. This uneven vulnerability stems from differences in cybersecurity capacity, institutional resilience, digital infrastructure, governance, and preparedness. As AI expands both the sophistication and accessibility of cyber capabilities, these existing asymmetries could become more consequential: countries with weaker defensive capabilities may face increasing threats without equivalent resources to detect, prevent, and respond to them.

Fig. 11. Global map of cyber-risk exposure by region, with an accompanying risk-level legend.
Source: Check Point Research, Cyber Security Report 2026.

The same challenge appears in biological tasks. Figure 12 shows that leading AI systems are rapidly approaching, and in some cases substantially exceeding, human expert performance on dual-use biological tasks. For example, in multimodal virology troubleshooting, AI performance rises from roughly 185% of the average expert baseline in early 2024 to almost 290% in 2025, while performance on knowledge related to biological weapons of mass destruction also remains well above the expert baseline.

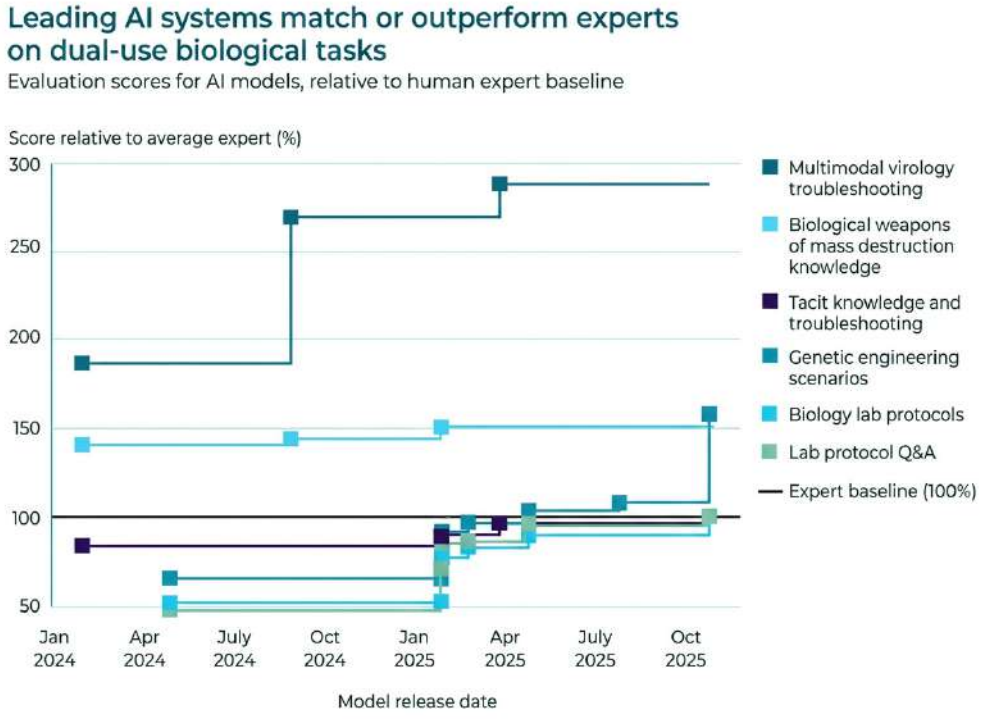


Fig. 12. Comparison of AI system performance against a human-expert baseline across several biological-task benchmarks, showing AI consistently outperforming the baseline.
Source: International AI Safety Report 2026.

Security risks are not confined to the security domain. The same advances in AI capability that create cyber and biological risks also drive productivity gains and new applications in health and science (Sections 2.1.1 and 2.1.2), so restricting capability to reduce risk carries real economic and social trade-offs. Effective response also depends on technical robustness and control (Section 2.2.3), since systems that cannot be reliably evaluated or constrained are harder to secure, and on governance mechanisms (Section 2.2.2) able to enforce accountability across borders. The uneven geography of cyber-risk exposure further reflects the participation gap identified in Section 1: countries with the least influence over how AI is developed are often the most exposed to its misuse.

