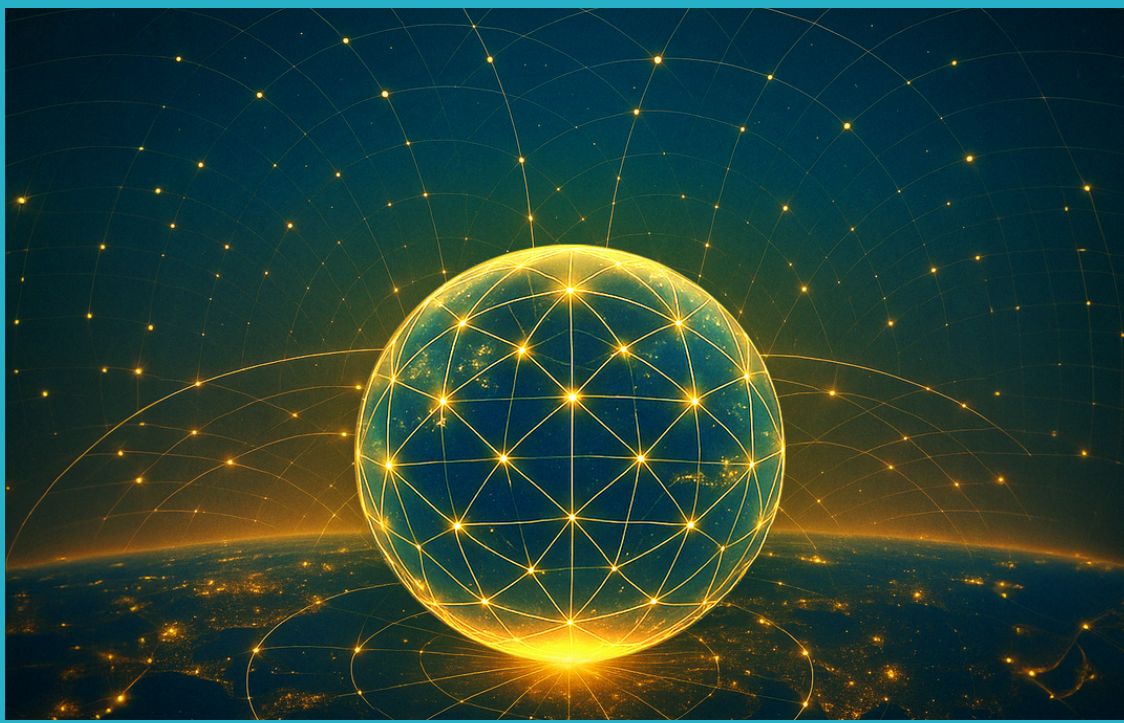
WHITEPAPER
December 2025

AI BENEFIT-SHARING FRAMEWORK: BALANCING ACCESS AND SAFETY



Authors: Sumaya Nur Adan, Joanna Wiaterek, Varun Sen Bahl, Ima Bello, Camila Beltran, Tobias Dierks, Luise Eder, Liam Epstein, Abra Ganz, Natalie Kiilu, Marianne Lu, Chinasa T. Okolo, Sydney Reis, Said Saillant, Krishna Sharma, Marjia Siddik, Jan Pieter Snoeij, José Jaime Villalobos, Anna Yelizarova

AI Benefit-Sharing Framework: Balancing Access and Safety

Sumaya Nur Adan[1 2 3], Joanna Wiaterek[1]

Varun Sen Bahl[4], Ima Bello[5], Camila Beltran[6], Tobias Dierks[7], Luise Eder[3], Liam Epstein[8 9], Abra Ganz[10], Natalie Kiilu[1 8], Marianne Lu, Chinasa T. Okolo[11], Sydney Reis[2], Said Saillant[12], Krishna Sharma[13], Marjia Siddik[14], Jan Pieter Snoeij, José Jaime Villalobos[5], Anna Yelizarova[15]

[1]Centre for AI Security and Access, [2]University of Oxford, [3]AI Governance Initiative, [4]Duke University, [5]Future of Life Institute, [6]Women in AI Colombia, [7]Office of the German Council for Sustainable Development, [8]University of Cambridge, [9]Convergence Analysis, [10]Center for AI Risk Management and Alignment, [11]Technēcultură, [12]Societas Sapiens, Inc, [13]Hoover Institution, [14]Dublin City University, [15]Windfall Trust

Acknowledgements

We would like to recognise all the experts who contributed to this whitepaper in various ways:

Expert working group who provided feedback on the direction and the key questions of this report: Aubra Anthony, Duncan Cass-Beggs, Matthijs Maas, Chinasa T. Okolo, Miro Pluckebaum, Saad Siddiqui, Jan Pieters Snoeij, José Jaime Villalobos, Anna Yelizarova.

Workshop participants who contributed to scenario deliberations and provided feedback on this report: Varun Sen Bahl, Ima Bello, Camila Beltran, Leticia Caminero, Deric Cheng, Tobias Dierks, Luise Eder, Liam Epstein, Abra Ganz, Haakon Huynh, Iman Khwaja, Marianne Lu, Sydney Reis, Said Sailant, Krishna Sharma, Marjia Siddik, Wai Kit Si Tou, Krupa Suchak, Philip Tomei.

Special thanks for invaluable conversations and insights: Bunmi Ajala, Sam Carter, Alex Cobham, Iason Gabriel, Timothy Hodges, Anton Korinek, Sam Manning, Kent Nnadozie, Nidhi Singh, Marina Mendes Tavares, Bosun Tijani, Robert Trager.

Disclaimer: Co-authorship or acknowledgment does not necessarily constitute endorsement of all views expressed in this paper. Names of co-authors and contributors appear in alphabetical order.

Cover page image: Victor Farias.

This report draws on expert consultations conducted between May and November 2025, supported by a working group that provided guidance on the research direction. During this period, we engaged stakeholders across multiple regions – including Africa, Latin America, Southeast Asia, and representatives from governments, industry, international organisations, and academia. Discussions examined what has and has not worked in existing benefit-sharing regimes, the needs of different actors, and priority areas for advancing benefit sharing in the context of AI. Additional insights emerged from ongoing research collaborations and paper discussions. The analysis was further informed by a multi-stakeholder workshop convened on 5 September 2025 at the Oxford Martin School, AI Governance Initiative.

The workshop and the report have been possible thanks to the financial support from the Future of Life Institute.

Contact: info@casa-ai.org
Cite as: Adan, Wiaterek et al. (2025) "AI Benefit-Sharing Framework: Balancing Access and Safety."

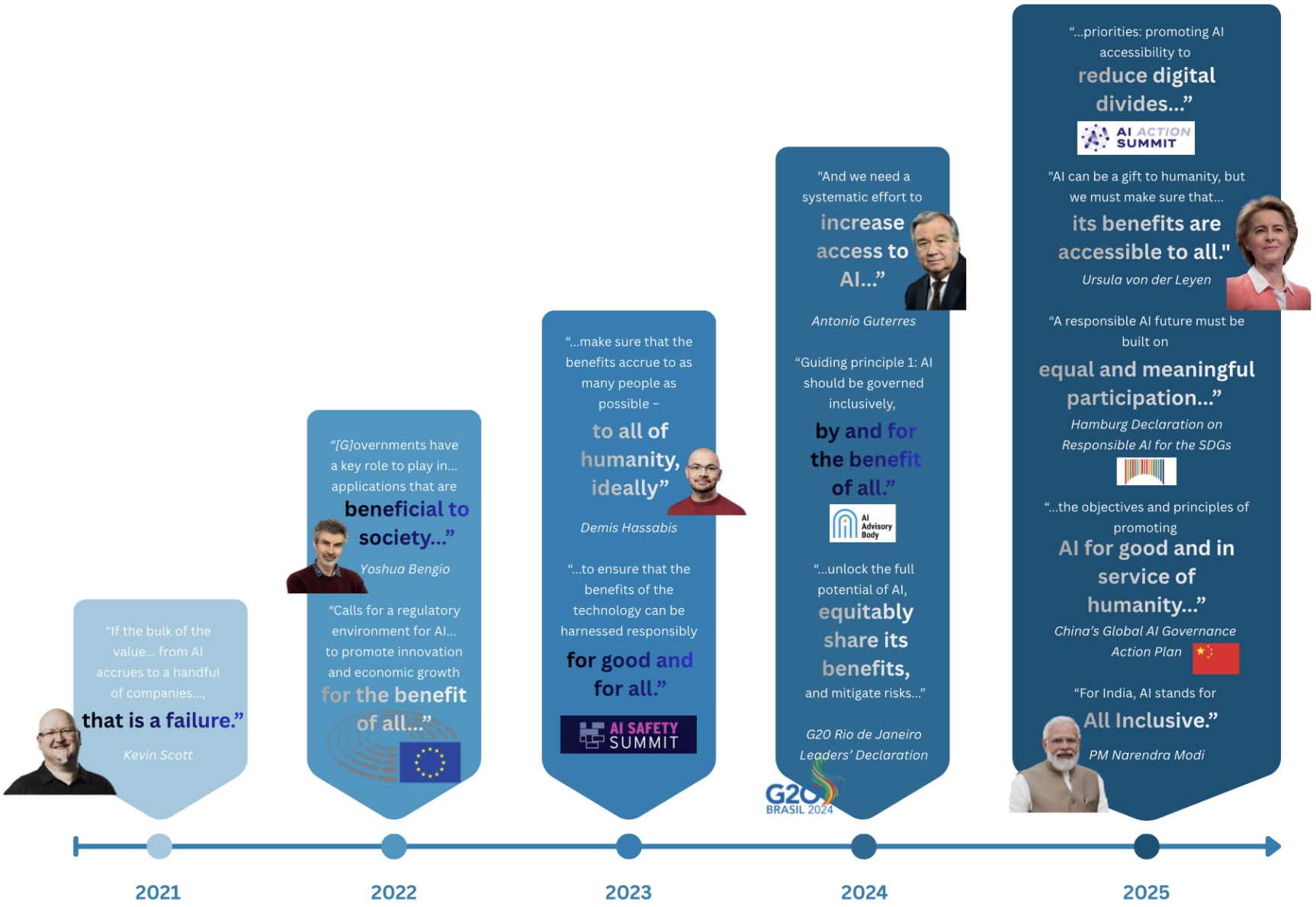


Image 1: Timeline of statements on equitable AI governance from AI experts and policy-makers.

"For India, AI stands for All Inclusive." – *PM Narendra Modi (2025)*[1]

"(...) we hereby put forward the Global AI Governance Action Plan, calling on all parties to take concrete and effective actions in advancing global AI development and governance based on the objectives and principles of promoting AI for good and in service of humanity, respecting national sovereignty, aligning with development goals, ensuring safety and controllability, upholding fairness and inclusiveness, and fostering open cooperation." – *China's Global AI Governance Action Plan (2025)*[2]

"A responsible AI future must be built on equal and meaningful participation, with actions to ensure that all stakeholders, especially those from emerging markets, developing economies, and vulnerable groups, have fair and equitable access to, as well as ownership of, computing, data, investment, and resources for capacity and talent development. AI's benefits must not remain concentrated among a privileged few." – *Hamburg Declaration on Responsible Artificial Intelligence for the Sustainable Development Goals (2025)*[3]

"AI can be a gift to humanity, but we must make sure that its benefits are widespread and that its benefits are accessible to all." – *Ursula von der Leyen, European Commission (Paris AI Action Summit, 2025)*[4]

"(...) we have affirmed the following main priorities: promoting AI accessibility to reduce digital divides; ensuring AI is open, inclusive, transparent, ethical, safe, secure and trustworthy, taking into account international frameworks for all (...)" – *Statement on Inclusive and Sustainable Artificial Intelligence for People and the Planet (Paris AI Action Summit, 2025)*[5]

"To unlock the full potential of AI, equitably share its benefits, and mitigate risks, we will work together to promote international cooperation and further discussions on international governance for AI, recognizing the need to incorporate the voices of developed and developing countries." – *G20 Rio de Janeiro Leaders' Declaration (2024)*[6]

---

[1] See: https://economictimes.indiatimes.com/tech/artificial-intelligence/for-india-ai-stands-for-all-inclusive-says-pm-narendra-modi/articleshow/124425100.cms?from=mdr
[2] See: https://www.fmprc.gov.cn/mfa_eng/xw/zyxw/202507/t20250729_11679232.html
[3] See: https://www.bmz-digital.global/wp-content/uploads/2025/06/250603_Hamburg_Declaration.pdf
[4] See: https://www.reuters.com/technology/artificial-intelligence/quotes-eu-chief-von-der-leyens-ai-speech-paris-summit-2025-02-11/
[5] See: https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet
[6] See: https://g20.org/wp-content/uploads/2024/11/G20-Rio-de-Janeiro-Leaders-Declaration-EN.pdf

“Guiding principle 1: AI should be governed inclusively, by and for the benefit of all.” – *Governing AI for Humanity (United Nations, 2024)*[7]

"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." *– Antonio Guterres, United Nations (2024)*[8]

“In recognition of the transformative positive potential of AI, and as part of ensuring wider international cooperation on AI, we resolve to sustain an inclusive global dialogue that (…) contributes in an open manner to broader international discussions, and to continue research on frontier AI safety to ensure that the benefits of the technology can be harnessed responsibly for good and for all.” *– The Bletchley Declaration (2023)*[9]

“I think we need to make sure that the benefits accrue to as many people as possible – to all of humanity, ideally.” *– Demis Hassabis, Google DeepMind (2023)*[10]

“Calls for a regulatory environment for AI that provides effective governance and protection of fundamental rights, while facilitating competitive access to digital markets for actors of all size to promote innovation and economic growth for the benefit of all; underlines that a competitive, accessible and fair data economy, based on common standards, is a prerequisite for the adequate development and training of AI (...)” *– European Parliament Resolution of 3 May 2022 on Artificial Intelligence in a Digital Age (2022)*[11]

“(...) applications of AI that are clearly beneficial to society should be encouraged, whether it be in health, in the fight against climate change, against injustice or in increasing access to knowledge and education. In all these areas, governments have a key role to play in directing the forces of AI research and entrepreneurship towards those applications that are beneficial to society but where the desire to make a profit is not always sufficient to stimulate the needed investments.” *– Yoshua Bengio, MILA (2022)*[12]

“I think we should have objectives around real democratisation of the technology. If the bulk of the value that gets created from AI accrues to a handful of companies in the West Coast of the United States, that is a failure.” *– Kevin Scott, Microsoft (2021)*[13]

---

[7] See: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf
[8] See: https://www.weforum.org/stories/2024/01/what-leaders-said-about-ai-at-davos-2024/
[9] See: https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023
[10] See: https://time.com/6246119/demis-hassabis-deepmind-interview/
[11] See: https://www.europarl.europa.eu/doceo/document/TA-9-2022-0140_EN.html
[12] See: https://yoshuabengio.org/2022/01/24/superintelligence-futurology-vs-science/
[13] See: https://exchange.scale.com/public/videos/democratizing-and-accelerating-the-future-of-ai-with-kevin-scott

Table of Contents

Abstract

Artificial intelligence offers unprecedented opportunities for innovation and economic growth. However, current development trajectories risk excluding billions, especially those in Global Majority countries, from meaningfully accessing these benefits. The existing AI governance ecosystem lacks systematic mechanisms to ensure equitable global AI diffusion while also managing risks. To fill this gap, this paper proposes a framework for AI benefit-sharing – the fair distribution of and access to AI's opportunities and gains under conditions of safety. It integrates three pillars: redistribution of economic gains, technology transfer and capacity-building, and non-proliferation of dangerous capabilities. We address factors hindering equitable AI governance, clarify the conceptual foundations of benefit-sharing, and outline prerequisites for readiness among Global Majority states. Combining the analytical and operational dimensions, this framework maps out specific AI benefit-sharing mechanisms, offering a viable toolkit for building a more equitable future.

Executive Summary

Introduction

As artificial intelligence (AI) races ahead, measured in months rather than decades, the world is witnessing a technological transformation unfolding faster than any in human history. However, while offering unprecedented opportunities for innovation and economic growth, the current development and governance trajectories risk excluding billions, especially those in Global Majority countries, from meaningfully accessing these benefits. Despite heightened rhetoric around fair global AI governance (*see Image 1*), the existing structures lack systematic mechanisms to ensure equitable AI diffusion while also managing risks. Therefore, this paper seeks to bridge that gap by proposing a comprehensive framework for AI benefit-sharing – the fair distribution of and access to AI's opportunities and gains under conditions of safety. Specifically, the paper's contribution lies in:

1. **Addressing the prevailing counter-arguments** which sideline the operationalisation of benefit-sharing and confine it to political rhetoric or mere redistributive measures.
2. **Situating benefit-sharing** within the related AI access governance terminology and identifying the specific tensions it must navigate.
3. **Outlining prerequisites for implementation of AI benefit-sharing and mapping out its mechanisms** across three dimensions: redistribution, technology transfer and capacity-building, and non-proliferation and safety.

In short, the paper defines and grounds AI benefit-sharing, addresses the barriers to its operationalisation, and offers a viable toolkit for moving from principle to practice.

Part I: Situating AI Benefit-Sharing

Part I situates the central argument of this paper that AI benefit-sharing requires integration across three distinct governance traditions: redistribution of economic gains, technology transfer and capacity-building, and non-proliferation and safety controls, by clarifying both the flawed arguments that sideline the issue and the conceptual terrain on which it must be redefined.

Chapter 1 interrogates three prevailing assumptions around the diffusion of AI's opportunities and gains that hinder serious consideration of benefit-sharing mechanisms:

Counter-Arguments to AI Benefit-Sharing	
Argument Against AI Benefit-Sharing	**Response Summary**
The Market Myth: diffusion of AI benefits will occur naturally through market forces	1. Market incentives alone often fall short. 2. Even where market failures are recognized and addressed, traditional intervention mechanisms may be too slow for the AI transition timeline. 3. Existing redistribution mechanisms, especially relevant for market failures, are inadequate to address global AI equity.
The U.S. AI approach: strategic restraint as the default, limiting pathways for meaningful global diffusion	1. The domestic foundation of the U.S. AI leadership is under strain. 2. Preserving U.S. leadership requires more than hoarding compute – it requires shaping the global distribution of AI through partnerships. 3. The economic logic of commercial diplomacy supports a more outward-facing approach.
U.S.–China AI geopolitical tensions overshadow international cooperation on AI diffusion	1. AI competition is not zero-sum in all dimensions. 2. Specific areas of convergent interest exist despite rivalry. 3. The costs of non-cooperation are rising.

Table 1: Summary of the common counter-arguments sidelining AI benefit-sharing and the paper's responses to them.

Chapter 2 begins by examining the existing terminology related to global access to AI: benefit-sharing, technology transfer, diffusion, and non-proliferation. While these terms are often used interchangeably, each of them carries distinct historical origins and normative implications, and their adaptation to AI reveals important shifts in how the technology is being politically framed:

- **Benefit-sharing** originates in environmental governance where it seeks fair distribution of profits derived from the use of genetic resources. In the context of AI, it is used to describe the equitable allocation of economic, social, and technological gains or benefits resulting from AI development and deployment. This includes, for example, access to AI-enabled services, financial returns, capacity-building opportunities, and decision-making influence. The use of benefit-sharing signals a normative view of AI as a global public good.

- **Technology transfer** in AI adapts an international trade and development concept of deliberately moving knowledge and tools from advanced to less-resourced actors. Here, it refers to the sharing of models, data, and compute while framing AI as a strategic asset whose access enables participation but remains constrained by competitiveness and intellectual-property concerns.
- **Diffusion** refers to the innovation-studies notion of technologies spreading through adoption and integration. In the context of AI, it is used descriptively to track global AI deployment, however, this seemingly neutral framing risks obscuring the structural barriers that limit equitable spread and wrongly implying that benefits will diffuse automatically.
- **Non-proliferation** applies an arms-control logic to AI by advocating restrictions on access to high-risk models, compute, and techniques. It frames AI as a potentially destabilizing technology whose tight control can conflict with more inclusive benefit-sharing aims.

The chapter then addresses the underlying tensions that fundamentally shape the feasibility and design of AI benefit-sharing mechanisms and that require strategic navigation:

Underlying Tensions of AI Benefit-Sharing	
Speed vs. Deliberation	AI development is advancing at unprecedented speed while institutional reform and governance mechanisms move slowly. This temporal mismatch creates significant challenges for designing and implementing inclusive and equitable AI access policies.
Private AI development vs. Public governance	The market incentives driving today a small number of private firms at the frontier of AI development may conflict with states' responsibility to safeguard the public interest and values such as transparency, accountability, equity, or risk mitigation. This divergence creates tensions between commercial objectives and the imperatives of public governance.
Non-proliferation vs. Widespread AI access	The dual-use nature of AI technologies creates a tension between preventing dangerous proliferation and enabling beneficial access. While traditional non-proliferation regimes assume clear boundaries between civilian and military applications, AI systems resist such clean distinctions.
Sovereignty vs. Coordination	The capacity to develop frontier AI lies almost exclusively with a handful of countries and corporations who, as a result, claim the power to dictate the rules of global diffusion. In this geopolitical context, Global Majority states face the challenge of balancing coordination for widespread AI access with strengthening their sovereignty.
Innovation incentives vs. Distributional imperatives	Governing AI distribution requires balancing competing imperatives. On one hand, aggressive AI benefit-sharing risks undermining the innovation incentives that drive AI development. On the other hand, allowing current concentration patterns to persist may fail to motivate cooperation on substantive benefit-sharing commitments."

Table 5: Summary of the tensions underlying AI benefit-sharing.

Part II: Operationalising Benefit-Sharing

Building on the conceptual foundations, Part II approaches **benefit-sharing** not a singular logic but an integration of three distinct traditions of governance: **redistribution**, which ensures that economic gains are spread across societies; **technology transfer and capacity-building**, which enables states to develop and govern AI themselves; and **non-proliferation and safety**, which manages the security risks of advanced systems while allowing inclusion:

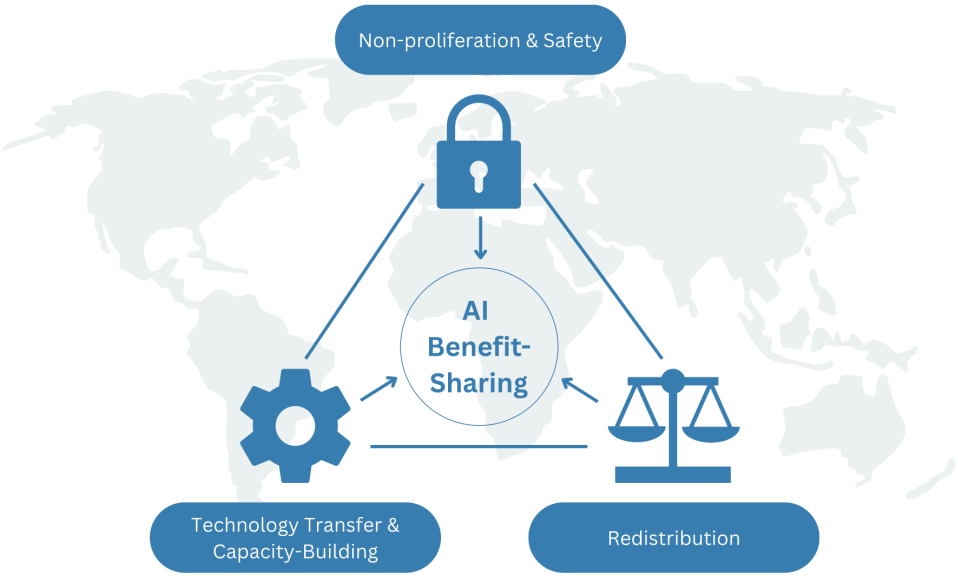


Image 2: AI benefit-sharing components.

Chapter 3 opens the more practical part of the paper by unwrapping domestic readiness for Global Majority states, that is, mapping out the prerequisites needed to prepare for the diffusion of AI benefits and capabilities:

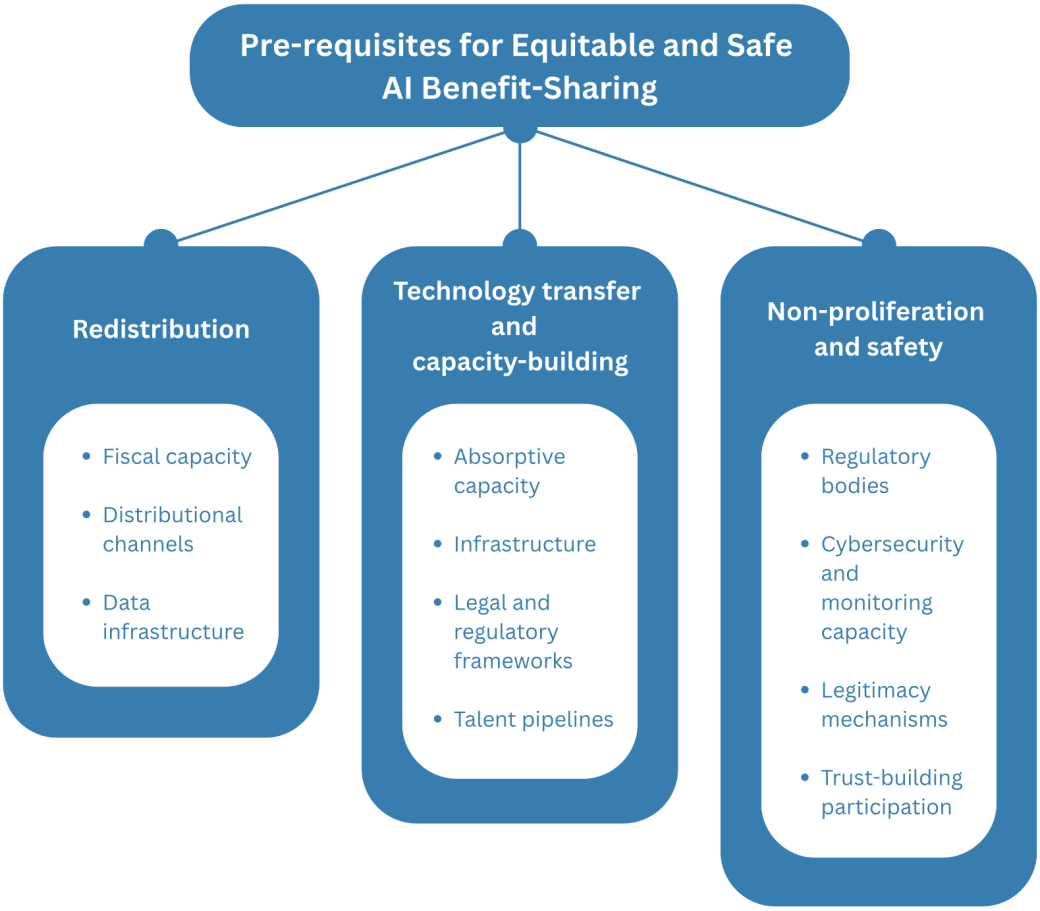


Image 3: Summary table of the prerequisites for equitable and safe AI benefit-sharing.

Chapter 4 outlines concrete mechanisms that states can use to structure the equitable redistribution of economic gains from AI. Three main avenues are explored:

Mechanisms for Redistribution through Rules and Laws	
Regulatory frameworks	1. AI licensing and authorisation regimes 2. Mandatory benefit assessment requirements 3. Distributional impact analysis obligations 4. Compliance and enforcement mechanisms 5. Special Economic Zones (SEZ)-based benefit-sharing mandates 6. Human rights due diligence
Fiscal rules and redistribution mechanisms	1. Taxation of AI-derived value a. Corporate taxation and windfall levies b. Value-added tax (VAT) adjustments c. Data dividend or usage fees 2. Revenue distribution mechanisms a. Sovereign technology funds b. Targeted social transfers c. Public service subsidisation 3. Procurement requirements with benefit-sharing conditions 4. Redistributive Special Economic Zones fiscal regimes 5. Public investment guidelines with distributional provisions
Institutional capacity for redistribution	1. National governance bodies a. Specialised benefit-sharing authorities b. Special Economic Zones oversight bodies c. Multi-stakeholder governance structures d. Coordination mechanisms across agencies 2. Fiscal institutions a. Sovereign technology funds b. Distribution mechanisms and accountability structures 3. International coordination institutions a. Alliances of committed states b. Regional capacity facilities c. Technical standards and harmonisation d. Dispute resolution and compliance mechanisms e. Bilateral Special Economic Zones and cross-border corridors

Table 6: Summary of the redistributive AI benefit-sharing mechanisms.

Chapter 5 moves beyond redistribution to consider how Global Majority states can expand their capacity to develop, govern, and adapt AI technologies themselves. It outlines the mechanisms of technology transfer and capacity-building:

Mechanisms for Technology Transfer and Capacity-Building	
Infrastructure mechanisms	1. Shared compute hubs 2. Cloud credits and subsidised access 3. Regional data centres and sovereign data governance 4. Connectivity and energy investments
Knowledge transfer and human capital	1. Training programmes and fellowships 2. Joint research labs 3. Curriculum development and standards alignment
Legal instruments for technology transfer	1. Intellectual property adaptations a. Compulsory licensing for AI systems b. Adapted copyright for AI-generated content c. Open licensing and shared intellectual property pools 2. Trade and competition law a. Competition law to prevent market concentration b. Technology transfer provisions in trade agreements 3. International legal regimes a. Trade-Related Aspects of Intellectual Property Rights (TRIPS) adjustments for AI b. Model laws and treaty provisions
Partnership and financing models	1. Public–private partnerships (PPPs) a. Joint ventures with benefit-sharing mandates b. Technology access funds 2. South–South cooperation a. Regional centres of excellence b. Cross-regional knowledge networks 3. International financing mechanisms a. Sovereign technology funds b. Global AI capacity facility c. Concessional loans and blended finance
SEZ-based platforms for technology transfer and capacity-building	1. SEZs provide an institutional template through which the various mechanisms of technology transfer and capacity-building can be bundled into a coherent regime.

Table 7: Summary of the technology transfer and capacity-building AI benefit-sharing mechanisms.

Lastly, Chapter 6 reconceptualised non-proliferation for the context of AI as **a graduated access regime** grounded in transparency, differentiated responsibilities, and compensatory measures

where diffusion is restricted. The chapter maps out the different rules and standards for high-risk capabilities:

Rules and Standards for High-Risk Capabilities	
Defining high-risk capabilities	1. Frontier models with dual-use potential 2. Training datasets for high-risk domains 3. Critical compute resources
Licensing and authorization regimes	1. Model licensing 2. User vetting and credentialing 3. Verification measures
International standards and norms	1. Red-lines agreements 2. Baseline safety standards 3. Differentiated responsibilities

Table 8: Summary of the non-proliferation and safety components of AI benefit-sharing.

In conclusion, this paper began with a straightforward observation: despite widespread rhetorical commitment to "sharing the benefits of AI," the mechanisms required to translate principle into practice remain underdeveloped, fragmented, or absent altogether. We are witnessing a technological transformation that could either reduce global inequalities or entrench them permanently, and the window for meaningful intervention is narrowing rapidly.

The paper's central argument is that AI benefit-sharing requires integration across three distinct governance traditions: redistribution of economic gains, technology transfer and capacity-building, and non-proliferation and safety controls.

These three pillars are not alternatives but complementary requirements. Redistribution without capacity-building risks perpetuating dependency; capacity-building without safety mechanisms risks catastrophic misuse; and safety regimes that ignore equity concerns risk reproducing the exclusionary patterns of past arms control regimes. Effective benefit-sharing requires all three, implemented in ways that are mutually reinforcing rather than contradictory.

Image 4: Summary of the selected AI benefit-sharing mechanisms relevant for different stakeholders.

Glossary

The following definitions apply to the terms used in this document and may differ from their meanings in other contexts:

AI-derived value refers here to the financial profits from AI-enabled activities and AI-related industries (e.g. chip production).

AI diffusion refers to the process by which AI technologies spread across societies, sectors, or geographies.

AI divide refers to the disparity across communities in access to AI systems, the necessary enablers fueling the AI industry (e.g. compute, data, expertise), and thus, in harnessing the benefits.

AI-enabled future refers to a vision of society in which AI is deeply integrated into economic, political, and social systems, creating new opportunities for innovation while also posing heightened risks of inequality, dependency, and systemic disruption.

AI safety and security represent two complementary aspects of mitigating AI threats: the former concerns predominantly harm prevention from advanced systems (e.g. catastrophic misuse, bias, labor market disruption), while the latter focuses more on protecting the integrity of models through their design, implementation, and deployment.[14]

AI utility follows here the definition in Solaiman et al. (2025)[15], referring to the qualities (e.g. multilinguality) that enable an individual to meaningfully use the available components of an AI system.

Benefit-sharing refers to the fair allocation of economic, social, and technological gains resulting from AI development and deployment, and involves both *ex ante* and *ex post* mechanisms enabling such fair allocation.

Compute divide refers to the geographic concentration of the key physical AI infrastructure which reinforces existing patterns of global inequality. This concentration means that only a small number of states have the power to turn their compute assets into influence over AI development and governance.

Coordination refers to the collective development of governance frameworks, standards, and cooperative mechanisms to manage AI's risks and distribute its benefits.

[14] See: https://internationalaisafetyreport.org/, https://www.aisi.gov.uk/
[15] Irene Solaiman et al., "Beyond Release: Access Considerations for Generative AI Systems" *arXiv*, 2025 https://arxiv.org/abs/2502.1670

Domestic readiness refers to the fiscal, infrastructural, and regulatory prerequisites that states need to develop so that they can systematically derive tangible benefits from AI.

Dual-use technology refers to technologies that can both be used for beneficial civilian applications, as well as more disruptive, and potentially catastrophic military ones.

Existential risk refers to the possibility that advanced AI systems could cause irreversible harm on a global scale, whether through loss of human control, misuse, or catastrophic systemic disruption.

Frontier AI refers to cutting-edge, highly capable AI technologies with transformative potential and elevated risks.

Global Majority is a political term used to describe the collective of countries in Africa, Asia, Latin America, and the Caribbean that represent most of the world's population as a counterpoint to "Global North,". In this text, the term is used interchangeably with low- and lower-middle income countries, excluding regional AI powerhouses, such as China.

Graduated access regime is an approach to governing access to AI systems aimed at preventing the proliferation of high-risk capabilities while being targeted, legitimate, and compatible with benefit-sharing.

Knowledge transfer refers to the sharing of AI-related expertise (e.g., red-teaming protocols, impact evaluation frameworks, the know-how of building an open-source AI system) between institutions and countries.

Non-proliferation refers to an approach to AI governance that seeks to restrict the spread of powerful models, compute, or techniques deemed dual-use or high-risk.

Redistribution refers here to the *ex ante* and *ex post* mechanisms for sharing of economic gains derived from AI to correct for inequalities.

Sovereignty refers to a state's ability to control its data flows, set its own regulatory priorities, and shape its domestic technological ecosystem without undue external interference or dependencies.

Technology transfer refers to the sharing of model architectures, training data, evaluation protocols, and compute infrastructure between institutions or countries.

Introduction

What kind of future are *we* building and for *whom*? As artificial intelligence (AI) races ahead, measured in months rather than decades, the world is witnessing a technological transformation unfolding faster than any in human history.

Despite widespread recognition of these challenges, systematic mechanisms to ensure equitable AI diffusion while managing risks remain absent from the global governance ecosystem. International declarations, such as the 2025 Paris AI Action Summit Statement[16] or the 2023 Bletchley Declaration,[17] routinely invoke the need for "sharing AI's benefits," "ensuring widespread access," and "bridging the AI divide." The G20 Rio Declaration,[18] the Hamburg Declaration,[19] and the UN's Governing AI for Humanity report[20] all stress similar principles. Even industry leaders like Sam Altman, Demis Hassabis, and Dario Amodei have stated publicly that distributing AI's benefits widely is core to their mission[21].

The paper's central argument is that AI benefit-sharing requires integration across three distinct governance traditions: redistribution of economic gains, technology transfer and capacity-building, and non-proliferation and safety controls. While multilateral institutions and private actors remain essential stakeholders, states possess unique capacity to establish enforceable rules, mobilize resources, build infrastructure, and negotiate binding international arrangements. The analysis places particular emphasis on states in the Global Majority – not because they must become frontier developers, but because they face the immediate challenge of building the institutional prerequisites (fiscal systems, regulatory frameworks, absorptive capacity) necessary to systematically benefit from AI technologies they do not themselves develop. This focus responds to a persistent gap in existing literature, which tends to concentrate on frontier development or high-level international coordination while neglecting the domestic readiness requirements essential for benefit-sharing mechanisms to function in practice.

The paper's contribution is threefold. First, it develops a conceptual framework that situates AI benefit-sharing within competing governance traditions – technology transfer, diffusion, and non-proliferation – each carrying distinct normative implications and institutional logics. By clarifying these framings and their underlying tensions, the framework provides analytical tools

---

[16] See: https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet

[17] See: https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023

[18] See: https://g20.org/wp-content/uploads/2024/11/G20-Rio-de-Janeiro-Leaders-Declaration-EN.pdf

[19] See: https://www.bmz-digital.global/wp-content/uploads/2025/06/250603_Hamburg_Declaration.pdf

[20] See: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

[21] For example, see their mission statements: https://openai.com/about/; https://deepmind.google/about/; https://www.anthropic.com/company

for navigating trade-offs between speed and deliberation, private development and public governance, access and security, sovereignty and coordination, and innovation incentives and distributional imperatives. Second, it identifies specific axes of policy action required for operationalization, demonstrating how redistribution, capacity-building, and safety mechanisms must function as integrated rather than isolated interventions. Third, it maps concrete institutional pathways and policy mechanisms, specifying both minimum prerequisites for readiness and viable implementation strategies across different governance contexts.

The analysis operates under a set of scope conditions that reflect current geopolitical and technological trajectories. We assume continued US-China strategic competition but not overt military conflict that would entirely subordinate AI governance to security imperatives. We recognize that frontier AI development remains concentrated in private firms, but assume states retain sufficient leverage through regulation, procurement, and infrastructure control to shape industry behavior. And we focus primarily on how Global Majority states can build capacity for AI adoption and governance rather than frontier development.These premises do not narrow the paper's argument; they clarify the environment in which benefit-sharing mechanisms will realistically operate.

Part I situates the central argument of this paper by addressing both the frequent arguments sidelining the issue and clarifying the conceptual terrain on which AI benefit-sharing must be redefined:

- Chapter 1 argues that prevailing assumptions (e.g. that markets will diffuse AI benefits organically, that restrictive national strategies are unavoidable, or that U.S.–China rivalry forecloses cooperation) are misleading and represent common obstacles to implementing AI benefit-sharing. Markets alone rarely deliver equitable diffusion; restrictive strategies deepen exclusion; and even in rivalry, selective cooperation is possible and often necessary.
- Chapter 2 unwraps and clarifies the conceptual framings underlying AI benefit-sharing (e.g. diffusion, technology transfers), analysing the specific implications of each narrative. It also outlines and unpacks the different tensions underpinning AI access governance to account for the challenges that AI benefit-sharing strategies may need to navigate.

Part II then moves to the pragmatic concern of how to operationalise AI benefit-sharing. Based on the identified three dimensions of action: *redistribution*, *technology transfers and capacity-building*, as well as *non-proliferation and safety*, the remaining chapters outline a framework for applying benefit-sharing into practice and the specific preparedness prerequisites that states need to ensure in the first place.

- Chapter 3 introduces domestic readiness – the foundations that would help states materialise the potential value from AI. It outlines three key components of this readiness (redistribution, technology transfer and capacity-building, and non-proliferation and safety) and their specific prerequisites.
- Chapter 4 focuses on the requirements for redistribution of AI-derived economic gains and the specific redistributive mechanisms that are available.
- Chapter 5 addresses the capacity required to develop, govern and adapt AI technologies, taking the approach that AI benefit-sharing starts already with inclusiveness in the AI development process. It also showcases multiple tools, such as infrastructure building, knowledge transfers and others.
- Chapter 6 is a response to the dual-use nature of AI and adds a lens of how to ensure that AI diffusion is safe and beneficial, rather than proliferating dangerous or capabilities.

Part I: Situating AI Benefit-Sharing

This part lays the conceptual and political foundations for the paper. It begins by asking why AI benefit-sharing, despite frequent invocation in policy discourse, remains underexplored in practice. Chapter 1 argues that prevailing assumptions that markets will diffuse AI benefits organically, that restrictive national strategies are unavoidable, or that U.S.–China rivalry forecloses cooperation are misleading. Markets alone rarely deliver equitable diffusion; restrictive strategies deepen exclusion; and even in rivalry, selective cooperation is possible and often necessary. By challenging these views, the chapter clarifies the stakes of continued neglect: without deliberate action, AI risks entrenching permanent global asymmetries.

Chapter 2 then turns to conceptual clarity. It shows that "benefit-sharing" in AI cannot be reduced to a narrow redistributive exercise. Instead, it intersects with neighboring frames such as technology transfer, diffusion, and non-proliferation, each carrying distinct histories and normative implications. The chapter identifies five tensions that cut across these debates: speed vs. deliberation, private development vs. public governance, non-proliferation vs. access, sovereignty vs. coordination, and innovation incentives vs. distributional imperatives. Taken together, these tensions explain why benefit-sharing has been so difficult to operationalise, and why different stakeholders align or misalign around it.

Part I therefore situates the central argument of this paper that AI benefit-sharing requires integration across three distinct governance traditions: redistribution of economic gains, technology transfer and capacity-building, and non-proliferation and safety controls, by clarifying both the flawed arguments that sideline the issue and the conceptual terrain on which it must be redefined.

Chapter 1: Why AI Benefit-Sharing Is Sidelined and Why It Matters

Despite frequent declarations about the need for equitable AI development,[22] benefit-sharing remains peripheral in most policy agendas. Policymakers often treat it as a problem that will resolve itself, a distraction from national strategy, or a casualty of great-power rivalry.[23] These views, while influential, are misleading. Despite arguments to the contrary,[24] markets alone rarely ensure fair diffusion of transformative technologies, restrictive strategies by leading states reinforce exclusion, and the current geopolitical tensions magnify these effects, further fragmenting the global AI landscape.

As such, this chapter sets out to challenge these assumptions. It examines the arguments most often used to deprioritise AI benefit-sharing, shows why they are inadequate in the case of these frontier technologies, clarifies the stakes of continued neglect, and finally, underscores the importance of actualising benefit-sharing. It also argues that even within rivalry, opportunities for selective cooperation remain, and that seizing them will be essential to prevent permanent asymmetries in the distribution of AI's benefits.

[22] United Nations Conference on Trade and Development, *Technology and Innovation Report 2025: Inclusive Artificial Intelligence for Development*, Chapter V: "Global collaboration for inclusive and equitable AI", 2025, UNCTAD https://unctad.org/system/files/official-document/tir2025ch5_en.pdf: World Economic Forum. *A Blueprint for Equity and Inclusion in Artificial Intelligence*, 2022, 8 https://www3.weforum.org/docs/WEF_A_Blueprint_for_Equity_and_Inclusion_in_Artificial_Intelligence_2022.pdf

[23] For example, during the Paris AI summit, the US and UK declined to sign inclusive AI statements likely due to an interest is asserting their own dominance, see: *From Global Governance to Nationalism: The Future of AI*, GIGA Focus Global, No. 2, 2025 https://www.giga-hamburg.de/en/publications/giga-focus/from-global-governance-to-nationalism-the-future-of-ai; Sammy Martin, Justin Bullock & Corin Katzke, "Analysis of Global AI Governance Strategies", *Convergence Analysis*, 4 Dec 2024, https://www.convergenceanalysis.org/research/analysis-of-global-ai-governance-strategies.