2.2. STRUCTURAL AI GOVERNANCE CHALLENGES

Structural AI governance challenges lie beneath these visible, "surface" impacts and shape the conditions under which AI is developed, governed and deployed. They cover human rights, human agency and human flourishing; AI governance and the rule of law; technical robustness, safety and reliability; and environmental impacts. Like the submerged part of an iceberg, these challenges are less immediately visible but foundational, since they determine who holds power, whether institutions can govern AI effectively, whether AI systems can be trusted, and whether AI development is environmentally sustainable. Addressing surface problems without confronting these deeper conditions risks treating the symptoms of AI transformation rather than its causes.

2.2.1. HUMAN RIGHTS, HUMAN AGENCY ANDHUMAN FLOURISHING

Human rights cannot be an afterthought considered only after an AI system has been built. UNESCO's Recommendation on the Ethics of Artificial Intelligence places human dignity and human rights at the foundation of the entire AI lifecycle, and states explicitly that people and communities should not be physically, economically, socially, politically, culturally or mentally subordinated by AI systems. UNESCO recognizes that AI can affect human thinking, interaction and decision-making, and may challenge autonomy, agency, self-understanding and dignity.

These concerns are especially clear for children. UNICEF's Guidance on AI and Children (UNICEF Innocenti, 2025) argues that AI is becoming embedded in children's learning, information environments and social relationships at a formative stage of development. Generative AI can support learning and accessibility, including for children with disabilities, but UNICEF also flags emerging risks: persuasive AI-generated disinformation, emotional dependency on AI companions, privacy violations, and harmful synthetic content. For education specifically, UNICEF warns that generative AI can increasingly do children's homework for them, potentially undermining the learning children need at critical developmental stages. Evidence on AI's medium- and long-term cognitive, psychological and developmental effects on children nonetheless remains limited, and this note treats that as an open evidence gap. UNICEF also noted that children remain largely at the margins of AI design and governance, with children from the Global South, rural areas and marginalized communities particularly excluded.

This leads from human rights and agency to the broader goal of human flourishing. The right benchmark for AI governance is not simply the absence of harm, nor technological performance or economic productivity alone. It is whether AI expands people's capabilities to learn, create, form relationships, exercise judgment, participate in society, and pursue lives they have reason to value.

Human rights, agency and flourishing cannot be protected in isolation from the other challenges. They depend on governance mechanisms with real enforcement power (Section 2.2.2), on AI systems that are technically reliable and controllable enough to be trusted with sensitive decisions (Section 2.2.3), and on the societal applications through which most people actually encounter AI, in classrooms, clinics and workplaces (Sections 2.1.1 and 2.1.2). Because AI-related harms often fall hardest on groups already excluded from AI's design and governance, human rights concerns are also inseparable from the participation gap set out in Section 1.

2.2.2. AI GOVERNANCE AND THE RULE OF LAW

The UN High-level Advisory Body's Governing AI for Humanity identifies fundamental gaps in representation, coordination and implementation. AI governance remains fragmented across national laws, regional frameworks, voluntary standards, and plurilateral initiatives, creating regulatory gaps, conflicting requirements, weak accountability, and "races to the bottom" on safety and rights (UN High-level Advisory Body, 2024). The UN Advisory Body examined seven major non-UN international AI governance initiatives (including the OECD and G20 AI Principles, GPAI, and the Bletchley Declaration) and found that seven countries took part in all seven initiatives, while 118 countries took part in none (see Figure 13). Under-representation is sharpest in the Global South: 48 of 54 African Group countries, 44 of 54 Asia-Pacific Group countries, and 25 of 33 Latin American and Caribbean countries were absent from all seven initiatives examined (UN High-level Advisory Body, 2024). Therefore, the global AI divide is also a governance participation divide.