[24] For example see: OECD, *Scale, Market Power and Competition in a Digital World*, 2023, https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/01/scale-market-power-and-competition-in-a-digital-world_2f43b51d/c1cff861-en.pdf

Counter-Arguments to AI Benefit-Sharing	
Argument Against AI Benefit-Sharing	**Response Summary**
The Market Myth: diffusion of AI benefits will occur naturally through market forces	1. Market incentives alone often fall short. 2. Even where market failures are recognized and addressed, traditional intervention mechanisms may be too slow for the AI transition timeline. 3. Existing redistribution mechanisms, especially relevant for market failures, are inadequate to address global AI equity.
The U.S. AI approach: strategic restraint as the default, limiting pathways for meaningful global diffusion	1. The domestic foundation of the U.S. AI leadership is under strain. 2. Preserving U.S. leadership requires more than hoarding compute – it requires shaping the global distribution of AI through partnerships. 3. The economic logic of commercial diplomacy supports a more outward-facing approach.
U.S.–China AI geopolitical tensions overshadow international cooperation on AI diffusion	1. AI competition is not zero-sum in all dimensions. 2. Specific areas of convergent interest exist despite rivalry. 3. The costs of non-cooperation are rising.

Table 1: Summary of the common counter-arguments sidelining AI benefit-sharing and the paper's responses to them.

1.1 Argument I: The Market Myth: diffusion of AI benefits will occur naturally through market forces

The market myth argument suggests that diffusion of AI benefits will occur naturally through commercial expansion, competition, or eventual cost reduction.[25] This view, often implicit in policy discussions, assumes that market incentives alone will drive AI companies to make their products widely available. After all, major players like OpenAI, Google, and Anthropic are already expanding globally, offering their products across multiple countries with free tiers alongside premium services. The competitive landscape drives companies to seek new markets

[25] Claire Dennis et al., *Options and Motivations for International AI Benefit Sharing*, Centre for the Governance of AI, 2025, 36 https://www.governance.ai/research-paper/options-and-motivations-for-international-ai-benefit-sharing.

and user bases.[26] From this perspective, as AI technologies mature, their spillover effects will likely reach broader populations without the need for proactive redistribution mechanisms.

The diffusion of mobile phones in the 1990s and early 2000s is often cited as a model: companies like Nokia and Ericsson drove the proliferation of GSM networks and low-cost handsets across Kenya, Nigeria, India, and other emerging markets.[27] Services like M-Pesa in Kenya later leveraged this infrastructure to enable widespread mobile banking, often cited as a case of "leapfrogging" development.[28]

Today, AI already appears relatively accessible. Consumer-facing tools such as ChatGPT, Gemini, and Claude are used by hundreds of millions worldwide, and open-source releases, distillation, and fine-tuning communities have lowered the barrier to experimentation. Yet, as Solaiman et al. show, the mere release of models does not guarantee meaningful access.[29] True diffusion depends not only on availability and release but also on whether populations can derive **utility** from accessible capabilities. That is because meaningful access to AI systems encompasses several prerequisites: the availability of necessary infrastructure, the technical skills required to engage with an AI system, and utility - the qualities (e.g. multilingual support) that enable an individual to meaningfully use the available components.[30] Subscription fees, licensing restrictions, hardware costs, and limited institutional capacity often prevent communities from turning nominal access into real empowerment. In practice, access is stratified: broad public use exists, but the most capable frontier systems remain gated and carefully controlled. The critical question is whether this pattern will continue, or whether the frontier will eventually diffuse in ways that provide usable and beneficial capabilities to a wider range of societies.

However, assuming that market forces are sufficient for effective AI diffusion risks resting on an overly generic analogy between AI and earlier general-purpose technologies. While some diffusion is likely to occur, this argument overlooks several critical dynamics:

---

[26] This assumption has roots in classical technology diffusion theory, which posits that innovations spread through populations via adoption curves driven primarily by individual choice and competitive market dynamics. See Everett M. Rogers, *Diffusion of Innovations*, 3rd ed., Free Press, 1983. For contemporary discussions of market-driven approaches to digital technology access, see OECD, Scale, Market Power and Competition in a Digital World, 2021, https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/01/scale-market-power-and-competition-in-a-digital-world_2f43b51d/c1cff861-en.pdf. The framework assumes that competitive pressures and declining costs naturally incentivize firms to expand access, thereby distributing technological benefits broadly over time.

[27] Garcia-Swartz, Daniel D., & Martin Campbell-Kelly**,** *Cellular: An Economic and Business History of the International Mobile-Phone Industry*. Cambridge, MA: The MIT Press, 2022, 288 https://direct.mit.edu/books/oa-monograph/5457/CellularAn-Economic-and-Business-History-of-the.

[28] Calestous Juma, "Leapfrogging Progress: The Misplaced Promise of Africa's Mobile Revolution," 7 *Breakthrough Journal*, 2017 https://thebreakthrough.org/journal/issue-7/leapfrogging-progress

[29] Irene Solaiman et al., "Beyond Release: Access Considerations for Generative AI Systems" *arXiv*, 2025 https://arxiv.org/abs/2502.16701; D.G. Widder, S. West & M. Whittaker, "Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI" Social Science Research Network 2023, 2 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4543807

[30] Irene Solaiman et al., "Beyond Release: Access Considerations for Generative AI Systems" *arXiv*, 2025 https://arxiv.org/abs/2502.16701.

Response 1: Market incentives alone often fall short

Historical examples show that even "neutral" or commodifiable technologies didn't diffuse equitably without intervention/required intervention to diffuse equitably:
Technologies often believed to be neutral or universally empowering such as the printing press, radio, and mobile phones have historically failed to diffuse evenly without intentional efforts. The internet was hailed as a global equalizer, yet today's digital divide reflects enduring disparities in infrastructure, language access, and institutional capacity.[31] Even mobile phones, often cited as a success story of market-driven innovation, required public-private partnerships,[32] donor-backed infrastructure,[33] and regulatory reform[34] to reach rural and underserved communities. For example, in Rwanda, government adoption vouchers substantially increased the diffusion of mobile phones in rural areas, generating widespread benefits for the targeted populations.[35] Similarly, in Bangladesh, the Grameen Telecom Village Phone program provided up to 2.8 million rural residents with phone access, markedly expanding diffusion across the country.[36] These cases show that seemingly neutral technologies do not automatically proliferate widely.[37]

For technologies with strong public good dimensions and high barriers to entry, market incentives failed on their own:
Where technologies carry high public value but face significant development costs, market mechanisms consistently fall short. Vaccines offer a clear case: the TRIPS waiver debate during the COVID-19 pandemic exposed how intellectual property regimes and limited manufacturing capacity in the Global South blocked timely access to life-saving mRNA technologies, despite

---

[31] Sangmoon Kim, "The Diffusion of the Internet: Trend and Causes," 40(2) *Social Science Research*, 2011 https://www.sciencedirect.com/science/article/abs/pii/S0049089X10001377.
[32] Dongesit Williams & Morten Falch, *Public-Private Partnerships and Next Generation Networks*, in A. M. Hadjiantonis & B. Stiller (eds.), *Telecommunication Economics* 7216 Lecture Notes in Computer Science, 2012 https://www.researchgate.net/publication/230610442_Public_Private_Partnerships_and_Next_Generation_Network Grace Langham, *Rural Coverage Initiatives: Stakeholders Can Learn from the Successes and Failures of Past Efforts Analysys Mason Research*, 2023 https://www.analysysmason.com/research/content/articles/rural-initiatives-lessons-rdns0
[33] Natalie Meyenn & Cina Lauson, *Bringing Cellular Phone Service to Rural Areas: Grameen Telecom and Village Pay Phones in Bangladesh*, World Bank Note No. 205, 2000 https://openknowledge.worldbank.org/server/api/core/bitstreams/738ca9a6-2bd9-5999-b9d8-2784dd29dd72/content
[34] Daniel Björkegren & Burak Ceyhun Karaca, "The Effect of Network Adoption Subsidies: Evidence from Digital Traces in Rwanda" *arXiv* 2020, 3 https://arxiv.org/pdf/2002.05791; Georges V. Houngbonon, Marc Ivaldi, Emil Palikot & Davide Strusani, "The Impact of Shared Telecom Infrastructure on Digital Connectivity and Inclusion" *arXiv* 2025, https://arxiv.org/pdf/2507.19693
[35] Daniel Björkegren & Burak Ceyhun Karaca, "The Effect of Network Adoption Subsidies: Evidence from Digital Traces in Rwanda" *arXiv* 2020, 3 https://arxiv.org/pdf/2002.05791.
[36] Natalie Meyenn & Cina Lauson, *Bringing Cellular Phone Service to Rural Areas: Grameen Telecom and Village Pay Phones in Bangladesh*, World Bank Note No. 205, 2000 https://openknowledge.worldbank.org/server/api/core/bitstreams/738ca9a6-2bd9-5999-b9d8-2784dd29dd72/content
[37] S. Sangwan & L. F. Pau, *Diffusion of Mobile Phones in China*, ERIM Report Series, 2005 https://ideas.repec.org/p/hal/psewpa/halshs-00844446.html.

unprecedented global need and profitability.[38] In pharmaceuticals, it continues to require massive public subsidies,[39] creating incentives to fulfill technology transfer agreements,[40] and push-pull funding mechanisms, like Gavi and CEPI,[41] to make life-saving treatments available beyond the richest markets. AI represents a particularly complex case within this pattern of market failure.

Much like the pharmaceutical sector, where both intellectual-property rights and limited regional manufacturing capacity have hindered equitable access,[42] the AI industry's exclusionary mechanisms also operate through resource concentration and infrastructural control.[43] Capabilities in both industries are concentrated in a handful of Global North countries, but the concentration in AI may be even more pronounced. For example, by some estimates the top twenty companies in the pharmaceutical sector account for roughly 66% of global sales,[44] reflecting a transnational oligopoly.[45] In AI, concentration is striking in key infrastructure segments: NVIDIA controls at least 70% of AI chips market[46] and major technology firms collectively dominate the foundation-model and platform market.[47] Despite decades of efforts to improve equitable access to medicines in low- and middle-income countries, disparities persist, with significant implications for their populace.[48] Given the even greater concentration of AI compute, models, and infrastructure, the challenge of ensuring fair and widespread access to AI technologies is likely at least as severe as in the pharmaceutical sector, if not worse.

---

[38] T. Chaudhary & A. Chaudhary, "TRIPS Waiver of COVID-19 Vaccines: Impact on Pharmaceutical Industry and What It Means to Developing Countries" 24(5–6) *Journal of World Intellectual Property*, 2021 https://pmc.ncbi.nlm.nih.gov/articles/PMC8661626/.

[39] James Love & Tim Hubbard, "Prizes for Innovation of New Medicines and Vaccines," 18(2) *Annals of Health Law*, 2009, 177 https://lawecommons.luc.edu/cgi/viewcontent.cgi?article=1111&context=annals.

[40] International Federation of Pharmaceutical Manufacturers & Associations (IFPMA), *Technology Transfer: A Collaborative Approach to Improve Global Health*, 2023 https://www.ifpma.org/wp-content/uploads/2023/01/i2023_IFPMA_Technology_Transfer_2015_Web.pdf.

[41] Coalition for Epidemic Preparedness Innovations, *Enabling Equitable Access,* https://cepi.net/equitable-access.

[42] Chamas C et al, *Intellectual Property and Medicine: Towards Global Health Equity* in Wong T and Dutfield G (eds), *Intellectual Property and Human Development: Current Trends and Future Scenarios* (Cambridge University Press 2010), 64-65.

[43] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5 2024, https://hdsr.mitpress.mit.edu/pub/xdukxlpp.

[44] Hardman & Co Life Sciences Research, *2024 Pharma Statistics: Strong 9.2% market growth – driven by obesity,* 2025 https://hardmanandco.com/wp-content/uploads/2025/04/250403-Hardman-Co-Insight-2024-Pharma-Statistics.pdf.

[45] Sparke, M., & Williams, O, "COVID and structural cartelisation: market-state-society ties and the political economy of Pharma" 29(4), *New Political Economy,* 2024, https://www.tandfonline.com/doi/full/10.1080/13563467.2024.2304180.

[46] *Nvidia AI Accelerator Market Outlook (2023–2027)*, SiliconAnalysts, https://siliconanalysts.com/nvidia-ai-accelerator-market-outlook-2023-2027/.

[47] Amba Kak, Sarah Myers West & Meredith Whittaker, "Make No Mistake—AI Is Owned by Big Tech", *MIT Technology Review*, 2023 https://www.technologyreview.com/2023/12/05/1084393/make-no-mistake-ai-is-owned-by-big-tech/.

[48] Oldfield, L., Penm, J., Mirzaei, A., & Moles, R, "Prices, availability, and affordability of adult medicines in 54 low-income and middle-income countries: evidence based on a secondary analysis" 13(1) *The Lancet Global Health*, 2025 https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(24)00442-X/fulltext.

Furthermore, in the case of AI, release does not equal access[49], which encompasses the broader physical and institutional ability to use or develop AI. However, training frontier models demands infrastructure investments affordable to only a handful of global actors; semiconductor export controls create hardware chokepoints more restrictive than any patent regime; and even open-source models depend on cloud platforms monopolized by a few concentrated providers.

Strategic AI capabilities are even less likely to diffuse:
AI is classified as a critical and foundational technology that fundamentally shapes national power, economic competitiveness, and security.[50] Current frontier AI development is concentrated in a handful of firms and states, primarily in the U.S. and China, whose motivations are shaped not only by commercial incentives but also by national security, geopolitical positioning, and control over emerging global standards.[51] These actors are not simply responding to market demand; they are actively shaping who gets access and on what terms. Additionally, the current market size of the Global Majority economy does not appear as a priority for AI labs. India, however, appears to be an exception,[52] likely due to the scale of its current and planned AI investments compared with other Global Majority countries.[53] Beyond India, we do not see the same momentum for market capture in other Global Majority regions. Moreover, diffusion depends not only on cost and competition, but also on infrastructure, institutional capacity, and absorptive readiness. In low-resource contexts, these barriers prevent even beneficial and profitable technologies from spreading.[54]

Response 2: Even where market failures are recognized and addressed, traditional intervention mechanisms may be too slow for the AI transition timeline

The speed mismatch between AI development and market diffusion timelines:
Historically, technology diffusion operated on multi-generational timescales, even for beneficial technologies with clear economic advantages.[55] The printing press, invented around 1450, took until the late 17th century to establish meaningful presence in sub-Saharan Africa and parts of

[49] Irene Solaiman et al., "Beyond Release: Access Considerations for Generative AI Systems" *arXiv*, 2025 https://arxiv.org/abs/2502.16701.
[50] See: https://www.rand.org/content/dam/rand/pubs/research_reports/RRA3200/RRA3295-1/RAND_RRA3295-1.pdf
[51] Lennart Heim, *Understanding the Artificial Intelligence Diffusion Framework: Can Export Controls Create a U.S.–Led Global Artificial Intelligence Ecosystem?*, RAND Corporation Perspective, 2025, https://www.rand.org/pubs/perspectives/PEA3776-1.html.
[52] *OpenAI's Altman Meets with India IT Minister to Discuss Country's AI Plans* Reuters, 2025 https://www.reuters.com/technology/openais-altman-meets-with-india-it-minister-discuss-countrys-ai-plans-2025-02-05/: "Students in India Just Got a Gemini Upgrade" *Google India Blog*, 2025, https://blog.google/intl/en-in/company-news/technology/students-in-india-just-got-a-gemini-upgrade/.
[53] Tambiama Madiega & Rafał Ilnicki, *AI Investment: EU and Global Indicators*, European Parliament, 2024, https://www.europarl.europa.eu/RegData/etudes/ATAG/2024/760392/EPRS_ATA(2024)760392_EN.pdf.
[54] Irene Solaiman et al., *Beyond Release: Access Considerations for Generative AI Systems*, *arXiv,* 2025 https://arxiv.org/abs/2502.16701.
[55] Jeffrey Ding, "The Rise and Fall of Technological Leadership: General-Purpose Technology Diffusion and Economic Power Transitions", 68(2) *International Studies Quarterly*, 2024 https://doi.org/10.1093/isq/sqae013.

Asia, despite its obvious commercial benefits.[56] Electricity required 60 years to reach 50% household penetration in the United States after Edison's first power station in 1882, and rural electrification remained incomplete until the 1950s despite federal intervention.[57] The telephone, patented in 1876, did not achieve widespread adoption in developing countries until the 1980s and 1990s, over a century later.[58] Even the internet, celebrated for its rapid growth, required 25 years to reach 50% global penetration and still leaves 2.9 billion people without access as of 2023.[59] AI capabilities, by contrast, are advancing exponentially. Leading models improve dramatically every 6-18 months, with some researchers predicting transformative AI systems within the current decade.[60] This creates an unprecedented temporal mismatch: by the time market forces naturally overcome adoption barriers and reduce costs sufficiently for global diffusion, the foundational architecture of the AI ecosystem – including control over key infrastructure, the setting of technical standards, and the shaping of governance frameworks – will already be locked in.

AI's unique characteristics create compounding advantages that may permanently lock in current asymmetries before market forces can respond:
Given the speed of AI advancement and infrastructure concentration, waiting for organic diffusion may result in permanent asymmetries.[61] By the time costs fall, foundational control may already be locked in. The dynamics of AI development display strong *returns to scale and information asymmetry*, advantaging early movers.[62] Labs and states with frontier capabilities are already shaping global norms, safety benchmarks, and regulatory templates. This risks "locking in" not only material advantages but also economic and political influence, cultural biases, loyalty and epistemic authority into the foundations of the AI ecosystem.[63] Further, this compounding occurs under conditions of extreme resource concentration amplifying global inequalities due to the AI divide. Unlike previous general-purpose technologies, frontier AI is unfolding under conditions of extreme capital intensity, strategic secrecy, and infrastructural

[56] Oxford Companion to the Book, *Oxford Reference*, https://www.oxfordreference.com/display/10.1093/acref/9780198606536.001.0001/acref-9780198606536-e-0039.
[57] Lorenzo Pellegrini & Luca Tasciotti, *Rural Electrification Now and Then: Comparing Contemporary Challenges in Developing Countries to the USA's Experience in Retrospect*, https://repub.eur.nl/pub/39064/Metis183450.pdf.
[58] *Telephone*, EBSCO Research Starters https://www.ebsco.com/research-starters/history/telephone.
[59] Teutem, Simon van, *Internet Use Became the Norm for Humanity Only Very Recently*. Our World in Data, 2025 https://ourworldindata.org/data-insights/internet-use-became-the-norm-for-humanity-only-very-recently.
[60] Zershaaneh Qureshi, *Timelines to Transformative AI: An Investigation*, LessWrong 2024, https://www.lesswrong.com/posts/ddj5HtnCHHMQGiQEM/timelines-to-transformative-ai-an-investigation.
[61] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5 2024 https://hdsr.mitpress.mit.edu/pub/xdukxlpp.
[62] Sandrine Kergroach & Julien Héritier, *Emerging Divides in the Transition to Artificial Intelligence*, OECD Regional Development Papers No. 147 June 2025, 8 https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/emerging-divides-in-the-transition-to-artificial-intelligence_eeb5e120/7376c776-en.pdf.
[63] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5 2024, https://hdsr.mitpress.mit.edu/pub/xdukxlpp.

centralization—especially in compute, data, and talent.[64] The global bifurcation of AI compute capacity indicates that waiting for natural diffusion risks entrenching structural exclusions that may become irreversible.[65] Eventually, the potential for rapid, transformative capability gains means that current advantages could translate into permanent global stratification.[66] As Dario Amodei has suggested, we may be entering a "compressed 21st century" where the first actors to harness transformative AI could leapfrog decades of scientific and technological progress within a single decade.[67]

Response 3: Existing redistribution mechanisms, especially relevant for market failures, are inadequate to address global AI equity

AI challenges the traditional redistribution mechanisms:
Some may suggest that the existing ecosystem of global redistribution mechanisms such as foreign aid, development finance, and capacity-building initiatives, if repurposed, would ensure AI benefit-sharing. However, this lens inappropriately conflates benefit-sharing, which includes both *ex ante* and *ex post* mechanisms, with mere *ex-post* redistribution and overlooks both the historical failures of such mechanisms and the novel demands of AI as a general-purpose, dual-use, and strategically sensitive technology.

Such efforts have often fostered dependency rather than empowerment,[68] and have been vulnerable to political cycles and traction (e.g. abrupt defunding of agencies like USAID;[69] UK aid cuts to shift money to defence[70]). They have also failed to eradicate global exploitation or systematically improve the geopolitical position of aid-receiving nations.[71] These precedents raise legitimate skepticism: if global redistribution hasn't worked effectively before, why would AI be any different?

---

[64] Jai Vipra & Sarah Myers West, *Computational Power and AI*, AI Now Institute, 2023 https://www.ainowinstitute.org/publications/compute-and-ai.
[65] Billy Perrigo, *AI Computing Global Divide*, *The New York Times*, 2025 https://www.nytimes.com/interactive/2025/06/23/technology/ai-computing-global-divide.html.
[66] Indermit Gill, Commentary, *Whoever Leads in Artificial Intelligence in 2030 Will Rule the World Until 2100*, 2020, Brookings, https://www.brookings.edu/articles/whoever-leads-in-artificial-intelligence-in-2030-will-rule-the-world-until-2100/.
[67] Dario Amodei, *Machines of Loving Grace*, 2024, https://www.darioamodei.com/essay/machines-of-loving-grace.
[68] Valentina Finckenstein, *How International Aid Can Do More Harm Than Good: The Case of Lebanon*, LSE IDEAS Strategic Update 2021 https://www.lse.ac.uk/ideas/Assets/Documents/updates/LSE-IDEAS-How-International-Aid-Can-Do-More-Harm-Than-Good.pdf
[69] The White House, *Reevaluating and Realigning United States Foreign Aid*, 2025, https://www.whitehouse.gov/presidential-actions/2025/01/reevaluating-and-realigning-united-states-foreign-aid/.
[70] Paul Johnson, *Cut Aid and Benefits to Boost Defence? The Sums Don't Quite Add Up*, Institute for Fiscal Studies 2025 https://ifs.org.uk/articles/cut-aid-and-benefits-boost-defence-sums-dont-quite-add.
[71] Todd Moss, Gunilla Pettersson & Nicolas van de Walle, *An Aid-Institutions Paradox? A Review Essay on Aid Dependency and State Building in Sub-Saharan Africa*, 255 in William Easterly, ed., *Reinventing Foreign Aid* MIT Press 2008.

AI is qualitatively different from past technologies redistributed through aid.[72] Unlike vaccines, mobile phones, or solar panels, frontier AI systems are (i) capital- and infrastructure-intensive, requiring vast compute, skilled human capital, data ecosystems, and regulatory frameworks; (ii) national-security sensitive, leading to active restrictions on their transfer (e.g. US export controls on GPUs and AI models);[73] and (iii) rapidly evolving, making slow-moving development mechanisms ineffective. Increasingly, the traditional foreign aid system is not a viable equity-promoting strategy,[74] and it is the tech companies, not multilaterals or state agencies, that play an increasing role in AI-enabled development.[75] As such, there is a need to rethink the nature of global redistribution mechanisms for AI benefits.

The weakening of multilateralism:
What is more, multilateral institutions that could, in theory, facilitate AI benefit-sharing are increasingly ill-suited to govern frontier AI development.

Multilateral bodies were established with the primary unit of governance being the nation-state. However, the frontier of AI research and deployment is dominated by a handful of private laboratories that are traditionally not direct subjects of international law.[76] These entities are not state-owned, and while state-market entanglements exist, multilateral institutions lack legal, financial, or strategic leverage over these corporate actors. For perspective, the United Nations' 2024 strained budget of $3.59 billion[77] is eclipsed by OpenAI's projected $10 billion annual revenue by mid-2025.[78]

Crucially, the UN has no regulatory enforcement power over the activities of AI companies, and no strategic resource (such as critical compute infrastructure) that could serve as leverage. Multilateral institutions are therefore dependent on state cooperation, which renders them largely symbolic actors in the AI governance space without state involvement. For example, the United

---

[72] Francesco Filippucci, Peter Gal, Cecilia Jona-Lasinio, Álvaro Leandro & Giuseppe Nicoletti, *The Impact of Artificial Intelligence on Productivity, Distribution and Growth: Key Mechanisms, Initial Evidence and Policy Challenges*, OECD Artificial Intelligence Papers No. 15 2024 https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/the-impact-of-artificial-intelligence-on-productivity-distribution-and-growth_d54e2842/8d900037-en.pdf.
[73] U.S. Department of Commerce, Bureau of Industry and Security, *Framework for Artificial Intelligence Diffusion*, https://public-inspection.federalregister.gov/2025-00636.pdf.
[74] *AI is the New Foreign Aid*, Financial Times, 2025, https://www.ft.com/content/d02eb244-8b48-48b1-bd17-f5e48677e22b.
[75] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5 2024, https://hdsr.mitpress.mit.edu/pub/xdukxlpp.
[76] Amba Kak, Sarah Myers West & Meredith Whittaker, "Make No Mistake—AI Is Owned by Big Tech", *MIT Technology Review*, 2023 https://www.technologyreview.com/2023/12/05/1084393/make-no-mistake-ai-is-owned-by-big-tech/.
[77] United Nations, *The Price of Peace and Development: Paying for the UN*, UN Affairs, 2024 https://news.un.org/en/story/2024/01/1145272.
[78] *OpenAI's Annualized Revenue Hits $10 Billion, Up from $5.5 Billion in December 2024*, Reuters, 2025 https://www.reuters.com/business/media-telecom/openais-annualized-revenue-hits-10-billion-up-55-billion-december-2024-2025-06-09/.

States' recent withdrawals from UN bodies depict the fragility of multilateralism when major powers see their strategic interests misaligned with global consensus mechanisms.[79] In 2025, the US announced its withdrawal from UNESCO and the suspension of funding to several UN-affiliated agencies, citing concerns over putting "America first" and opposing China's growing influence within international governance bodies.[80] These actions erode the operational capacity of multilateral institutions by cutting financial lifelines and undermining their legitimacy as neutral conveners.

Moreover, the divergence in AI priorities and capabilities among member states creates a potential policy paralysis within multilateral forums. For instance, the United States approaches AI governance through a lens of national security and strategic competition with China, whereas emerging economies are more concerned with AI's role in their development, digital inclusion, and capacity-building.[81] Attempting to reconcile these competing agendas within a single multilateral framework often leads to lowest-common-denominator outcomes, vague declarations, or outright deadlock.

1.2 Argument II: The U.S. AI approach: strategic restraint as the default, limiting pathways for meaningful global diffusion

As the global AI leader, the U.S. plays an outsized role in shaping access to compute and the development of frontier models or global safety norms. Therefore, the next two arguments address the common U.S.-centric perspectives against the prioritisation of international AI diffusion. The first obstacle centres on the United States' own strategic posture[82]: its current policy, as reflected in the 2025 AI Action Plan[83], is built on a logic that emphasizes strategic restraint and protectionism, not broad diffusion. Congressional testimonies[84] from national security officials and industry leaders consistently portray AI as a domain of zero-sum

[79] The White House, *Withdrawing the United States from and Ending Funding to Certain United Nations Organizations and Reviewing United States Support to All International Organizations*, 2025 https://www.whitehouse.gov/presidential-actions/2025/02/withdrawing-the-united-states-from-and-ending-funding-to-certain-united-nations-organizations-and-reviewing-united-states-support-to-all-international-organizations/.
[80] U.S. Department of State, *The United States Withdraws from the United Nations Educational, Scientific and Cultural Organization,* 2025 https://www.state.gov/releases/office-of-the-spokesperson/2025/07/the-united-states-withdraws-from-the-united-nations-educational-scientific-and-cultural-organization-unesco.
[81] A. Ishkhanyan, "The Sovereignty-Internationalism Paradox in AI Governance: Digital Federalism and Global Algorithmic Control" 5 *Discover Artificial Intelligence* 2025, 123 https://link.springer.com/article/10.1007/s44163-025-00374-x.
[82] The White House, *America's AI Action Plan*, 2025 https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf.
[83] See: https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf
[84] Charles W. "Chip" Pickering Jr., *Written Testimony Before the Subcommittee on Communications and Technology, House Committee on Energy and Commerce*, 119th Congress 2025 https://www.congress.gov/119/meeting/house/118333/witnesses/HHRG-119-IF16-Wstate-PickeringC-20250604.pdf

competition with China.[85] For instance, hearings have highlighted concerns over the US lead narrowing to mere months, with witnesses comparing it to an "AI Sputnik moment" and stressing the need for rapid infrastructure scaling to counter China's advancements.[86]

Contrary to the assumption that the US has limited incentives for global AI diffusion, several factors suggest otherwise:

Response 1: The domestic foundation of the U.S. AI leadership is under strain

The ability to expand large-scale compute clusters which are essential for training, deployment, and experimentation now collides with hard constraints in energy availability and permitting. Estimates suggest that training a frontier model by 2030 will require gigawatt-scale energy inputs which is on par with multiple nuclear plants.[87] Yet the U.S. power grid is aging, congested, and bureaucratically gridlocked.[88] Some data centres already face multi-year delays for grid connection.[89] In contrast, geopolitical competitors such as China[90] and the UAE[91] can mobilize energy and infrastructure at speeds the U.S. regulatory apparatus cannot match. This asymmetry creates a paradox: even as the U.S. seeks to hoard compute domestically, it lacks the physical and institutional capacity to scale at the pace the frontier demands.

Response 2: Preserving U.S. leadership requires more than hoarding compute – it requires shaping the global distribution of AI through partnerships

The report by the Center for a New American Security (CNAS) on the interlink between compute diffusion and national security emphasizes that American AI strategy must operate on two levels simultaneously: sustaining control over frontier-scale systems that push the edge of capability, and enabling U.S. firms to distribute subfrontier compute globally in ways that align

---

[85] Marina Yue Zhang, *The China–U.S. AI Race Enters a New (and More Dangerous) Phase*, The Diplomat, 2025 https://thediplomat.com/2025/05/the-china-us-ai-race-enters-a-new-and-more-dangerous-phase/.
[86] Adam Thierer, *Submitted Statement before the House Committee on Science, Space, and Technology, Subcommittee on Research and Technology: "DeepSeek: A Deep Dive" Hearing*, 119th Congress, 2025 https://www.congress.gov/119/meeting/house/118111/witnesses/HMTG-119-SY15-Wstate-PickeringC-20250604.pdf: Julia Stoyanovich, *Testimony Before the Subcommittee on Research and Technology, Committee on Science, Space, and Technology, U.S. House of Representatives: "DeepSeek: A Deep Dive"*, 119th Congress, 2025 https://www.congress.gov/119/meeting/house/118111/witnesses/HMTG-119-SY15-Wstate-StoyanovichJ-20250408.pdf.
[87] Konstantin F. Pilz, Yusuf Mahmood & Lennart Heim, *AI's Power Requirements Under Exponential Growth: Extrapolating AI Data Center Power Demand and Assessing Its Potential Impact on U.S. Competitiveness* RAND Corporation, 2025 https://www.rand.org/pubs/research_reports/RRA3572-1.html.
[88] Shalini Bhat, *Aging Electric Infrastructure in the United States*, Interdisciplinary Professional Programs, University of Wisconsin–Madison, 2025 https://interpro.wisc.edu/aging-electric-infrastructure-in-the-united-states/.
[89] Tim McLaughlin, *Power-Boosting Project for U.S. Grid to Miss July Deadline*, Reuters 31 January 2025 https://www.reuters.com/business/energy/power-boosting-project-us-grid-miss-july-deadline-2025-01-31/.
[90] Kyle Chan et al., *Full Stack: China's Evolving Industrial Policy for AI*, RAND Corporation, 2025 https://www.rand.org/pubs/perspectives/PEA4012-1.html.
[91] Gregory C. Allen, Georgia Adamson, Lennart Heim & Sam Winter-Levy, *The United Arab Emirates' AI Ambitions* Center for Strategic and International Studies, 2025 https://www.csis.org/analysis/united-arab-emirates-ai-ambitions.

with U.S. interests.[92] Access to compute determines which countries can participate meaningfully in the AI era not only as users, but as co-creators of standards, safety protocols, and deployment norms.[93] If the U.S. does not fill this space, China is already poised to do so. Through initiatives like the Digital Silk Road, Chinese firms such as Huawei, Alibaba Cloud, and Tencent are rapidly constructing AI-ready infrastructure – data centers, cloud platforms, and edge computing hubs – across Africa, Southeast Asia, and the Middle East.[94]

Response 3: The economic logic of commercial diplomacy supports a more outward-facing approach

The US has strong incentives to advance business interests in the global marketplace through diplomatic channels, particularly through "commercial diplomacy".[95] AI-driven economic growth in emerging markets creates new consumer bases and investment opportunities that benefit US companies. In the AI context, commercial diplomacy involves leveraging trade partnerships, development finance, export guarantees, and technology transfer arrangements to position US firms as key actors in the AI ecosystems of tomorrow. Leading in global compute allows the U.S. to maintain AI leadership, thereby accelerating innovation, shaping global norms, and safeguarding national security.[96] Such efforts have precedent. The U.S. International Development Finance Corporation (DFC), for instance, has invested in secure telecommunications infrastructure across the Indo-Pacific[97] to counterbalance Huawei's dominance. This model has now been adapted for cloud and compute infrastructure under initiatives like the U.S. Global Infrastructure Partnership (PGII).[98] Promoting American AI abroad is part of a broader "commercial diplomacy" strategy that reinforces U.S. industrial strength while drawing other countries into the U.S.-led technological sphere.

---

[92] Janet Egan, *Global Compute and National Security: Strengthening American AI Leadership Through Proactive Partnerships* Center for a New American Security 2025, https://www.cnas.org/publications/reports/global-compute-and-national-security.

[93] Jai Vipra & Sarah Myers West, *Computational Power and AI*, AI Now Institute, 2023 https://www.ainowinstitute.org/publications/compute-and-ai.

[94] Paul Triolo, Kevin Allison, Clarise Brown & Kelsey Broderick, *The Digital Silk Road: Expanding China's Digital Footprint,* Eurasia Group, 2020 https://www.eurasiagroup.net/files/upload/Digital-Silk-Road-Expanding-China-Digital-Footprint.pdf.

[95] The White House, Executive Order No. 14280, *Promoting the Export of the American AI Technology Stack*, 2025 https://www.whitehouse.gov/presidential-actions/2025/07/promoting-the-export-of-the-american-ai-technology-stack/.

[96] Janet Egan, *Global Compute and National Security: Strengthening American AI Leadership Through Proactive Partnerships* Center for a New American Security 2025, https://www.cnas.org/publications/reports/global-compute-and-national-security.

[97] *Supporting Resilient Telecommunications Infrastructure in the South Pacific*, US International Development Finance Corporation, 2025 https://www.dfc.gov/investment-story/supporting-resilient-telecommunications-infrastructure-south-pacific.

[98] The White House, *Fact Sheet: Partnership for Global Infrastructure and Investment at the G7 Summit,* 2024 https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/06/13/fact-sheet-partnership-for-global-infrastructure-and-investment-at-the-g7-summit-2/.

1.3 Argument III: U.S.–China AI geopolitical tensions overshadow international cooperation on AI diffusion

A second strand of Argument II focuses less on American internal dynamics and more on the broader geopolitical environment, particularly the breakdown of trust between the United States and China.[99] Here, the claim is not just that cooperation between the US and China is politically inconvenient, but that it is strategically unwise. For example, the strategic competition can be used to justify and necessitate closed models. U.S.–China relations are increasingly framed in zero-sum terms across trade, security, and technology domains. AI now sits at the centre of this rivalry.

Both countries have adopted competing theories of AI diffusion, underpinned by divergent political values and governance styles. The United States treats diffusion both as a matter of national interest[100] and a strategic economic and political tool.[101] Its strategy is one of selective openness: compute exports are tightly controlled, infrastructure partnerships are limited to allies, and frontier model access is governed through licensing, not openness. In contrast, China's strategy places more emphasis on deliberate export of its AI technologies. Its July 2025 *Global AI Governance Action Plan* calls for equitable access to AI infrastructure, public-private compute partnerships, and broad model availability for Global South states.[102] These diffusion strategies are not merely technical, they reflect deep ideological differences over sovereignty, governance, and the purpose of AI development.

[99] Ryan Hass, Ryan McElveen & Lily McElwee, *Advancing U.S.-China Coordination amid Strategic Competition: An Emerging Playbook* CSIS 2025, https://www.csis.org/analysis/advancing-us-china-coordination-amid-strategic-competition-emerging-playbook.
[100] Barath Harithas, *The AI Diffusion Framework: Securing U.S. AI Leadership While Preempting Strategic Drift* Center for Strategic International Studies, 2025 https://www.csis.org/analysis/ai-diffusion-framework-securing-us-ai-leadership-while-preempting-strategic-drift.
[101] Sam Manning, *Beyond Export Controls: How Strategic Promotion of American AI Abroad Can Protect the Homeland*, The Republic, 2025 https://therepublicjournal.com/web-exclusives/beyond-export-controls/.
[102] Ministry of Foreign Affairs of the People's Republic of China, *Global AI Governance Action Plan*, 2025 https://www.fmprc.gov.cn/eng./xw/zyxw/202507/t20250729_11679232.html.

Axes of U.S.–China Divergence		
Axis	**United States**	**China**
AI Diffusion Strategy	***America's AI Action Plan*** (July 2025) emphasizes American dominance: controlled diffusion via strategic export controls (especially on compute and dual-use models), with licensing regimes and "trusted" AI infrastructure limited to allies. Supports open-source innovation but restricts access to frontier systems. It aims at "full AI stack" diffusion among partners to strengthen "American values" and counteract the increasing presence of China in emerging markets.	***Global AI Governance Action Plan*** (July 2025) promotes broad diffusion through infrastructure investment in the Global South, advocating equitable access to AI capacity and governance rights. Positions Chinese AI models and tools as affordable, sovereign-compatible alternatives.
Conditions for AI Cooperation	Frames cooperation around shared democratic values, voluntary norms, and safety protocols—largely via existing like-minded coalitions (e.g., OECD-GPAI merger, Quad, Hiroshima AI Process Friends Group of the G7). Multilateralism is limited and selective.	Frames sovereignty as a non-negotiable precondition for cooperation. Reaffirms commitment to multilateralism under the UN system; calls for "equal participation" of Global South states in norm-setting and infrastructure.
Approach to AI Safety	Centred on frontier safety and existential risk mitigation: mandates model evaluations above compute thresholds, red-teaming, and pre-deployment disclosure. Safety is increasingly framed as both a technical and national security problem.	Emphasises controllability and accountability, especially in social, ethical, and content domains. Safety is increasingly framed as a public security priority, comparable to cyber threats and pandemics. China has rapidly expanded national safety standards and red-teaming protocols, yet safety remains tightly state-controlled, with limited industry transparency and international engagement.
AI Development Model	Privileges market-led frontier innovation, reinforced by public R&D investment (e.g., CHIPS Act, NSF AI institutes) and immigration reform for technical talent. Open competition among private labs is encouraged.	Anchored in state-directed strategic planning with strong public-private coordination. National AI champions are shaped by long-term five-year plans, with civil-military fusion embedded in the model.

Table 2: A summary table comparing the US and China AI Action Plans through the lens of AI diffusion and international AI governance.[103]

[103] See: The White House, *America's AI Action Plan,* 2025 https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf; Ministry of Foreign Affairs of the People's Republic of China, *Global AI Governance Action Plan,* 2025

This divergence has three implications. First, it erodes the normative foundations of multilateralism. The U.S. preference for "club governance" (e.g., the G7, the recently merged OECD-GPAI, and Quad) creates exclusionary dynamics,[104] while China's inclusive multilateralism approach is frequently criticised as instrumentalist and norm-shaping.[105] Second, it fragments the AI ecosystem into rival infrastructure zones. Without a shared global framework, we risk the emergence of parallel AI spheres of influence with incompatible safety approaches, interoperability, and access standards. Third, it turns cooperative institutions into theatres of strategic contestation. The U.S. AI Action Plan explicitly pledges to counter Chinese influence in global governance bodies,[106] further weakening the credibility of forums like the UN, which already suffer from underfunding and limited authority over frontier actors.

The argument that the geopolitical rift between the United States and China renders cooperation on AI not only unlikely but unwise may appear compelling at first glance. But this conclusion rests on a flawed premise: that strategic rivalry and international collaboration are mutually exclusive. On closer examination, this view fails to account for the complexity of interdependence, the nature of global risk, and the shared incentives embedded within the very technologies at stake.

Response 1: AI competition is not zero-sum in all dimensions

Much of the prevailing rhetoric surrounding the U.S.–China AI dynamics rests on the metaphor of a race.[107] It is largely framed as a contest in which one nation's gain must come at the other's expense.[108]

Shared risks are structurally non-zero-sum and neither the U.S. nor China can mitigate these threats alone. Unlike nuclear weapons, where deterrence regimes have been internationally

---

https://www.fmprc.gov.cn/eng/xw/zyxw/202507/t20250729_11679232.html; Organisation for Economic Co-operation and Development (OECD) & Global Partnership on Artificial Intelligence (GPAI), *GPAI and OECD Unite to Advance Coordinated International Efforts for Trustworthy AI*, OECD (3 July 2024), https://www.oecd.org/en/about/news/speech-statements/2024/07/GPAI-and-OECD-unite-to-advance-coordinated-international-efforts-for-trustworthy-AI.html; Gabriel Wagner, Jason Zhou, Kwan Yee Ng & Brian Tse, *State of AI Safety in China* Concordia AI, 2025, https://concordia-ai.com/wp-content/uploads/2025/07/State-of-AI-Safety-in-China-2025.pdf.

[104] Stewart Patrick, *Four Contending U.S. Approaches to Multilateralism*, Carnegie Endowment for International Peace, 2023 https://carnegieendowment.org/research/2023/01/four-contending-us-approaches-to-multilateralism?lang=en.

[105] Ding Xuexiang, *Keeping to the Right Path of Multilateralism and Promoting Open and Inclusive Development*, Ministry of Foreign Affairs of the People's Republic of China, 2025 https://www.fmprc.gov.cn/eng/xw/zyjh/202501/t20250121_11542131.html.

[106] The White House, *America's AI Action Plan*, 2025 https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf.

[107] Notably, while U.S.-China cooperation in research fell by 5% in 2021 across most domains, cooperation in AI rose by 3%: *U.S.–China Research Collaboration May Be Falling—but Not in AI*, Emerging Technology Observatory, 2023 https://eto.tech/blog/datapoints-us-china-research-falling-not-ai/

[108] Seán Ó hÉigeartaigh, *The Most Dangerous Fiction: The Rhetoric and Reality of the AI Race,* Social Science Research Network 2025, https://ssrn.com/abstract=5278644.

instituted and enforcement relies on observable hardware and stockpile parity,[109] AI risk depends on opaque systems, untested model behaviors, and distributed deployment. In this context, a single catastrophic failure, whether in safety alignment, training data poisoning, or exploitability, could have cascading global consequences.