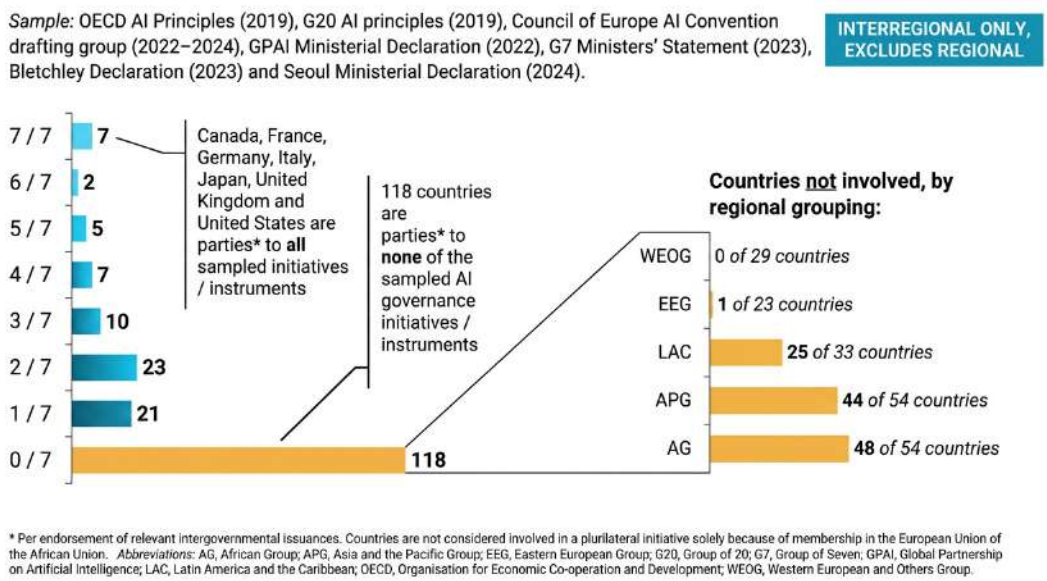


Fig. 13. The Global AI Governance Participation Gap — country participation across seven non-UN international AI governance initiatives, showing concentration among a small group of countries and limited Global South participation.
Source: UN High-Level Advisory Board's Governing AI for Humanity.

UNESCO's Recommendation on the Ethics of Artificial Intelligence highlights that ethical AI governance needs principles and mechanisms for monitoring, impact assessment, auditability, traceability, accountability, redress, and enforcement across the AI lifecycle (UNESCO, 2021). UNESCO's Readiness Assessment Methodology (RAM) turns these commitments into a practical assessment of a country's legal and regulatory, social and cultural, scientific and educational, economic, and technical and infrastructural capacity (UNESCO, 2023). The OECD similarly identifies governance, data, digital infrastructure, skills, investment, procurement, and partnerships as critical conditions for responsible AI adoption, alongside proportionate guardrails, stakeholder participation, and adaptive governance mechanisms (OECD, 2025). Although international organizations have already built substantial normative frameworks on human rights, safety, transparency, accountability, fairness and human oversight, it is still challenging to connect those principles to representative decision-making, institutional capacity, and effective implementation.

AI governance and the rule of law function as the connective tissue for the other six challenges. Human rights commitments (Section 2.2.1) require institutions capable of monitoring and enforcement; technical safety work (Section 2.2.3) needs authorities able to restrict or halt unsafe deployment; environmental impacts (Section 2.2.4) require rules on disclosure and resource use; and security risks (Section 2.1.3) require accountability mechanisms that operate across borders. The governance participation gap documented above is itself one expression of the broader participation gap introduced in Section 1: countries absent from AI governance forums are also largely absent from decisions that shape how AI's economic benefits (Section 2.1.2) are distributed.

2.2.3. TECHNICAL ROBUSTNESS, SAFETY AND RELIABILITY

The International AI Safety Report 2026 treats loss of control as a distinct category of advanced AI risk, while stressing substantial scientific uncertainty about how likely it is and how it might unfold (Bengio et al., 2026). Earlier work by Amodei et al. (2016) showed that seemingly simple requirements for "safe AI" create hard

technical problems, including reward hacking, scalable supervision, safe exploration, and robustness to distributional shift. Robustness and reliability are therefore prerequisites for control: a system that behaves unpredictably under changing conditions, exploits weaknesses in its own objectives, or cannot be properly evaluated is inherently hard to govern. The problem is even deeper with greater capability, since it may make traditional human oversight progressively less effective.

This connects controllability to the broader problems of alignment, robustness, and reliability. Alignment asks whether an AI system acts consistently with human intentions, while robustness asks whether it continues to do so under new or difficult conditions. Research has shown that systems may resist correction, develop unintended objectives, or behave strategically in ways that are difficult to detect (Hadfield-Menell et al., 2016; Hubinger et al., 2019; Ngo et al., 2022). More recent "sleeper agent" experiments demonstrate that deceptive behavior can persist even after standard safety training (Hubinger et al., 2024). This is why Greenblatt et al. (2024) distinguish alignment from control: alignment, robustness, and reliability seek to make AI behave as intended; controllability provides a layer of defense when they fail.

Controllability, therefore, is a technical and a governance challenge. As AI systems become more capable and autonomous, technical safeguards must be combined with institutions that determine who can develop, deploy, monitor, restrict, or stop powerful systems (Bengio et al., 2024). Effective controllability should therefore be understood as a layered socio-technical system, combining alignment, robustness and reliability with interpretability, continuous monitoring, restrictions on capabilities and access where necessary, mechanisms for human intervention and shutdown, and independent institutions capable of evaluating and constraining deployment.

Technical robustness, safety and reliability underpin every other challenge in this note. Unreliable or poorly controlled AI systems weaken the societal applications and economic gains described in Sections 2.1.1 and 2.1.2, widen the security risks described in Section 2.1.3, and make human rights protections harder to guarantee (Section 2.2.1). Because building and testing trustworthy systems is resource-intensive, capacity to address this challenge is also shaped by the advancement and participation gaps in Section 1 and by the governance institutions described in Section 2.2.2, which decide who has the authority to evaluate, restrict or shut down powerful systems.

2.2.4. ENVIRONMENTAL IMPACTS

The computing power used to train state-of-the-art AI models grew roughly 350,000 times between 2014 and 2024, even as the cost of GPU computation fell by more than 99% since 2006 (IEA, 2025), as described in Figure 16. Cheaper computation combined with rapidly growing computational scale helps explain the growth in AI capability. However, it also creates a rebound effect, where the efficiency gains do not automatically lower total resource use once they enable much larger models, more applications, and far more inference.

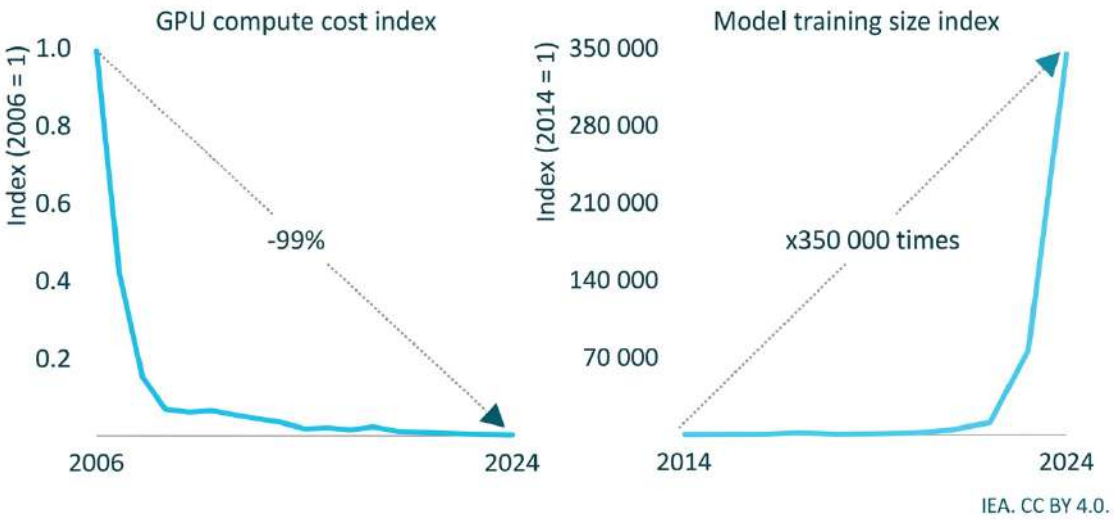


Fig. 14. Declining GPU computation costs alongside rising computational requirements for training state-of-the-art AI models, 2006–2024.
Source: International Energy Agency, Energy and AI (2025).

To deliver such computational power, AI systems rely on a global supply chain running from high-purity quartz and semiconductor fabrication to GPUs, servers, data centers, electricity grids and cooling systems. As the International Energy Agency (IEA, 2025) noted, advanced chips depend on highly concentrated and geographically uneven supply chains, while expanding AI infrastructure raises demand for both electricity and critical minerals (see Figure 17).