That makes AI safety a non-zero-sum public good: the more actors contribute to it, the safer the system is for all.[110] Attempts to exclude China from model evaluation efforts,[111] or vice versa, do not reduce existential risk; they externalize it. Likewise, failure to coordinate on compute verification and red-teaming protocols can fuel escalation cycles based on misperception or mistrust.

Response 2: Specific areas of convergent interest exist despite rivalry

Rather than requiring comprehensive cooperation, benefit-sharing can focus on areas where U.S. and Chinese interests naturally align:

- **Safety and Risk Mitigation**: Both nations share interests in preventing AI-enabled global instability, cyberattacks, and loss of governmental authority.[112] Research shows that geopolitical rivals can cooperate effectively on technical AI safety measures, including verification mechanisms and shared protocols.[113]
- **Economic Stability**: Uncontrolled AI-driven economic disruption threatens both U.S. and Chinese domestic stability. Both nations have incentives to: coordinate gradual AI deployment to prevent mass unemployment; share best practices for economic transition frameworks, and jointly develop international norms preventing destabilising AI use. Historical patterns reveal that even amid rivalry, both nations have consistently prioritised economic stabilization during periods of systemic disruption.[114]

---

[109] Treaty on the Non-Proliferation of Nuclear Weapons, 729 UNTS 161; 7 ILM 109 (1968).
[110] Kayla Blomquist et al., *Examining AI Safety as a Global Public Good: Implications, Challenges, and Research Priorities* Oxford Martin School, 2025, https://aigi.ox.ac.uk/publications/examining-ai-safety-as-a-global-public-good-implications-challenges-and-research-priorities/.
[111] Bureau of Industry and Security, *Framework for Artificial Intelligence Diffusion*, 90 Fed. Reg. 4544, 2025 https://www.federalregister.gov/documents/2025/01/15/2025-00636/framework-for-artificial-intelligence-diffusion
[112] The White House, *America's AI Action Plan,* 2025 https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf: Ministry of Foreign Affairs of the People's Republic of China, *Global AI Governance Action Plan*, 2025, https://www.fmprc.gov.cn/eng./xw/zyxw/202507/t20250729_11679232.html.
[113] Bucknall B et al, "In Which Areas of Technical AI Safety Could Geopolitical Rivals Cooperate?" *arXiv,* 2025, https://www.arxiv.org/pdf/2407.06232.
[114] Christopher S. Chivvis et al., *U.S.-China Relations for the 2030s: Toward a Realistic Scenario for Coexistence,* Carnegie Endowment for International Peace, 2024, https://carnegieendowment.org/research/2024/10/us-china-relations-for-the-2030s-toward-a-realistic-scenario-for-coexistence?lang=en.

U.S.–China Responses to Systemic Economic Disruption				
Case	**Disruption**	**U.S. Response**	**China Response**	**Shared Incentive**
2008 Global Financial Crisis	Global recession, financial system collapse.	$700B TARP bailout; liquidity injections; stimulus spending.	$586B stimulus focused on infrastructure and domestic demand.	Prevent global economic collapse; stabilize trade and employment.
Post-WTO Accession ("China Shock")	Trade-driven labor displacement and structural adjustment.	2–2.4M manufacturing job losses; rise in populism and regional inequality.	Urban unemployment spike; large-scale rural-urban migration; social tensions.	Preserve domestic stability amid rapid globalization and structural change.

Table 3: U.S.–China responses to systemic economic disruption.[115]

Response 3: The costs of non-cooperation are rising

While strategic rivalry between the United States and China has defined much of the global AI discourse, the costs of sustained non-cooperation are becoming increasingly difficult to ignore. Both states have adopted defensive postures in pursuit of technological sovereignty and geopolitical leverage. Yet this mutual entrenchment imposes significant structural burdens on innovation, influence, and institutional legitimacy. The asymmetrical but mutually restrictive effects of ongoing fragmentation across infrastructure, market access, governance, and research ecosystems are described in *Table 4* below. It demonstrates that while competition may be unavoidable, total disengagement is neither strategically sound nor practically sustainable.

[115] Alan S. Blinder & Mark Zandi, *The Financial Crisis: Lessons for the Next One*, Center on Budget and Policy Priorities, 2015 https://www.cbpp.org/research/economy/the-financial-crisis-lessons-for-the-next-one; World Bank, *China's Stimulus Policies Are Key for Growth in 2009 and an Opportunity for More Rebalancing*, Press Release No. 2008/11/25, 24 November 2008, https://www.worldbank.org/en/news/press-release/2008/11/24/chinas-stimulus-policies-key-growth-2009-opportunity-more-rebalancing-says-world-bank-update.

Rising Costs of U.S.–China AI Non-Cooperation		
Dimension	**United States**	**China**
Infrastructure Constraints	Domestic energy bottlenecks limit frontier AI scaling; delays in grid expansion and permitting restrict new data center deployment.	Semiconductor restrictions constrain access to advanced chips, limiting the training and deployment of frontier models.
Global Market Access	Export controls weaken U.S. firms' competitiveness abroad; restrictions on model access and compute push partners toward alternative suppliers.	Defensive posturing reduces opportunities for global market expansion, particularly in emerging economies.
Normative Influence	Tightening "trusted partner" regimes erode U.S. influence over global AI norms, governance frameworks, and deployment standards.	Overemphasis on sovereignty and national self-reliance limits participation in shared governance regimes and multilateral institutions.
Alliance Management	Conditional access frameworks cause ally fatigue; demand grows for non-U.S. infrastructure and AI systems.	Political isolation and tight state control discourage foreign collaboration and investment, increasing dependence on domestic firms.
Innovation Ecosystem	Isolation risks losing access to diverse deployment contexts and slowing diffusion of U.S.-aligned tools and models.	Reduced access to global research communities and talent flows risks epistemic stagnation in frontier research domains.
Strategic Leverage	Fragmentation undermines the U.S. ability to coordinate collective responses to AI risks or influence emerging norms.	Strategic overreach and self-imposed limitations may slow China's ability to shape global standards or lead on safety.

Table 4: Rising costs of U.S.–China AI non-cooperation.

The responses in this chapter to the common arguments sidelining AI benefit-sharing highlight the likely trends for AI diffusion and thus, the likely future of access to its benefits. Together, they show how, without adequate mechanisms in place, the continued neglect of pro-active diffusion strategies risks consolidating global asymmetries. We now turn to the very notion of AI benefit-sharing and its conceptual foundations.

Chapter 2: Conceptual Foundations

The main aim of this chapter is to understand the underlying conceptual framings for sharing AI benefits, including its gains, tools or enablers for participation, and provide an analytical background for discussing the potential solutions in later chapters.

2.1 Evolving meanings in the governance of AI access

Discussions of global access to artificial intelligence frequently invoke a range of terms as benefit-sharing, technology transfer, diffusion, and non-proliferation, each of which carries distinct historical origins and normative implications. While these terms are often used interchangeably, their original meanings differ significantly, and their adaptation to AI reveals important shifts in how the technology is being politically framed.

Benefit-sharing originates in environmental governance, particularly in treaties such as the Convention on Biological Diversity (CBD)[116] and the Nagoya Protocol,[117] where it refers to the fair and equitable distribution of benefits derived from the use of genetic resources.[118] In that context, benefit-sharing was designed, in part, to compensate local or Indigenous communities for their contributions to environmental protection[119] and to avoid extractive dynamics by ensuring that profits, technologies, or knowledge generated from those resources would be shared with the providers.[120] Notably, benefit-sharing frameworks are transactional and have thus had limited success in promoting biodiversity conservation and public health due to power imbalances embedded in the global marketplace.[121] When applied to AI, the term has been adapted to refer to the fair allocation of economic, social, and technological gains or benefits resulting from AI development and deployment.[122] This includes, for example, access to AI-enabled services, financial returns, capacity-building opportunities, and decision-making influence. The use of this term in the AI context signals a normative view of AI as a global

---

[116] *Convention on Biological Diversity*, opened for signature 5 June 1992, 1760 U.N.T.S. 79 (entered into force 29 December 1993).
[117] *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from Their Utilization to the Convention on Biological Diversity*, opened for signature 29 October 2010, UN Doc. UNEP/CBD/COP/DEC/X/1 (entered into force 12 October 2014).
[118] Elisa Morgera, "The Need for an International Legal Concept of Fair and Equitable Benefit Sharing" 27(2) *European Journal of International Law*, 2016, 353 https://academic.oup.com/ejil/article/27/2/353/1748393.
[119] Article 8 (j), *Convention on Biological Diversity*, opened for signature 5 June 1992, 1760 U.N.T.S. 79 (entered into force 29 December 1993).
[120] Elisa Morgera, *Under the Radar: Fair and Equitable Benefit-Sharing and the Human Rights of Indigenous Peoples and Local Communities Related to Natural Resources,* Working Paper 2016, 2 https://www.researchgate.net/publication/311824205_Under_the_Radar_Fair_and_Equitable_Benefit-Sharing_and_the_Human_Rights_of_Indigenous_Peoples_and_Local_Communities_Related_to_Natural_Resources.
[121] Eccleston-Turner M, and Rourke M, "Arguments against the inequitable distribution of vaccines using the access and benefit sharing transaction" 70(4) *International & Comparative Law Quarterly*, 2021, 842, https://doi.org/10.1017/S0020589321000294.
[122] Claire Dennis et al., *Options and Motivations for International AI Benefit Sharing*, Centre for the Governance of AI, 2025, 7 https://www.governance.ai/research-paper/options-and-motivations-for-international-ai-benefit-sharing.

public good, one that, like natural resources, produces extractable value that should be governed through distributional justice mechanisms.[123]

By contrast, **technology transfer** comes from the domain of international trade and development. Traditionally, it refers to the deliberate movement of technical knowledge, tools, and expertise across actors, typically from technologically advanced states or firms to those with less capacity.[124] Its aim is to support local innovation, reduce dependency, and promote self-sufficiency.[125] In the context of AI, technology transfer has come to mean the sharing of model architectures, training data, evaluation protocols, and compute infrastructure between institutions or countries.[126] This framing presumes that access to the underlying technical systems is a prerequisite for meaningful participation in the AI ecosystem. Simultaneously, the use of this term reflects a view of AI as a rather strategic asset – for example, in the U.S.,[127] the European Union,[128] or China[129] – one that can be transferred, but still remains bound up in concerns about competitiveness and intellectual property.

AI diffusion is derived from the definition of technology diffusion rooted in innovation studies and economics, and traditionally refers to the process by which new technologies spread across societies, sectors, or geographies. Diffusion occurs through adoption, replication, and integration into everyday systems.[130] In AI governance discourse, diffusion is typically used in a descriptive sense: to track where and how AI is being deployed globally.[131] Framed as such, AI diffusion is presented as likely to occur automatically and in a manner largely shaped by adoption and economic incentives. However, this neutral framing often obscures the structural constraints—such as infrastructure gaps or compute monopolies—that limit diffusion in practice.[132] As such, appeals to diffusion can reinforce the mistaken assumption that AI benefits will inevitably "trickle down" without intervention.

---

[123] Iason Gabriel, *Toward a Theory of Justice for Artificial Intelligence*, American Academy of Arts & Sciences, 2021, 220 https://doi.org/10.1162/daed_a_01911.
[124] Cooper C and Sercovich F, *The Channels and Mechanism for the Transfer of Technology from Developed to Developing Countries: A Study,* UNCTAD Secretariat, 1971, 2.
[125] Commission on Intellectual Property Rights, *Integrating Intellectual Property Rights and Development Policy* Report, September 2002, 20 http://www.iprcommission.org/papers/pdfs/final_report/ciprfullfinal.pdf.
[126] Claire Dennis et al., *Options and Motivations for International AI Benefit Sharing*, Centre for the Governance of AI, 2025, 19 https://www.governance.ai/research-paper/options-and-motivations-for-international-ai-benefit-sharing.
[127]See: : https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf
[128] See: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN
[129] See: https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/
[130] Everett M. Rogers, *Diffusion of Innovations*, 3rd ed. Free Press, 1983, 6.
[131] Emmanouil Papagiannidis, Patrick Mikalef & Kieran Conboy, "Responsible Artificial Intelligence Governance: A Review and Research Framework", *Journal of Strategic Information Systems*, 2025, https://www.sciencedirect.com/science/article/pii/S0963868724000672.
[132] Maslej, Nestor, et al., *Artificial Intelligence Index Report 2024*, Stanford Institute for Human-Centered Artificial Intelligence, 2024 https://hai.stanford.edu/ai-index/2024-ai-index-report.

Finally, **non-proliferation** stems from arms control and international security. In its original context, it refers to efforts to prevent the spread of dangerous technologies, particularly nuclear weapons, to actors that are not already in possession of them.[133] Mechanisms of non-proliferation include export controls, verification regimes, and diplomatic agreements. In the AI context, the term is increasingly used to justify restrictions on access to powerful frontier models, compute resources, or algorithmic techniques that are deemed dual-use or high-risk.[134] This framing positions AI as a potentially destabilizing force, akin to strategic weapons, and supports the idea that its development and use should be tightly controlled. It often stands in tension with benefit-sharing, as the imperative to prevent harm can justify exclusivity and opacity.

2.2 Underlying tensions

The feasibility and design of benefit-sharing mechanisms are fundamentally shaped by underlying tensions that require explicit acknowledgment and strategic navigation. These are not merely academic distinctions or unfortunate complications, they represent fundamental design choices that will determine how AI benefits are distributed globally and who gets to participate in shaping the AI-enabled future.[135]

Understanding these tensions is crucial because they reveal why AI benefit-sharing is inherently challenging and why simple solutions are likely to fail. They also point toward more sophisticated governance approaches that can navigate multiple competing pressures simultaneously. Rather than viewing these as fatal contradictions, we can approach them as design challenges that require adaptive institutional capacity.

Tension 1: Speed vs. Deliberation

AI development operates on exponential timescales that fundamentally misalign with democratic governance processes. Leading models improve dramatically every 6-18 months, with researchers predicting transformative capabilities within the current decade.[136] Meanwhile, institutional reform, from international treaty negotiation to domestic policy implementation,

[133] Allan S. Krass, *Uranium Enrichment and Nuclear Weapon Proliferation*, in *SIPRI Yearbook 1983: World Armaments and Disarmament* Almqvist & Wiksell, 1983, Chapter 7 https://www.sipri.org/sites/default/files/files/books/SIPRI83Krass/SIPRI83Krass07.pdf.
[134] Helen Toner, *Nonproliferation Is the Wrong Approach to AI Misuse*, Rising Tide, 2025 https://helentoner.substack.com/p/nonproliferation-is-the-wrong-approach.
[135] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5, 2024 https://hdsr.mitpress.mit.edu/pub/xdukxlpp.
[136] Shirin Ghaffary, *Anthropic CEO Thinks AI May Outsmart Most Humans as Soon as 2026*, Bloomberg, 2024 https://www.bloomberg.com/news/newsletters/2024-10-18/anthropic-ceo-thinks-ai-may-outsmart-most-humans-as-soon-as-2026.

unfolds over years or decades.[137] This creates a structural disadvantage for inclusive governance approaches.

The European Union's AI Act,[138] for instance, took four years to develop and will require additional years for full implementation.[139] During that same period, AI capabilities progressed from GPT-3 to GPT-4 to multimodal systems approaching human-level performance across diverse domains.[140] By the time comprehensive regulatory frameworks are operational, the technological landscape they were designed to govern may have fundamentally shifted.

This temporal mismatch creates pressure for emergency deployment without adequate consideration of distributional consequences. When AI systems promise rapid solutions to urgent challenges such as climate modeling, pandemic response, or economic forecasting, democratic deliberation can seem like a luxury rather than a necessity. Yet, rushing to deploy without ensuring equitable access risks permanently entrenching existing inequalities into the technological infrastructure of the future.

Tension 2: Private AI development vs. Public governance

The concentration of frontier AI development within private entities, particularly a small number of technology firms, creates fundamental tensions between commercial objectives and the imperatives of public governance.[141] Private developers operate under market-based incentives: rapid capability scaling, competitive advantage, and capital returns.[142] On the other side, states, particularly democratic ones, are mandated to safeguard public values: transparency, accountability, equity, and risk mitigation.[143] This divergence creates a tension in both the means and ends of AI governance between the private AI developers and the public interests.[144]

This tension also echoes Karl Polanyi's argument that modern economies are structured by a "double movement": the advance of economic liberalism (the disembedding of markets from social control), and the reactive emergence of social protection (measures to re-embed markets

---

[137] Nicole Simonelli, *Bargaining Over International Multilateral Agreements: The Duration of Negotiations*, 37 International Negotiation, 2011, 145 https://www.researchgate.net/publication/233434460_Bargaining_Over_International_Multilateral_Agreements_The_Duration_of_Negotiations.
[138] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence http://data.europa.eu/eli/reg/2024/1689/oj.
[139] Artificial Intelligence Act, "Implementation Timeline," ArtificialIntelligenceAct.eu (2024), https://artificialintelligenceact.eu/implementation-timeline/.
[140] OpenAI, *ChatGPT Release Notes*, https://help.openai.com/en/articles/6825453-chatgpt-release-notes.
[141] Swati Srivastava & Justin Bullock, *AI, Global Governance, and Digital Sovereignty*, 2024, 1 https://arxiv.org/pdf/2410.17481.
[142] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5, 2024, https://hdsr.mitpress.mit.edu/pub/xdukxlpp.
[143] Swati Srivastava & Justin Bullock, *AI, Global Governance, and Digital Sovereignty*, 2024, 2 https://arxiv.org/pdf/2410.17481.
[144] T. Züger & H. Asghari, "AI for the Public: How Public Interest Theory Shifts the Discourse on AI", 38 *AI & Society*, 2023 https://link.springer.com/article/10.1007/s00146-022-01480-5#citeas.

within normative and institutional constraints).[145] In the AI context, there is a rapid commodification of core societal functions including data, cognition, decision-making authority, knowledge, and social coordination, and this process is unfolding faster than public institutions can adapt, leaving governance efforts reactive, fragmented, and disadvantaged.[146] As such, AI benefit redistribution emerges not just as a normative aspiration, but rather, a strategy to avoid societal harm (e.g. displacement of workers, algorithmic discrimination, erosion of trust, and social fragmentation due to misinformation and manipulation).[147]

However, the view that private firms act in isolation from the state would be incomplete. While development is led by private labs, their scaling is reliant on publicly provisioned inputs: national energy infrastructure,[148] public R&D subsidies, access to talent shaped by public education, and permissive regulatory environments.[149] This state–market interdependence creates a paradox: while firms retain formal autonomy, they are functionally dependent on public ecosystems. As such, their choices have public consequences, while governments remain partially complicit in the concentration of power, resources, and access.[150]

This entanglement transcends domestic governance, extending into geopolitical arenas. For example, OpenAI's "AI for Countries" programme is coordinated in tandem with the U.S. government.[151] The establishment of sovereign AI infrastructure in the UAE (e.g., Stargate UAE) illustrates how private AI ventures serve dual purposes: expanding corporate market capture while simultaneously advancing national geopolitical leverage.[152] Both actors stand to benefit from such "commercial diplomacy". For example, OpenAI gains strategic footholds and investment flows, while the U.S. consolidates influence over emergent AI ecosystems abroad and fosters political alliances.[153] This model of coordinated techno-industrial diplomacy blurs the boundary between private enterprise and state politics. The private-public governance tensions are thus deeply entangled with the market-state co-dependence.[154]

---

[145] Karl Polanyi, *The Great Transformation: The Political and Economic Origins of Our Time*, Beacon Press, 1944, 75.
[146] Anka Reuel & Trond Arne Undheim, "Generative AI Needs Adaptive Governance" *arXiv*, 2024, https://arxiv.org/pdf/2406.04554.
[147] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5, 2024 https://hdsr.mitpress.mit.edu/pub/xdukxlpp.
[148] Ben Hayum, *The Weakest Link: Strategic Inputs in U.S.-China AI Competition*, Americans for Responsible Innovation, Policy Bytes, 2024 https://ari.us/policy-bytes/the-weakest-link-strategic-inputs-in-u-s-china-ai-competition/.
[149] Jaime Sevilla et al., *Can AI Scaling Continue Through 2030?*, Epoch AI, 20 August 2024 https://epoch.ai/blog/can-ai-scaling-continue-through-2030.
[150] Swati Srivastava & Justin Bullock, "AI, Global Governance, and Digital Sovereignty" *arXiv*, 2024, 2 https://arxiv.org/pdf/2410.17481.
[151] OpenAI, *Introducing OpenAI for Countries*, 2025 https://openai.com/global-affairs/openai-for-countries/.
[152] OpenAI, *Introducing Stargate UAE*, 2025 https://openai.com/index/introducing-stargate-uae/.
[153] Janet Egan, *Global Compute and National Security: Strengthening American AI Leadership Through Proactive Partnerships*, Center for a New American Security, 2025 https://www.cnas.org/publications/reports/global-compute-and-national-security.
[154] Swati Srivastava & Justin Bullock, "AI, Global Governance, and Digital Sovereignty" *arXiv,* 2024, 2 https://arxiv.org/pdf/2410.17481.

Tension 3: Non-proliferation vs. Widespread AI access

The dual-use nature of AI technologies creates a fundamental tension between preventing dangerous proliferation and enabling beneficial access. Non-proliferation advocates argue that certain AI capabilities pose significant security risks, ranging from enabling sophisticated cyberattacks to accelerating bioweapons development, and therefore, require strict controls on their distribution.[155] Conversely, broader access proponents emphasize that these same AI systems offer transformative benefits for global development, scientific research, and human welfare, making restrictive approaches potentially harmful to global equity and progress.[156]

This tension becomes particularly acute because of the strongly coupled nature of AI capabilities, where beneficial and dangerous uses often stem from the same underlying technical capabilities. A foundation model capable of accelerating drug discovery through protein folding predictions could equally assist in designing biological weapons.[157] Similarly, AI systems that enhance cybersecurity defenses can also enable more sophisticated offensive capabilities.[158] Unlike previous dual-use technologies where civilian and military applications might require different technical specifications, AI systems are inherently general-purpose, making it difficult to provide access to beneficial capabilities while restricting dangerous ones.