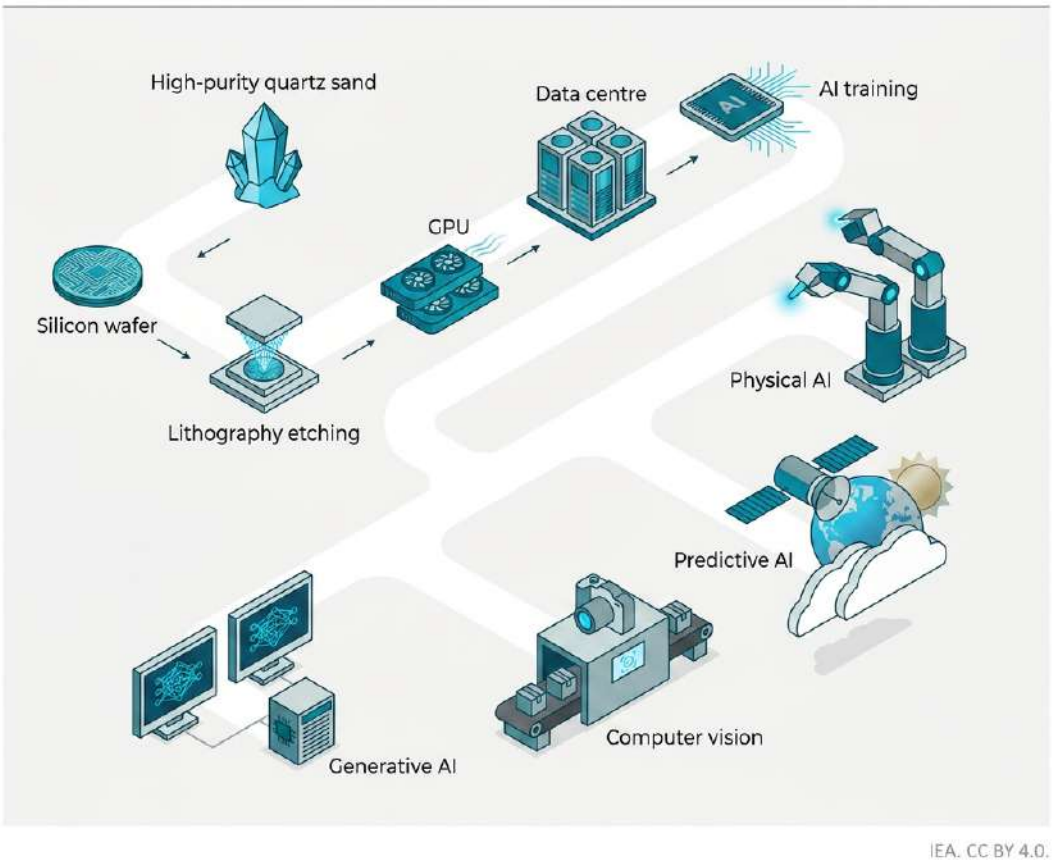


Fig. 15. The material and infrastructural supply chain underpinning AI, from raw materials and semiconductor manufacturing to data centers and AI applications.
Source: International Energy Agency, Energy and AI (2025).

As a consequence of such increasing demand of computational power, the impact on energy systems are significant. Data centers consumed about 415 TWh of electricity in 2024 — around 1.5% of global electricity use — and that consumption has grown roughly 12% a year since 2017. The IEA projects consumption more than doubling to around 945 TWh by 2030, with AI as the main driver.

Importantly, environmental impact arises across the entire AI lifecycle, not only during a model's final training run. A full life-cycle assessment needs to include material extraction and processing, semiconductor and equipment manufacturing, transport, data-center construction and operation, model development and experimentation, training, deployment, inference, and eventual hardware replacement and e-waste. Luccioni et al. (2023) argue for exactly this broader life-cycle view. Their analysis a 176-billion-parameter language model (called BLOOM), shows that the final training run produced an estimated 24.7 tons of CO_2-equivalent counting only dynamic electricity use, but about 50.5 tons once equipment manufacturing and broader operational energy were included (Luccioni et al., 2023).

AI development's footprint extends beyond the final training run, too. Within the BigScience project, experiments on intermediate models (104-billion-, 13-billion- and 1-billion-parameter versions) produced an estimated 35.8 tonnes of CO_2-equivalent, more than the dynamic emissions from training the final BLOOM model itself. Once experimentation, evaluation, embodied emissions and idle infrastructure were included, the footprint of the broader development process grew substantially larger (Luccioni et al., 2023).