However, the degree of coupling varies across different AI capabilities and applications, challenging simplistic framings of this tension. Research on AI safety and governance distinguishes between "tightly coupled" systems where beneficial and harmful capabilities are nearly inseparable, and "loosely coupled" systems where some degree of access control might be feasible.[159] Cybersecurity applications illustrate this complexity: while defensive and offensive capabilities often overlap, researchers have identified potential technical approaches for "differential access" that could provide defensive benefits while limiting offensive potential.[160] Yet these solutions remain largely theoretical, and empirical evidence for their effectiveness at scale is limited. Moreover, studies of technology transfer and proliferation suggest that even

---

[155] Miles Brundage et al., *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* Future of Humanity Institute, February 2018, 6 https://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/MaliciousUseofAI.pdf?ver=1553030594217.

[156] Ido Alon, Hazar Haidar, Ali Haidar & José Guimón, "The Future of Artificial Intelligence: Insights from Recent Delphi Studies", 165 *Futures* 2025 https://www.sciencedirect.com/science/article/pii/S0016328724001976.

[157] Soice E, Rocha R, Cordova K, Specter M, Esvelt K, *Can large language models democratize access to dual-use biotechnology*, 1 https://arxiv.org/ftp/arxiv/papers/2306/2306.03809.pdf

[158] Kiyemba Edris, E.K., "Utilisation of Artificial Intelligence and Cybersecurity Capabilities: A Symbiotic Relationship for Enhanced Security and Applicability", 14(10) *Electronics*, 2025 https://www.mdpi.com/2079-9292/14/10/2057#:~:text=Artificial%20Intelligence%20(AI)%20is%20used,systems%20and%20mitigating%20their%20risks.

[159] Alexandros Gazos, James Kahn, Isabel Kusche, Christian Büscher & Markus Götz, "Organising AI for Safety: Identifying Structural Vulnerabilities to Guide the Design of AI-Enhanced Socio-Technical Systems", 184 *Safety Science*, 2025 https://www.sciencedirect.com/science/article/pii/S0925753524003217.

[160] Christopher Covino & Shaun Ee, *Policy Actions for Enabling Cyber Defense Through Differential Access* Institute for AI Policy and Strategy, 2025 https://www.iaps.ai/research/policy-actions-for-enabling-cyber-defense-through-differential-access.

sophisticated access controls face challenges from reverse engineering, model extraction attacks, and the rapid pace of capability development that can quickly outdate control mechanisms.[161]

This creates genuine policy dilemmas for benefit-sharing mechanisms that existing governance frameworks struggle to address. Traditional non-proliferation regimes assume clear boundaries between civilian and military applications. For instance, the International Atomic Energy Agency's approach[162] works partly because civilian nuclear reactors and weapons programs require different materials and processes. AI systems resist such clean distinctions, leading some scholars to propose novel governance approaches like "differential access",[163] or “structured access"[164] that could evolve with technological development. However, these remain largely conceptual while the need for policymakers to make tradeoffs between security concerns that favor restrictive access and equity concerns that favor broad diffusion, continues to rise.

Tension 4: Sovereignty vs. Coordination

Effective benefit-sharing may require unprecedented levels of international coordination, yet states seek to safeguard their sovereignty over their own domestic economic and social policy. This tension is particularly acute for Global Majority states concerned about neo-imperial impositions of external governance frameworks.

The central challenge is to chart a path that allows states to preserve sovereignty while also participating in the level of international coordination necessary to share the benefits of AI effectively.[165] The question is not simply whether states should cooperate, but to what extent sovereignty is possible in a domain where the capacity to develop frontier AI lies almost exclusively with a handful of countries and corporations.

In AI governance, sovereignty is often defined as a state’s ability to control its data flows, set its own regulatory priorities, and shape its domestic technological ecosystem without undue external interference.[166] Sovereignty-first thinkers raise three primary concerns. First, deepening reliance on foreign AI infrastructure risks eroding sovereignty, especially for states dependent on

---

[161] Kaixiang Zhao, Lincan Li, Kaize Ding, Neil Zhenqiang Gong, Yue Zhao & Yushun Dong, “A Survey on Model Extraction Attacks and Defenses for Large Language Models” *arXiv* 2025 https://arxiv.org/pdf/2506.22521.
[162] For example, different resources are required making access to nuclear technologies contingent on safety capacity: Technical Co-operation Programme, Co-ordinated Research Programme, Nuclear Safety Standards Programme, Guidebook on the introduction of nuclear power | IAEA (1982), Manpower development for nuclear power | IAEA (1980) Guidebook, Building Capacity for Nuclear Security.
[163] Christopher Covino & Shaun Ee, *Policy Actions for Enabling Cyber Defense Through Differential Access* Institute for AI Policy and Strategy, 2025 https://www.iaps.ai/research/policy-actions-for-enabling-cyber-defense-through-differential-access.
[164] Toby Shevlane, “Structured Access: An Emerging Paradigm for Safe AI Deployment” *arXiv*, 2022, https://arxiv.org/pdf/2201.05159.
[165] Ishkhanyan, A, “The sovereignty–internationalism paradox in AI governance: digital federalism and global algorithmic control”, 5 *Discover Artificial Intelligence*, 2025 https://link.springer.com/article/10.1007/s44163-025-00374-x.
[166] Yu Chen, *AI Sovereignty: Navigating the Future of International AI Governance*, 2024 https://philarchive.org/rec/CHEASN-2.

imported compute, cloud services, or model access.[167] Second, international frameworks may be shaped by powerful states and corporations, embedding their norms into "global" rules in ways that constrain weaker states' policy space.[168] Third, harmonised regulatory frameworks can reduce regulatory diversity, imposing rigid one-size-fits-all approaches that may not fit local contexts.[169] These challenges are difficult to address because the technological capacity gap between leading AI powers and the rest of the world is vast, and the ability to develop frontier AI independently is at present beyond reach for most states.

Coordination in this context means the collective development of governance frameworks, standards, and cooperative mechanisms to manage AI's risks and distribute its benefits.[170] It is desirable for reasons few contest: AI's challenges are inherently cross-border.[171] Even those advocating for sovereignty often accept that without some form of coordination, many states will struggle to access the full benefits of AI.[172] Yet, coordination carries risks that mirror sovereignty advocates' concerns: it can deepen dependency, entrench the dominance of current AI powers under the guise of capacity-building, and formalise corporate capture by giving multinational technology firms a permanent role in governance, often at the expense of democratic oversight.[173]

The main tension here is the seeming difficulty in optimising both coordination on widespread AI access and strengthening state sovereignty due to the geopolitical context where only a few countries have the resources to build frontier AI technologies and thus the power to dictate the rules of global diffusion.[174] The governance challenge lies in identifying the points of strategic coordination where states can balance their sovereignty with advancing AI benefit-sharing. For Global Majority countries, those with low existing leverage over AI development and the

---

[167] Pankaj Pandey, "Digital Sovereignty and AI: Developing India's National AI Stack for Strategic Autonomy", 254 *Procedia Computer Science*, 2025, 250 https://www.sciencedirect.com/science/article/pii/S187705092500434X.
[168] Nico Krisch, "International Law in Times of Hegemony: Unequal Power and the Shaping of the International Legal Order", 16 *European Journal of International Law* 2005, 369 https://www.ejil.org/pdfs/16/3/301.pdf.
[169] Ishkhanyan, Artur, "The sovereignty–internationalism paradox in AI governance: digital federalism and global algorithmic control", 5 *Discover Artificial Intelligence*, 2025 https://link.springer.com/article/10.1007/s44163-025-00374-x.
[170] Ishkhanyan, A, "The sovereignty–internationalism paradox in AI governance: digital federalism and global algorithmic control", 5 *Discover Artificial Intelligence*, 2025 https://link.springer.com/article/10.1007/s44163-025-00374-x.
[171] Mira Lane & Stacey King, *Common Goals and Cooperation – Towards Multi-Stakeholderism in AI*, in Alex Krasodomski (ed.), *Artificial Intelligence and the Challenge for Global Governance* 58 Chatham House Digital Society Initiative, 2024 https://www.chathamhouse.org/2024/06/artificial-intelligence-and-challenge-global-governance/09-common-goals-and-cooperation.
[172] The White House, *America's AI Action Plan*, 2025 https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf.
[173] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5, 2024 https://hdsr.mitpress.mit.edu/pub/xdukxlpp.
[174] United Nations Conference on Trade and Development, *AI at the Technology Frontier*, in *Technology and Innovation Report 2025: Inclusive Artificial Intelligence for Development*, Chapter I, 2025, 10-12 https://unctad.org/system/files/official-document/tir2025ch1_en.pdf.

pre-existing power dynamics affecting their favours, the sovereignty-coordination concerns are disproportionately acute.[175]

Tension 5: Innovation incentives vs. Distributional imperatives

Benefit-sharing mechanisms that significantly redistribute value away from developers risk undermining innovation incentives, however, approaches that preserve existing distributional patterns fail to address equity concerns that motivate benefit-sharing commitments.

The fundamental tension between promoting AI innovation and ensuring equitable distribution has evolved beyond traditional market failure analyses to encompass global development imperatives and democratic participation rights. Academic research reveals this tension as increasingly more complex, with innovation incentives potentially undermining both distributional justice and long-term safety.

Industry responses have become more sophisticated but remain insufficient in scope. The Partnership for Global Inclusivity on AI launched in 2024 represents an unprecedented private sector commitment to addressing global AI access gaps, with Microsoft investing $12+ billion in Global South AI infrastructure while Meta contributes $10+ million for open-source innovation.[176] These initiatives generate significant multiplier effects where every $1 spent on AI solutions produces $4.9 in global economic value according to industry data but scale remains inadequate relative to global needs and the benefits remain highly concentrated.[177] Crucially, financial investment must be complemented by capacity-building and governance measures that address sociotechnical challenges, including gaps in local expertise, inadequate digital infrastructure, fragmented regulatory frameworks, and uneven access to high-quality datasets, all of which can pose significant barriers to innovation.[178]

---

[175] Sumaya N. Adan, Robert Trager, Kayla Blomquist, Claire Dennis, Gemma Edom, Lucia Velasco, Cecil Abungu, Ben Garfinkel, Julian Jacobs, Chinasa T. Okolo, Boxi Wu & Jai Vipra, *Voice and Access in AI: Global AI Majority Participation in Artificial Intelligence Development and Governance*, 2024 https://aigi.ox.ac.uk/publications/voice-and-access-in-ai-global-ai-majority-participation-in-artificial-intelligence-development-and-governance/.

[176] U.S. Department of State. *United States and Eight Companies Launch the Partnership for Global Inclusivity on AI: Fact Sheet*, 2024 https://2021-2025.state.gov/united-states-and-eight-companies-launch-the-partnership-for-global-inclusivity-on-ai/.

[177] For example: Alysa Taylor, *AI-Powered Success—with More Than 1,000 Stories of Customer Transformation and Innovation*, Microsoft Cloud Blog, 2025, https://www.microsoft.com/en-us/microsoft-cloud/blog/2025/07/24/ai-powered-success-with-1000-stories-of-customer-transformation-and-innovation/.

[178] United Nations Conference on Trade and Development, *Technology and Innovation Report 2025: Inclusive Artificial Intelligence for Development*, Chapter III: "Preparing to Seize AI Opportunites", 2025, 79, https://unctad.org/system/files/official-document/tir2025ch5_en.pdf.

The "compute divide" between developed and developing nations creates new forms of technological dependency that complicate traditional innovation-distribution frameworks.[179] Academic research documents "Compute Desert" countries with limited AI access, while export controls on semiconductors exacerbate these disparities.[180] This dynamic transforms innovation-distribution tensions from primarily domestic concerns to international development challenges requiring unprecedented coordination.

Open-source versus proprietary model debates reflect deeper tensions about innovation control and access. Open-source models, where code, and in some cases full model weights, are made freely available, can accelerate research, lower barriers, and broaden participation, shifting control outward to developers and users worldwide. They redistribute value more widely, but in doing so may weaken firms' incentives to invest and create new risks: once weights are public, they can be replicated, modified, and misused without oversight. The UK AI Security Institute warns that once weights are public they can be replicated, modified, and misused without oversight;[181] technical research highlights vulnerabilities such as data leakage, backdoors, and weight poisoning[182], while others[183] point to their potential for automating cyberattacks. Proprietary models, by contrast, retain control within a small number of firms. This concentration allows stronger safeguards, clearer liability, and the resources needed to advance frontier capabilities, but at the cost of limiting access and reinforcing existing inequities.

[179] United Nations, *Mind the AI Divide: Shaping a Global Perspective on the Future of Work,* 2024, 13 https://www.un.org/digital-emerging-technologies/sites/www.un.org.techenvoy/files/MindtheAIDivide.pdf.
[180] Vili Lehdonvirta, Bóxī Wú & Zoe Hawkins, "Compute North vs. Compute South: The Uneven Possibilities of Compute-Based AI Governance Around the Globe", 7(1) *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2024 https://ora.ox.ac.uk/objects/uuid:6306c118-58ca-49ba-b7a0-4ee7b9423d5a.
[181] AI Security Institute. *Managing risks from increasingly capable open-weight AI systems*, 2025 https://www.aisi.gov.uk/blog/managing-risks-from-increasingly-capable-open-weight-ai-systems.
[182] For example: Younis Al-Kharusi, Amjal khan, Muhammad Rizwan & Mohammed M. Bait-Suwailam, "Open-Source Artificial Intelligence Privacy and Security: A Review", 13(12), 311 *Computers*, 2024 https://www.mdpi.com/2073-431X/13/12/311; Dominik Hintersdorf, Lukas Struppek & Kristian Kersting, "Balancing Transparency and Risk: The Security and Privacy Risks of Open-Source Machine Learning Models", *arXiv*, 2023 https://arxiv.org/abs/2308.09490.
[183] For example: Alfonso de Gregorio, "Mitigating Cyber Risk in the Age of Open-Weight LLMs: Policy Gaps and Technical Realities", *arXiv, 2025* https://arxiv.org/abs/2505.17109.

Underlying Tensions of AI Benefit-Sharing	
Speed vs. Deliberation	AI development is advancing at unprecedented speed while institutional reform and governance mechanisms move slowly. This temporal mismatch creates significant challenges for designing and implementing inclusive and equitable AI access policies.
Private AI development vs. Public governance	The market incentives driving today a small number of private firms at the frontier of AI development may conflict with states' responsibility to safeguard the public interest and values such as transparency, accountability, equity, or risk mitigation. This divergence creates tensions between commercial objectives and the imperatives of public governance.
Non-proliferation vs. Widespread AI access	The dual-use nature of AI technologies creates a tension between preventing dangerous proliferation and enabling beneficial access. While traditional non-proliferation regimes assume clear boundaries between civilian and military applications, AI systems resist such clean distinctions.
Sovereignty vs. Coordination	The capacity to develop frontier AI lies almost exclusively with a handful of countries and corporations who, as a result, claim the power to dictate the rules of global diffusion. In this geopolitical context, Global Majority states face the challenge of balancing coordination for widespread AI access with strengthening their sovereignty.
Innovation incentives vs. Distributional imperatives	Governing AI distribution requires balancing competing imperatives. On one hand, aggressive AI benefit-sharing risks undermining the innovation incentives that drive AI development. On the other hand, allowing current concentration patterns to persist may fail to motivate cooperation on substantive benefit-sharing commitments."

Table 5: Summary of the tensions underlying AI benefit-sharing.

This chapter has laid out the distinct perspectives on AI access. Together, they help understand the varying logics and implications of the different AI benefit-sharing components examined in the remaining chapters. The clarification of existing tensions also highlights the evolving challenges that require strategic navigation.

Part II: Operationalising AI Benefit-Sharing

Part I argues that AI benefit-sharing has been systematically neglected, sidelined by both market and geopolitics and AI race imperatives. Beyond political goodwill, benefit-sharing requires mechanisms, including rules, institutions, fiscal frameworks, and international arrangements that translate abstract commitments into enforceable outcomes.[184] This Part provides a structured toolkit for operationalisation.

The analysis proceeds by recognising that **benefit-sharing** is not a singular logic but an integration of three distinct traditions of governance: **redistribution**, which ensures that economic gains are spread across societies; **technology transfer and capacity-building**, which enables states to develop and govern AI themselves; and **non-proliferation and safety**, which manages the security risks of advanced systems while allowing inclusion. Each tradition generates its own repertoire of mechanisms, but they cannot function in isolation. Benefit-sharing requires hybrid approaches that combine redistribution, transfer, and safety in mutually reinforcing ways.

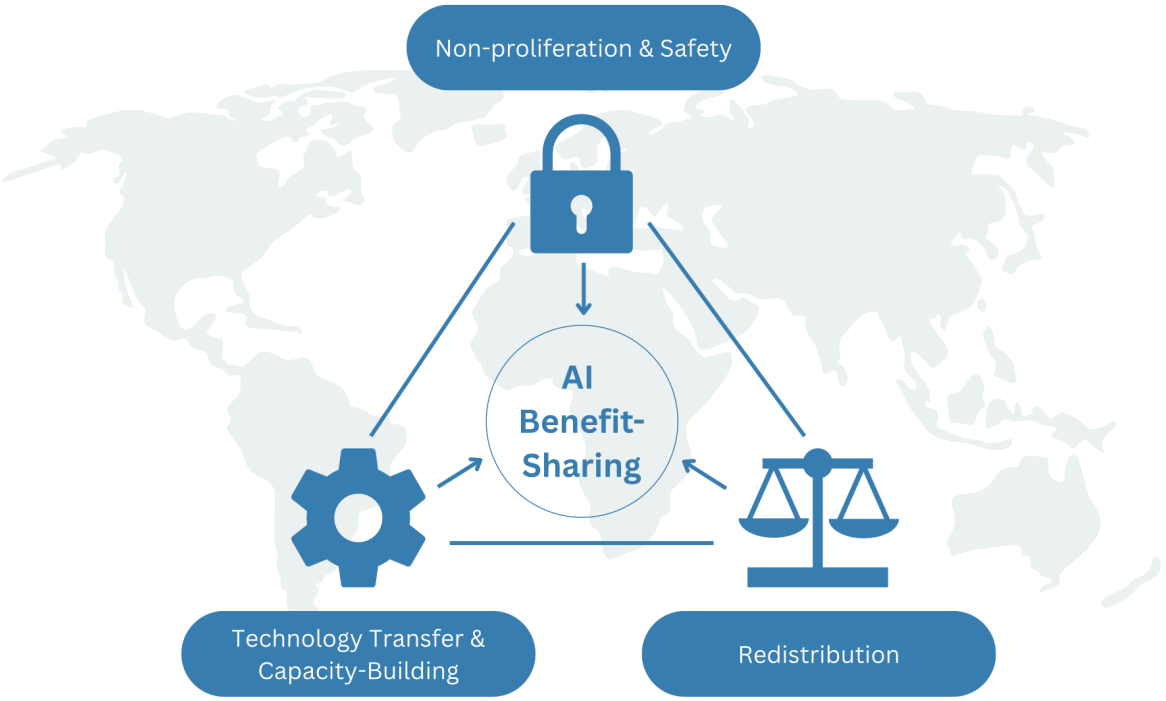


Image 2: AI benefit-sharing components.

Before setting out these mechanisms, however, it is necessary to establish what states should have in order to be in a position where they could harness AI benefits and systematically access

[184] Elisa Morgera, "The Need for an International Legal Concept of Fair and Equitable Benefit Sharing", *European* 27(2) *Journal of International Law*, 2016, 353 https://academic.oup.com/ejil/article/27/2/353/1748393.

its transformative powers. The mechanisms presuppose certain baseline capacities, without which efforts risk being purely symbolic or producing uneven outcomes.[185] The next chapter therefore outlines the **prerequisites for readiness**, providing a practical lens through which low and middle income countries can assess their preparedness for harnessing AI benefits and accessing its transformative powers.

[185] Claire Dennis et al., *Options and Motivations for International AI Benefit Sharing*, Centre for the Governance of AI, 2025, 20 https://www.governance.ai/research-paper/options-and-motivations-for-international-ai-benefit-sharing.

Chapter 3: Prerequisites and Readiness

Operationalisation depends not only on global negotiations or corporate commitments, but also on **domestic readiness**. The sustainability and effectiveness of benefit-sharing initiatives in states with limited AI capacities will depend on their strength of fiscal, infrastructural, and regulatory foundations which determine how well international frameworks can be translated into tangible local benefits. Where these prerequisites are absent, benefit-sharing mechanisms risk entrenching dependency rather than expanding capacity.[186] Given the readiness asymmetries sustained by broader power imbalances in the global order, applying the principle of *common but differentiated responsibilities* places an onus on frontier AI firms and their home states to support the readiness of Global Majority countries in parallel with those countries' own capacity-building efforts. Three categories of readiness correspond to the three components of benefit-sharing:

3.1 Redistribution readiness

For redistribution mechanisms to function, states require:

- **Fiscal capacity:** reliable taxation and revenue-collection systems that can capture AI-derived value (e.g. corporate taxation, VAT systems).
- **Distributional channels:** functioning welfare institutions, public investment vehicles, special economic zones, or sovereign funds capable of redistributing revenues transparently.
- **Data infrastructure:** the ability to collect, store, and analyse socio-economic data to target redistribution and monitor equity outcomes.