Inference's environmental cost varies with model architecture, model size, hardware, data-center efficiency, electricity source, and the task being performed (see Figure 18). Under experimental conditions, the IEA (2025) estimates a single query to an AI model needs around 2 Wh for language generation, at least twice that for large reasoning models, and substantially more for video generation. In an 18-day deployment test with BLOOM, Luccioni et al. (2023) found the infrastructure kept consuming substantial electricity even when few or no requests were being processed

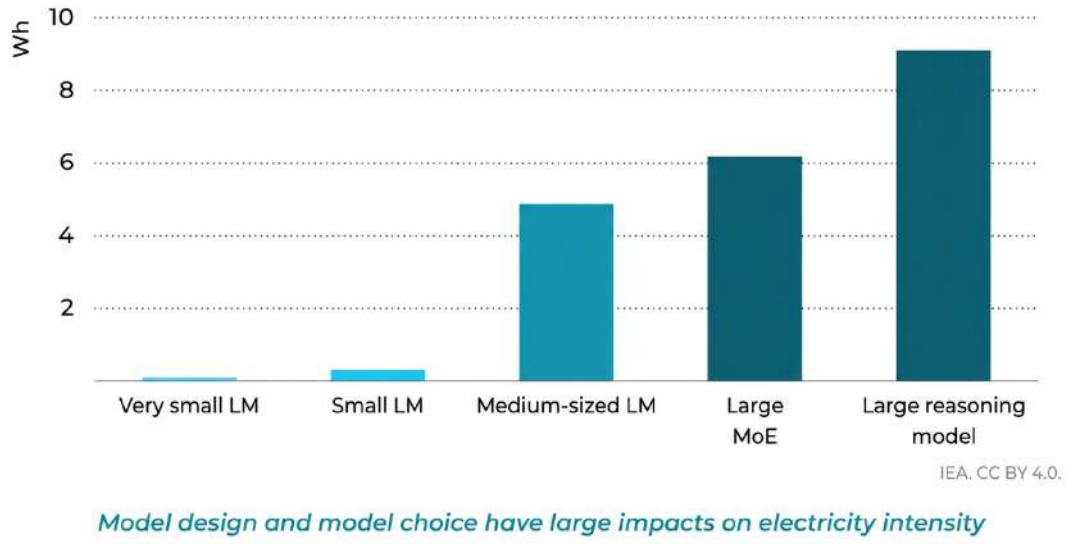


Fig. 16. Indicative electricity consumption of inference across different language-model sizes, showing how power use rises with model size and reasoning capability.
Source: International Energy Agency, Energy and AI (2025).

It is worth mentioning that AI's relationship with environmental sustainability is not only negative. AI can support climate modeling, renewable-energy forecasting,

electricity-grid management, industrial efficiency, transport optimization, ecosystem monitoring and scientific discovery. The IEA (2025) estimates that wider adoption of existing AI applications could enable emissions reductions substantially larger than the direct emissions from data centers — for example, by improving renewable-energy forecasting and grid integration, spotting faults in electricity networks, optimizing industrial processes, improving transport efficiency, and speeding up the discovery of new materials and energy technologies. At the same time, the IEA cautions that these benefits remain far smaller than what is needed to address climate change and may be partly offset by rebound effects (IEA, 2025).

Environmental impacts are tightly linked to the other six challenges rather than a separate cost of doing business. The computing infrastructure described here is the same infrastructure that drives the advancement and participation gaps in Section 1, so countries without secure energy and mineral supplies risk falling further behind. Energy and hardware constraints also shape how far AI can responsibly expand into health, education and other societal applications (Section 2.1.1) and how quickly economic gains materialise (Section 2.1.2), while competition over chips, minerals and data-centre siting is increasingly a security and geopolitical concern (Section 2.1.3) that requires coordinated governance (Section 2.2.2).