Rationale: Without these foundations, instruments such as AI taxes or data dividends cannot achieve their intended purpose. Historical evidence from extractive industries shows that resource rents often exacerbate inequality where fiscal and governance structures are weak.[187]

3.2 Technology transfer and capacity-building readiness

For technology transfer to succeed, states need:

---

[186] Noam Unger & Madeleine McLean, *An Open Door: AI Innovation in the Global South amid Geostrategic Competition,* Center for Strategic and International Studies 2025, 6 https://www.csis.org/analysis/open-door-ai-innovation-global-south-amid-geostrategic-competition

[187] Arezki, Rabah & Markus Brückner. *Rents, Corruption, and State Stability: Evidence From Panel Data Regressions*. IMF Working Paper No. 09/267, International Monetary Fund, 2009 https://www.elibrary.imf.org/view/journals/001/2009/267/article-A001-en.xml: Pr Atangana Ondoa Henri, "Natural Resources Curse: A Reality in Africa", 63 *Resources Policy,* 2019 https://www.sciencedirect.com/science/article/abs/pii/S0301420719300388.

- **Absorptive capacity:** universities, research institutes, and training programmes able to incorporate and apply new technologies. Beyond these traditional institutions, states can also cultivate absorptive capacity through designated special economic zones,[188] science parks, or innovation districts.[189] By concentrating infrastructure, talent, and core institutions in a single, well-governed location,[190] these zones establish a core foundation that could enable the diffusion of knowledge and innovation across the entire state.
- **Infrastructure:** reliable energy supply, data connectivity, and compute infrastructure to host and use transferred technologies.
- **Legal and regulatory frameworks:** for example, IP systems that allow technology licensing and adaptation, while preventing misuse, regulatory sandboxes, data governance frameworks.
- **Talent pipelines:** STEM education and fellowship programmes that ensure knowledge transfer does not dissipate into brain drain.

Rationale: Without absorptive capacity, transferred technologies often remain underutilised. The history of climate technology transfer under the UNFCCC demonstrates that technical assistance without domestic absorptive capacity leads to low uptake and dependency.[191]

3.3 Non-proliferation and safety readiness

To manage safe access regimes, states need:

- **Regulatory bodies:** authorities empowered to oversee AI deployment, enforce licensing conditions, and manage export controls.
- **Cybersecurity and monitoring capacity:** the technical means to prevent attacks and theft of sensitive models and verify compliance with safety protocols.
- **Legitimacy mechanisms:** public accountability frameworks to ensure restrictions are not viewed as neo-imperial exclusion, particularly in Global Majority contexts.
- **Trust-building participation:** the ability to join international monitoring regimes and exchange information credibly.

Rationale: Without baseline safety capacity, non-proliferation regimes risk locking Global Majority states out of advanced AI altogether. Precedents from nuclear governance show that

[188] Douglas Zhihua Zeng, *Global Experiences with Special Economic Zones: With a Focus on China and Africa*, World Bank 2015, https://www.worldbank.org/content/dam/Worldbank/Event/Africa/Investing%20in%20Africa%20Forum/2015/investing-in-africa-forum-global-experiences-with-special-economic-zones-with-a-focus-on-china-and-africa.pdf.
[189] Bruce Katz & Julie Wagner, *The Rise of Innovation Districts: A New Geography of Innovation in America*, Brookings Institution, 2014, https://www.brookings.edu/wp-content/uploads/2016/07/InnovationDistricts1.pdf.
[190] Michael E. Porter, "Clusters and the New Economics of Competition" 76 *Harvard Business Review* 1998, 77, https://hbr.org/1998/11/clusters-and-the-new-economics-of-competition.
[191] Damilola S. Olawuyi,"From Technology Transfer to Technology Absorption: Addressing Climate Technology Gaps in Africa" 36 *Journal of Energy & Natural Resources Law* 2017, https://www.tandfonline.com/doi/full/10.1080/02646811.2017.1379667.

access restrictions without capacity-building foster enduring technological inequality and political resentment.[192]

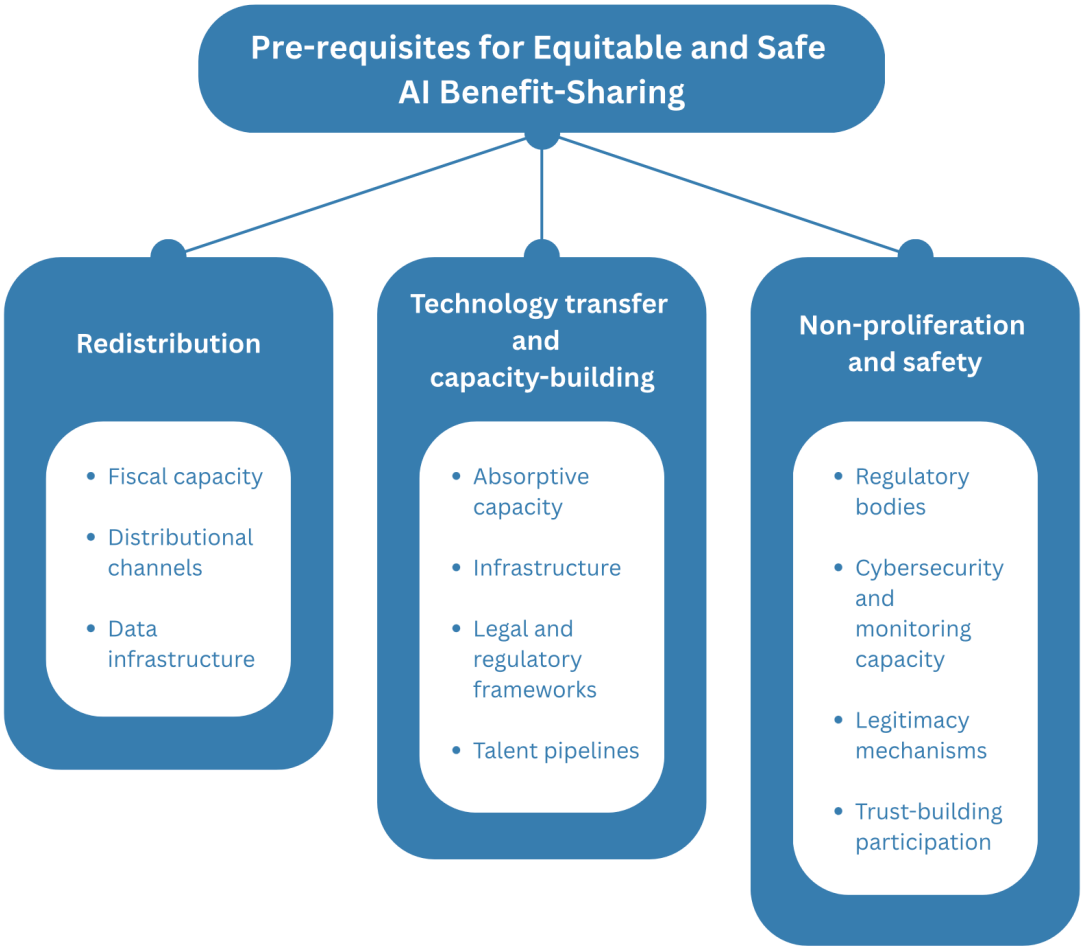


Image 3: Summary table of the prerequisites for equitable and safe AI benefit-sharing.

[192] Aslı Bâli & Aziz Rana, "Unequal Power and the Institutional Design of Global Governance: The Case of Arms Control", 40 *Review International Studies,* 2014, 9 https://www.researchgate.net/publication/286011111_Unequal_power_and_the_institutional_design_of_global_governance_The_case_of_arms_control: *The Japan Institute of International Affairs, Nuclear Proliferation and the Incentive Approach,* JIIA Research Report, 2010 https://www2.jiia.or.jp/report/j-report/100428-Nuclear_Proliferation.html.

Chapter 4: Mechanisms for Redistribution through Rules and Laws

As states establish key fiscal and institutional foundations, they can begin deploying mechanisms to ensure that the economic gains from AI are distributed across society. The most direct route is through **rules and legal frameworks**. These mechanisms embed benefit-sharing into the regulatory fabric of AI governance, making redistribution not a matter of voluntary goodwill but a legal and enforceable obligation.

4.1 Regulatory frameworks

Several regulatory tools can be adapted from existing domains to govern AI in ways that prioritise redistribution:

- **AI licensing and authorisation regimes:** States can condition market access for AI developers on explicit benefit-sharing obligations. AI licensing could require companies to demonstrate how their products contribute to equitable outcomes, for example, through open access to models with safeguards, discounted services for public institutions, or mandatory financial contributions to sovereign technology funds.
- **Mandatory benefit assessment requirements:** Drawing on analogies from environmental impact assessments (EIAs) and social impact assessments (SIAs), states could require that frontier AI projects submit benefit assessments before deployment. These assessments would identify who is likely to gain and who might be excluded, providing regulators with evidence to condition approvals or mandate remedial measures.
- **Distributional impact analysis obligations:** Beyond *ex ante* assessments, firms could be required to conduct regular distributional impact audits, tracking whether benefits (e.g. productivity gains, cost reductions, or public service improvements) are accruing beyond concentrated user groups. This provides governments with a monitoring tool to adjust taxation, procurement, or subsidy regimes accordingly.
- **Compliance and enforcement mechanisms:** To avoid regulatory capture, these frameworks must include independent enforcement authorities with the power to levy fines, revoke licenses, or impose corrective measures. Evidence from competition law suggests that without strong enforcement, formal obligations are easily circumvented by well-resourced actors.
- **SEZ-based benefit-sharing mandates:** States could establish Special Economic Zones (SEZs) through enabling laws that set the legal status, governance, and obligations of zone authorities, developers, and users.[193] Unlike economy-wide regulatory measures, SEZs function as institutional containers where fiscal, legal, and governance arrangements are bundled and tested in bounded geographies. Applied to AI, SEZ legislation could tie regulatory flexibility to public-interest mandates, such as compulsory

[193] United Nations Conference on Trade and Development, *World Investment Report 2019: Special Economic Zones* 2019, 161–176 https://unctad.org/system/files/official-document/wir2019_en.pdf.

skills transfer, supplier development, or earmarking tax incentives for national infrastructure. This design ensures that the legal autonomy granted for innovation is not a free-floating privilege but a structured bargain, where concentrated benefits are redirected into systemic national development.

- **Human rights due diligence:** Finally, benefit-sharing can be framed within human rights due diligence requirements. Under frameworks such as the UN Guiding Principles on Business and Human Rights, firms are already expected to prevent and mitigate harms.[194] The UN Global Digital Compact also emphasizes the centrality of adherence with international human rights as a foundation to digital cooperation.[195] Extending these principles to include distributive justice would push firms to proactively demonstrate how their business models promote broad social gains rather than exacerbate exclusion.

4.2 Fiscal rules and redistribution mechanisms

Legal frameworks create the conditions for redistribution, but fiscal tools are the primary levers through which states can directly capture and reallocate AI-derived value. Effective fiscal mechanisms ensure that economic rents generated by frontier AI systems are not captured solely by a narrow group of firms and investors but are instead channelled into broad-based public benefit.

I. Taxation of AI-derived value

- **Corporate taxation and windfall levies:** Governments may impose targeted taxes on the extraordinary profits generated by AI companies, akin to windfall taxes on extractive industries. Such levies could be structured progressively, ensuring that firms benefiting disproportionately from frontier models contribute proportionately to public revenues.[196]
- **Value-added tax (VAT) adjustments:** Given that AI systems often replace labour inputs, VAT frameworks may need to be adapted to capture value added through automation and digital services. Without such adjustments, states risk an erosion of their tax base.[197]
- **Data dividend or usage fees:** Inspired by resource royalties, states could require companies monetising local data to pay a "data dividend" into sovereign funds.[198] This

[194] United Nations, *Guiding Principles on Business and Human Rights: Implementing the United Nations 'Protect, Respect and Remedy' Framework,* 2011, https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf.
[195] United Nations, *Global Digital Compact*, Annex I to the *Pact for the Future*, adopted at the Summit of the Future, New York, 2024 https://www.un.org/global-digital-compact/sites/default/files/2024-09/Global%20Digital%20Compact%20-%20English_0.pdf.
[196] Cullen O'Keefe et al., *The Windfall Clause: Distributing the Benefits of AI for the Common Good,* Centre for the Governance of AI 2020 https://www.governance.ai/research-paper/the-windfall-clause-distributing-the-benefits-of-ai-for-the-common-good.
[197] Michael J. Ahn, *Navigating the Future of Work: A Case for a Robot Tax in the Age of AI*, Brookings Institution, 2024 https://www.brookings.edu/articles/navigating-the-future-of-work-a-case-for-a-robot-tax-in-the-age-of-ai/.
[198] Nicholas Vincent, Yichun Li, Renee Zha & Brent Hecht, "Mapping the Potential and Pitfalls of 'Data Dividends' as a Means of Sharing the Profits of Artificial Intelligence", *arXiv* 2019, https://arxiv.org/pdf/1912.00757.

recognises data as a collective asset and ensures local populations benefit from its exploitation.

II. Revenue distribution mechanisms

- **Sovereign technology funds:** Revenues from AI taxation can be channelled into sovereign funds dedicated to long-term investment in infrastructure, education, and social protection.[199] This mirrors models such as Norway's oil fund, which has converted resource rents into intergenerational wealth.
- **Targeted social transfers**: Direct redistribution mechanisms, such as cash transfers, universal service subsidies or wage top-ups, can ensure that AI-induced dislocations are offset by tangible benefits. For instance, data dividends could be partly disbursed as direct cash payments to citizens.[200]
- **Public service subsidisation:** AI revenues can also be earmarked to subsidise digital public goods, such as AI-enabled education platforms, healthcare diagnostics, or climate resilience systems. This ensures that redistribution is not purely monetary but also functional, embedding AI gains in public services.

III. Procurement requirements with benefit-sharing conditions

Public procurement is a powerful fiscal tool. By embedding benefit-sharing clauses in contracts, states can ensure that AI vendors provide discounted access, training, or infrastructure support as a condition for supplying government services. This approach mirrors "local content" requirements in natural resource contracts, where foreign firms are obligated to contribute to domestic capacity-building.[201]

IV. Redistributive SEZ fiscal regimes

Far from marginal experiments, SEZs have become a mainstream industrial policy tool: UNCTAD counts more than 5,400 across 147 countries, and China alone hosts over half.[202] The most famous case, Shenzhen, grew from a fishing village into a US$510 billion technology hub employing over 11 million people and contributing roughly 3% of China's GDP.[203] The lesson is

---

[199] Liam Epstein, *Lead, Own, Share: Sovereign Wealth Funds for Transformative AI* Convergence Fellowship Program, 2025, 28 https://drive.google.com/file/d/1pEZmGxq77ZkywXXg9_nUlkbjVa4uHv6x/view.

[200] Liam Epstein, *Lead, Own, Share: Sovereign Wealth Funds for Transformative AI* Convergence Fellowship Program, 2025, 36 https://drive.google.com/file/d/1pEZmGxq77ZkywXXg9_nUlkbjVa4uHv6x/view; Cullen O'Keefe et al., *The Windfall Clause: Distributing the Benefits of AI for the Common Good* Centre for the Governance of AI 2020, 31 https://www.governance.ai/research-paper/the-windfall-clause-distributing-the-benefits-of-ai-for-the-common-good.

[201] Tony Addison & Alan Roe (eds), *Extractive Industries: The Management of Resources as a Driver of Sustainable Development* Oxford 2018, 511 https://academic.oup.com/book/27405.

[202] As Zeng (2021) emphasizes in a recent survey of the field, SEZs have shifted from being marginal "enclaves" to becoming mainstream policy tools. His article traces their policy rationales, governance models, and evolving impacts, underscoring how well-designed SEZs can serve national development goals rather than operate in isolation.

[203] Shenzhen's 2024 GDP reached 3.68 trillion yuan (≈US$510 billion), representing approximately 3% of China's national nominal GDP. The city reported 11.7 million employed persons as of June 2020. See "Shenzhen hits

that when governments align policy experimentation with clear incentives and legal guarantees, zones can accelerate structural transformation.[204]

For AI, the same template can be adapted to ensure redistribution: preferential tax breaks, royalty schemes, or procurement rights could be tied to demonstrable benefit-sharing, with zone revenues earmarked for sovereign technology funds, targeted transfers, or domestic innovation funds.[205] Properly designed, such SEZ fiscal regimes align investment de-risking with national development goals.

V. Public investment guidelines with distributional provisions

Governments that are already investing substantial sums towards the development of their AI industries can condition public R&D investments in AI on demonstrable social returns. For instance, state-funded AI research could be required to publish outputs as open-source models, or to allocate a share of intellectual property to public institutions. These provisions align public investment with long-term distributional goals.

Historical precedent shows that without fiscal governance, resource rents often exacerbate inequality and corruption.[206] The "resource curse" literature illustrates how extractive industries have entrenched elite capture in contexts lacking strong fiscal and institutional frameworks. Conversely, models such as Norway's sovereign fund[207] or Botswana's diamond revenue governance[208] show that carefully designed fiscal rules can convert rents into broad-based public benefit. In the AI context, the risk of elite capture is even greater given the concentration of frontier capabilities in a handful of firms and jurisdictions.[209]

---

national firsts in 2024, targets 5.5% GDP growth in 2025," *ECNS*, Feb 27, 2025; and "City has over 11 million employed people," *Shenzhen Government Online*, Aug 11, 2020.

[204] Farole and Akinci's 2011 World Bank volume remains the definitive handbook on SEZ policy design. It details the legal and fiscal frameworks through which states establish zones, as well as governance and performance criteria, and shows how governments can channel zone-generated rents toward industrial upgrading and broader development objectives.

[205] Said Saillant, *From Tax Holidays to Tech Havens: A Playbook for AI-Ready Special Economic Zones in Latin America*, Working paper, 2025.

[206] Arezki, Rabah & Markus Brückner. *Rents, Corruption, and State Stability: Evidence From Panel Data Regressions*. IMF Working Paper No. 09/267. International Monetary Fund, 2009 https://www.elibrary.imf.org/view/journals/001/2009/267/article-A001-en.xml

[207] Einar Lie, "Learning by Failing: The Origins of the Norwegian Oil Fund", 43(11) *Scandinavian Journal of History,* 2018 https://www.researchgate.net/publication/322876420_Learning_by_Failing_The_Origins_of_the_Norwegian_Oil_Fund.

[208] Maria Sarraf & Moortaza Jiwanji, *Beating the Resource Curse: The Case of Botswana* World Bank, October 2001, 12 https://openknowledge.worldbank.org/bitstreams/a4c0dd5c-2155-5504-ba0a-aecd92a2093c/download.

[209] David Leslie et al., "'Frontier AI,' Power, and the Public Interest: Who Benefits, Who Decides?," *Harvard Data Science Review*, Special Issue 5, 2024, https://hdsr.mitpress.mit.edu/pub/xdukxlpp.

4.3 Institutional capacity for redistribution

Rules and fiscal tools provide the legal and financial foundations for redistribution, but their effectiveness ultimately depends on the **institutions tasked with implementation and oversight**. Weak or fragmented institutions risk turning benefit-sharing into symbolic commitments, whereas strong, well-coordinated institutions can translate revenues and rules into tangible outcomes. Three levels of institutional capacity are central: national governance bodies, fiscal institutions, and international coordination.

I. National governance bodies

- **Specialised benefit-sharing authorities:** States may establish independent agencies tasked with monitoring AI's economic and social impacts, conducting benefit assessments, and enforcing redistribution provisions. Such authorities would function analogously to environmental protection agencies, with a mandate to evaluate compliance and propose corrective action.
- **SEZ oversight bodies:** National SEZ authorities act as the statutory gatekeepers for zone development.[210] They license zones, approve tenant projects, and evaluate whether proposals align with national endowments and priority segments of global AI value chains. Beyond approvals, these bodies monitor zone performance against developmental benchmarks, such as job creation, technology transfer, or export diversification, ensuring that fiscal and regulatory privileges deliver measurable national benefits.
- **Multi-stakeholder governance structures:** Involving civil society, industry, and academia in advisory or oversight councils enhances legitimacy and reduces the risk of regulatory capture. Experience from extractive industry transparency initiatives suggests that well-coordinated multi-stakeholder involvement can improve compliance and public trust.[211]
- **Coordination mechanisms across agencies:** Given AI's cross-sectoral impact, ministries of finance, science and technology, labour, and education must coordinate benefit-sharing measures. Without such horizontal coordination, fiscal tools risk working at cross-purposes or failing to target priority areas.

II. Fiscal institutions

- **Sovereign technology funds**: Establishing well-governed sovereign funds allows revenues from AI taxation or royalties to be pooled and managed transparently.

---

[210] Institutional arrangements vary: some states centralize oversight nationally, while others delegate authority to sub-national bodies, as in China, where municipal authorities in Shenzhen directed approvals toward electronics manufacturing as part of a broader industrial upgrading strategy. For the variation in institutional arrangements (centralized vs. delegated, one-stop SEZ authorities), see *Special Economic Zones: Progress, Emerging Challenges and Future Directions* (Farole & Akinci, 2011), which discusses SEZ governance models and "one-stop" authorities.

[211] Paul Fenton Villar, "The Extractive Industries Transparency Initiative (EITI) and Trust in Politicians", 68 *Resources Policy*, 2020 https://www.sciencedirect.com/science/article/abs/pii/S0301420720302439.

Governance structures like independent boards, clear investment rules, and regular audits are essential to prevent elite capture.

- **Distribution mechanisms and accountability structures**: Whether through direct transfers or subsidies, fiscal institutions must be equipped with systems for disbursement and accountability. Conditional cash transfer programmes in Latin America demonstrate how robust institutions can channel revenues to households effectively when backed by digital public infrastructure and monitoring.[212]

III. International coordination institutions

- **Alliances of committed states**: Redistribution regimes may be more successful where Global Majority states act collectively. Regional or sub-regional blocs (e.g. the African Union, ASEAN, ECOWAS, Working Group on the Ethics of Artificial Intelligence in Latin America and the Caribbean) could establish minimum standards for taxation and redistribution, preventing a "race to the bottom" in fiscal policy.
- **Regional capacity facilities:** International hubs that pool resources and funding can foster capacity-building and collaboration among stakeholders by offering training for regulators and developing AI governance instruments such as model-evaluation and data-sharing protocols. These institutions can be modeled on the IMF Regional Capacity Development Centers and UNEP regional programs.
- **Technical standards and harmonisation**: International organisations can support harmonisation of benefit-sharing requirements, such as common guidelines for distributional impact assessments or reporting standards for AI revenues. The UK-led AI Standards Hub[213] provides an early example of such efforts.
- **Dispute resolution and compliance mechanisms**: International institutions can provide useful fora for resolving disputes between states and firms, ensuring that benefit-sharing obligations are enforceable beyond domestic jurisdictions.
- **Bilateral SEZs and cross-border corridors:** States can formalise jointly governed zones or corridor arrangements that harmonise customs, fast-track permits, and share revenue/oversight. For AI, such cross-border SEZs could pool investment in compute infrastructure, streamline data-sharing frameworks, and jointly develop regulatory sandboxes.[214]

---

[212] Pablo Ibarrarán, Nadin Medellín, Ferdinando Regalia & Marco Stampini (eds), *How Conditional Cash Transfers Work: Good Practices after 20 Years of Implementation,* Inter-American Development Bank, 2021 https://publications.iadb.org/en/how-conditional-cash-transfers-work.

[213] See https://aistandardshub.org/.

[214] A current template is the Johor–Singapore Special Economic Zone (JS-SEZ), formalised via a January 11, 2024 Memorandum of Understanding between Malaysia's Minister of Economy and Singapore's Ministry of Trade & Industry, designed to streamline cross-border movement of goods, people, and services (*Malaysia: The Johor–Singapore Special Economic Zone (JS-SEZ),* AHK / SGC InsightPlus, 2024). Similarly, a May 30, 2025 MoU between the Government of Ghana and the United Arab Emirates announced a US$1 billion Innovation and Technology Hub near Ningo-Prampram, signalling a bilateral, zone-anchored industrial policy model that could embed joint governance and benefit-sharing (*Ghana, UAE Partner to Establish Innovation and Technology Hub,* Government of Ghana, 2025).

4.4 Building national buy-In

Institutional capacity is not purely technical. Political legitimacy is essential.[215] Public communication campaigns can frame benefit-sharing not as an elite bargain but as a social contract: AI is developed and deployed under conditions that guarantee tangible benefits for all. Without this legitimacy, redistribution risks being undermined by public distrust or populist backlash.

Institutional capacity has been the decisive factor in whether redistribution succeeds or fails in other sectors. The Extractive Industries Transparency Initiative (EITI) illustrates how multi-stakeholder structures can enhance accountability in revenue management.[216] Similarly, sovereign funds demonstrate that fiscal governance can convert volatile rents into stable, long-term benefits but only under strict institutional safeguards. In contrast, weak fiscal institutions in resource-rich states have often led to rent-seeking and inequality.[217] The same dynamics will determine whether AI redistribution mechanisms generate broad-based welfare or reinforce concentration.

[215] J. J. Woo, M. Ramesh & M. Howlett, "Legitimation Capacity: System-Level Resources and Political Skills in Public Policy", 34(3–4) *Policy and Society*, 2015, 275 https://www.tandfonline.com/doi/epdf/10.1016/j.polsoc.2015.09.008?needAccess=true.
[216] Paul Fenton Villar, "The Extractive Industries Transparency Initiative (EITI) and Trust in Politicians", 68 *Resources Policy*, 2020 https://www.sciencedirect.com/science/article/abs/pii/S0301420720302439.
[217] J. Narh, 'The Resource Curse and the Role of Institutions Revisited, Environment', 27 *Development and Sustainability*, 2025, 8190 https://link.springer.com/article/10.1007/s10668-023-04279-6.

Mechanisms for Redistribution through Rules and Laws	
Regulatory frameworks	1. AI licensing and authorisation regimes 2. Mandatory benefit assessment requirements 3. Distributional impact analysis obligations 4. Compliance and enforcement mechanisms 5. Special Economic Zones (SEZ)-based benefit-sharing mandates 6. Human rights due diligence
Fiscal rules and redistribution mechanisms	1. Taxation of AI-derived value a. Corporate taxation and windfall levies b. Value-added tax (VAT) adjustments c. Data dividend or usage fees 2. Revenue distribution mechanisms a. Sovereign technology funds b. Targeted social transfers c. Public service subsidisation 3. Procurement requirements with benefit-sharing conditions 4. Redistributive Special Economic Zones fiscal regimes 5. Public investment guidelines with distributional provisions
Institutional capacity for redistribution	1. National governance bodies a. Specialised benefit-sharing authorities b. Special Economic Zones oversight bodies c. Multi-stakeholder governance structures d. Coordination mechanisms across agencies 2. Fiscal institutions a. Sovereign technology funds b. Distribution mechanisms and accountability structures 3. International coordination institutions a. Alliances of committed states b. Regional capacity facilities c. Technical standards and harmonisation d. Dispute resolution and compliance mechanisms e. Bilateral Special Economic Zones and cross-border corridors

Table 6: Summary of the redistributive AI benefit-sharing mechanisms.

Chapter 5: Mechanisms for Technology Transfer and Capacity-Building

Redistribution mechanisms ensure that the economic rents of AI are more fairly allocated. Yet redistribution alone cannot close global divides in AI. To truly operationalise benefit-sharing, states must also expand their **capacity to develop, govern, and adapt AI technologies themselves**. This requires mechanisms of technology transfer and capacity-building.

The principle here is distinct from redistribution: it is not about reallocating value after the fact but about enhancing inclusion and equipping more states to participate directly in AI development.[218] In this sense, capacity-building is both a form of benefit-sharing and a safeguard against long-term dependency, since it allows countries to generate their own AI capabilities rather than relying exclusively on rents distributed by others.

Historical experience underscores the stakes. In sectors ranging from nuclear energy to pharmaceuticals, access to technology and the absorptive capacity of recipient states has determined whether international agreements entrenched dependency or catalysed independent development. AI will be no different: without credible mechanisms for transfer and diffusion, the Global Majority risks being permanently locked into subordinate roles in the AI economy.

5.1 Infrastructure mechanisms: building the material base

At the core of capacity-building is physical and digital infrastructure: compute, connectivity, and data.[219] Without these, knowledge transfer alone has little impact.

- **Shared compute hubs:** Regional compute centres, established through multilateral funding or public–private partnerships, can pool resources for states lacking frontier-level infrastructure.[220] These centres could be governed through equitable access rules, preventing monopolisation by a single state or corporate actor. Early precedents exist in regional climate data hubs and the EU's EuroHPC initiative.[221] Alternatively, states could establish jointly governed special economic zones or cross-border corridors that pool

[218] Iason Gabriel, "Toward a Theory of Justice for Artificial Intelligence", *arXiv* 2020, 226 https://arxiv.org/pdf/2110.14419.
[219] United Nations Conference on Trade and Development, *Technology and Innovation Report 2025: Inclusive Artificial Intelligence for Development*, Chapter III: "Preparing to Seize AI Opportunites", UNCTAD 2025, 79, https://unctad.org/system/files/official-document/tir2025ch5_en.pdf: Muath Alduhishy, *Sovereign AI: What It Is, and 6 Strategic Pillars for Achieving It*, World Economic Forum, 2024 https://www.weforum.org/stories/2024/04/sovereign-ai-what-is-ways-states-building/.
[220] Jamille Tran, *China's AI Firms Are Going Regional for Compute Power – and South-east Asia Is Cashing In*, The Business Times, 2025, https://www.businesstimesintl.com/asean/chinas-ai-firms-are-going-regional-compute-power-and-south-east-asia-cashing.
[221] European Commission, *Seven Consortia Selected to Establish AI "Factories" to Boost AI Innovation in the EU,* 2024 https://ec.europa.eu/commission/presscorner/detail/en/ip_24_6302

investment in compute infrastructure, harmonise regulatory frameworks, and enable shared oversight and innovation in AI.

- **Cloud credits and subsidised access:** As a transitional measure, companies or consortia can provide subsidised cloud credits earmarked for research institutions and public-interest projects in the Global Majority. While this does not substitute for domestic infrastructure, it lowers entry barriers and creates time for states to build local capacity.
- **Regional data centres and sovereign data governance:** Localised data centres ensure that training data and sensitive information remain under national or regional jurisdiction. Models such as the African Union's Smart Africa initiatives illustrate how pooled investments in infrastructure can align with sovereignty and capacity-building goals.[222] These investments are key to AI sovereignty, enabling the Global Majority to develop regionally tailored AI systems such as LatamGPT, Latin America's first large language model.[223] By establishing robust, context-appropriate data governance frameworks, states can ensure that sensitive datasets remain within regional jurisdictions, promote responsible AI development aligned with local priorities, and reduce reliance on foreign-controlled infrastructure.
- **Connectivity and energy investments:** High-performance computing requires stable electricity grids and connectivity.[224] Without parallel investments in energy and network infrastructure, compute hubs will remain underutilised. Here, AI benefit-sharing must be explicitly tied to broader development finance.

5.2 Knowledge transfer and human capital

Infrastructure alone is insufficient without the human expertise to use it effectively. Absorptive capacity depends critically on investments in knowledge transfer.[225]

- **Training programmes and fellowships:** Frontier states and firms can be obligated to fund training programmes and fellowships, modelled on IAEA fellowships in nuclear sciences or WHO-sponsored training for public health. These schemes must be multi-year and designed to build durable expertise, not one-off exchanges.
- **Joint research labs:** Establishing joint labs between frontier and recipient states, with shared governance, creates channels for both technological diffusion and political

[222] European Union, *Global Gateway: EU and Smart Africa Strengthen Partnership for Africa's Digital Transformation,* 2024 https://international-partnerships.ec.europa.eu/news-and-events/news/global-gateway-eu-and-smart-africa-strengthen-partnership-africas-digital-transformation-2024-12-03_en.
[223] See https://www.latamgpt.org/en
[224] Yihong Zhou, Ángel Paredes Parrilla, Chaimaa Essayeh & Thomas Morstyn, "AI-Focused HPC Data Centers Can Provide More Power Grid Flexibility and at Lower Cost", *arXiv* 2024 https://doi.org/10.48550/arXiv.2410.17435
[225] Sergio Cuéllar, María Teresa Fernández-Bajón & Félix De Moya-Anegón, "Convergence between Absorptive Capacity and Knowledge Appropriation: A New Methodology Mapping the Hidden Links", 10(1) *Journal of Open Innovation: Technology, Market, and Complexity,* 2024 https://www.sciencedirect.com/science/article/pii/S219985312300272X.

legitimacy. Similar cooperation between AI safety institutes across the Global North-South is crucial for democratizing safety standards and enhancing the capacity and autonomy of Global Majority countries to manage AI-related risks. This mirrors past collaborations in space exploration and genomics, where joint governance was essential to credibility.

- **Curriculum development and standards alignment:** Support for local universities and technical institutes to develop AI curricula aligned with international standards is a prerequisite for sustainable capacity-building. Absent this, training pipelines risk being fragmented and misaligned with global practices.

5.3 Legal instruments for technology transfer

Beyond infrastructure and human capital, technology transfer depends on the legal frameworks that govern how knowledge, models, and tools move across borders. Current international regimes, particularly in intellectual property (IP) and trade, are designed to protect ownership rather than promote equitable diffusion.[226] If benefit-sharing is to be credible, these legal architectures must be rebalanced to support transfer while maintaining sufficient safeguards for innovation and security.[227]

I. Intellectual property adaptations

- **Compulsory licensing for AI systems:** Drawing on precedents in pharmaceuticals, compulsory licensing mechanisms could allow states to mandate access to proprietary AI technologies under defined conditions of public interest such as for healthcare, education, or climate adaptation.
- **Adapted copyright for AI-generated content:** AI systems generate outputs that blur traditional categories of authorship. Legal clarity is needed to prevent concentration of rights in the hands of frontier labs while enabling states and communities to claim fair use and derivative rights.
- **Open licensing and shared IP pools:** Frontier actors could be incentivised or required to contribute to shared IP pools, particularly for models and datasets relevant to the Sustainable Development Goals (SDGs). These pools would lower entry barriers for developing states while still protecting commercially sensitive frontier models.

II. Trade and competition law

- **Competition law to prevent market concentration:** Antitrust frameworks can be adapted to prevent dominant AI firms from using IP protections and closed licensing to lock out competitors, especially in smaller markets. This may involve mandating access

[226] B. S. Chimni, "International Institutions Today: An Imperial Global State in the Making", 15(1) *European Journal of International Law*, 2004 https://academic.oup.com/ejil/article/15/1/1/418237.
[227] Simon Chesterman, "Good Models Borrow, Great Models Steal: Intellectual Property Rights and Generative AI", 44(1) *Policy and Society,* 2025 https://academic.oup.com/policyandsociety/article/44/1/23/7606572.

to essential facilities (e.g. APIs, foundational datasets) on fair and non-discriminatory terms.

- **Technology transfer provisions in trade agreements:** Bilateral and regional trade agreements can incorporate binding clauses on technology transfer. Unfortunately, the existing examples[228] reveal a persistent weakness: technology transfers are typically legally unenforceable aspirations. For AI benefit-sharing, this gap is critical because effective technology transfers require explicit enforcement details.

III. International legal regimes

- **TRIPS adjustments for AI:** The Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) has been amended before to enhance the ability of states with limited pharmaceutical manufacturing capacity to utilise TRIPS flexibilities such as compulsory licensing. For example, the Article 31bis mechanism allows states to issue compulsory licenses permitting manufacturers in third states to produce the essential medicines on their behalf.[229] Similarly, TRIPS could be adapted to support more equitable governance of AI as an essential global public good.
- **Model laws and treaty provisions:** International organisations could develop model legal frameworks on AI technology transfer, creating soft law standards that states can adapt domestically. Voluntary initiatives like the UN Digital Compact, which recognises open AI models as digital public goods, could gradually pave the way for binding treaty provisions.[230]
- **Differentiated responsibilities:** Recognising asymmetries, obligations could be tiered: frontier states and firms bear greater duties to share technology, while recipient states commit to developing absorptive capacity and adhering to safeguards (e.g. security screening, ethical deployment).

5.4 Partnership and financing models

Technology transfer also depends on institutional partnerships and financing models that make transfer politically acceptable, commercially viable, and administratively feasible.[231] Well-designed partnerships not only facilitate access but also embed reciprocity, reducing the perception that transfer is one-sided.

---

[228] See Japan-Indonesia Economic Partnership Agreement, 2007 https://www.mofa.go.jp/region/asia-paci/indonesia/epa0708/ or EU-CARIFORUM Economic Partnership Agreement, 2008 https://trade.ec.europa.eu/access-to-markets/en/content/eu-cariforum-economic-partnership-agreement

[229] Article 31*bis, TRIPS Agreement* (1 January 1995).

[230] United Nations, *Global Digital Compact*, Annex I to the *Pact for the Future*, adopted at the Summit of the Future, New York, 2024 https://www.un.org/global-digital-compact/sites/default/files/2024-09/Global%20Digital%20Compact%20-%20English_0.pdf.

[231] E. Harris & M. Tanner, "Health Technology Transfer", 321(7264) *BMJ*, 2000 https://pmc.ncbi.nlm.nih.gov/articles/PMC1118623/.

I. Public–private partnerships (PPPs)

- **Joint ventures with benefit-sharing mandates:** Governments can negotiate joint ventures with frontier AI firms that require local capacity-building as a condition of market access. For example, licensing agreements could include commitments to establish local labs, train domestic engineers, or share infrastructure.
- **Technology access funds:** PPPs could establish dedicated funds financed through a mix of corporate contributions and public revenues to subsidise compute credits, research collaborations, and training programmes for Global Majority institutions.

II. South–South cooperation

- **Regional centres of excellence:** Emerging AI leaders in the Global South (e.g., India, Brazil, South Africa) can serve as regional anchors, hosting training hubs and compute facilities that benefit neighbouring countries. Efforts such as the Working Group on the Ethics of Artificial Intelligence in Latin America and the Caribbean[232] illustrate how such initiatives can foster regional cooperation and capacity-building. This reduces reliance on North–South transfer and strengthens collective bargaining power.[233]
- **Cross-regional knowledge networks:** Platforms for South–South collaboration can accelerate diffusion of best practices, model governance approaches, and technical expertise, much like Global North countries collaborate on AI safety e.g. through the OECD-GPAI merger.[234] These networks can mirror global health collaborations where knowledge sharing across developing states has proved decisive (e.g., antiretroviral rollouts in Africa and Latin America).[235]

III. International financing mechanisms

- **Sovereign technology funds:** States can capitalise sovereign funds using revenues from AI taxation or data dividends, earmarking them specifically for technology acquisition and capacity-building. These funds reduce dependency on ad hoc donor contributions and embed technology transfer into long-term fiscal planning.
- **Global AI capacity facility:** Multilateral development banks or UN agencies could establish a dedicated facility to finance AI infrastructure and skills development.

---

[232] See https://forojalac.org/en/

[233] United Nations Industrial Development Organization, *UNIDO-SSTIC Side Event at the 22nd High-Level Committee on South-South Cooperation: Concept Note*, 2025, https://unsouthsouth.org/wp-content/uploads/2025/05/UNIDO-SSTIC-Side-Event-at-the-22nd-HLC-on-SSC-Concept-Note-v20250516.pdf.

[234] See https://oecd.ai/en/about/about-gpai

[235] Jessica de Mattos Costa, Thiago Silva Torres, Lara Esteves Coelho & Paula Mendes Luz, "Adherence to Antiretroviral Therapy for HIV/AIDS in Latin America and the Caribbean: Systematic Review and Meta-Analysis", 25 *Journal of International AIDS Society*, 2022, https://onlinelibrary.wiley.com/doi/full/10.1002/jia2.25066: J. B. Nachega, P. Musoke, P. H. Kilmarx et al., "Global HIV Control: Is the Glass Half Empty or Half Full?", 10(9) *Lancet HIV,* 2023, https://pmc.ncbi.nlm.nih.gov/articles/PMC10733629.

Modeled on the Global Fund for HIV/AIDS or the Green Climate Fund, such a facility would pool resources and distribute them based on need and readiness.

- **Concessional loans and blended finance:** Concessional financing mechanisms (e.g., combining grants, loans, and private investment) can reduce the upfront costs of establishing AI infrastructure in low-income countries, while ensuring repayment terms remain sustainable.

IV. Conditionality and reciprocity

To be sustainable, partnership and financing models must integrate reciprocity: recipient states commit to building safeguards (e.g., data protection, security vetting, ethical frameworks), while frontier states and firms commit to providing access. This mutual obligation reduces fears of misuse and improves political viability in frontier jurisdictions.

5.5 SEZ-based platforms for technology transfer and capacity-building

SEZs provide an institutional template through which the various mechanisms of technology transfer and capacity-building can be bundled into a coherent regime. Unlike piecemeal measures, SEZs combine fiscal incentives, regulatory adaptations, infrastructure commitments, and governance structures within a delimited jurisdiction. For AI, SEZs could serve as living laboratories for both technology transfer and redistribution, aligning global benefit-sharing with national development strategies.[236]

[236] Said Saillant, *From Tax Holidays to Tech Havens: A Playbook for AI-Ready Special Economic Zones in Latin America*, Working paper, 2025.

Mechanisms for Technology Transfer and Capacity-Building	
Infrastructure mechanisms	1. Shared compute hubs 2. Cloud credits and subsidised access 3. Regional data centres and sovereign data governance 4. Connectivity and energy investments
Knowledge transfer and human capital	1. Training programmes and fellowships 2. Joint research labs 3. Curriculum development and standards alignment
Legal instruments for technology transfer	1. Intellectual property adaptations a. Compulsory licensing for AI systems b. Adapted copyright for AI-generated content c. Open licensing and shared intellectual property pools 2. Trade and competition law a. Competition law to prevent market concentration b. Technology transfer provisions in trade agreements 3. International legal regimes a. Trade-Related Aspects of Intellectual Property Rights (TRIPS) adjustments for AI b. Model laws and treaty provisions
Partnership and financing models	1. Public–private partnerships (PPPs) a. Joint ventures with benefit-sharing mandates b. Technology access funds 2. South–South cooperation a. Regional centres of excellence b. Cross-regional knowledge networks 3. International financing mechanisms a. Sovereign technology funds b. Global AI capacity facility c. Concessional loans and blended finance
SEZ-based platforms for technology transfer and capacity-building	1. SEZs provide an institutional template through which the various mechanisms of technology transfer and capacity-building can be bundled into a coherent regime.

Table 7: Summary of the technology transfer and capacity-building AI benefit-sharing mechanisms.

Chapter 6: Non-Proliferation and Safety

If redistribution seeks to share AI's gains and technology transfer aims to expand who can participate in AI development, then non-proliferation and safety serve to mitigate the severe and potentially catastrophic risks from AI misuse.

AI is a dual-use technology. The same systems that enable medical breakthroughs or climate modelling can also facilitate cyberattacks, disinformation campaigns, or bioweapons development. This dual-use character creates a structural dilemma: the wider AI capabilities are shared, the more inclusive and developmental their impact; but the more they diffuse, the higher the risks of catastrophic misuse or strategic instability.

The governance challenge is therefore not whether to restrict access, but how to do so in ways that are targeted, legitimate, and compatible