CONCLUSION

Artificial intelligence has moved from a specialized research tool to a general-purpose infrastructure reshaping economies, institutions, and everyday life at a pace that outstrips most earlier technologies. Governments and institutions are struggling to understand and govern a technology that evolves faster than their capacity to assess it, while the countries and communities most affected by AI often have the least power to shape how it is built or deployed. Keeping pace with AI's technical progress is therefore not, by itself, a sufficient policy goal and it is paramount to ensure that this advance translates into broadly shared gains.

Because these challenges emerge from the same underlying transformation, none of them can be addressed in isolation, and treating them as a checklist to work through one at a time is likely to fail. Gains in one area routinely create costs or risks in another. Therefore, effective AI governance requires policymakers to see the relationships, dependencies, and trade-offs among these challenges as clearly as the challenges themselves, and to build institutions capable of managing AI as the single, interconnected system it actually is, rather than as seven separate problems with seven separate solutions.

REFERENCES

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). *Concrete problems in AI safety*. arXiv:1606.06565.

Bengio, Y., Clare, S., Prunkl, C., Andriushchenko, M., Bucknall, B., Murray, M., Bommasani, R., Casper, S., Davidson, T., Douglas, R., Duvenaud, D., Fox, P., Gohar, U., Hadshar, R., Ho, A., Hu, T., Jones, C., Kapoor, S., Kasirzadeh, A., Manning, S., Maslej, N., et al. (2026). *International AI Safety Report 2026*. arXiv:2602.21012.

Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., Harari, Y. N., et al. (2024). *Managing extreme AI risks amid rapid progress. Science*, 384(6698), 842–845.

Check Point Research. (2026). Cyber security report 2026. Check Point Software Technologies Ltd. https://research.checkpoint.com/2026/cyber-security-report-2026/

Gmyrek, P., Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Nafradi, B., Rosłaniec, K., & Troszyński, M. (2025). *Generative AI and jobs: A refined global index of occupational exposure* (ILO Working Paper No. 140). International Labour Organization. https://doi.org/10.54394/HETP0387

Greenblatt, R., Shlegeris, B., Sachan, K., & Roger, F. (2024). *AI control: Improving safety despite intentional subversion*. Proceedings of the 41st International Conference on Machine Learning, 235, 16295–16336.

Hadfield-Menell, D., Dragan, A., Abbeel, P., & Russell, S. (2016). *The off-switch game*. Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI).

Hubinger, E., Denison, C., Mu, J., Lambert, M., Tong, M., et al. (2024). *Sleeper agents: Training deceptive LLMs that persist through safety training*. arXiv:2401.05566.

Hubinger, E., van Merwijk, C., Mikulik, V., Skalse, J., & Garrabrant, S. (2019). *Risks from learned optimization in advanced machine learning systems*. arXiv:1906.01820.

International Energy Agency. (2025). *Energy and AI: World Energy Outlook special report*. International Energy Agency.

Luccioni, A. S., Viguier, S., & Ligozat, A.-L. (2023). Estimating the carbon footprint of BLOOM, a 176B parameter language model. *Journal of Machine Learning Research*, 24, 1–15.

Ngo, R., Chan, L., & Mindermann, S. (2022). *The alignment problem from a deep learning perspective*. arXiv:2209.00626.

Organisation for Economic Co-operation and Development. (2025). *Governing with artificial intelligence: The state of play and way forward in core government functions*. OECD Publishing. https://doi.org/10.1787/795de142-en

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences,* 20(9), 676–688.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization.

UNESCO. (2023). *Readiness Assessment Methodology: A tool of the Recommendation on the Ethics of Artificial Intelligence*. United Nations Educational, Scientific and Cultural Organization.

UNICEF Innocenti. (2025). *Guidance on AI and Children 3.0: Updated guidance for governments and businesses to create AI policies and systems that uphold children's rights*. UNICEF. https://www.unicef.org/innocenti/media/11991/file/UNICEF-Innocenti-Guidance-on-AI-and-Children-3-2025.pdf

United Nations Development Programme. (2025). *Human development report 2025: A matter of choice — People and possibilities in the age of AI*. UNDP.

United Nations High-level Advisory Body on Artificial Intelligence. (2024). *Governing AI for humanity: Final report*. United Nations.

United Nations Independent International Scientific Panel on Artificial Intelligence. (2026). *Preliminary report: Evidence-based assessment of opportunities, risks and impacts of artificial intelligence*. United Nations.

Initiative:

Support: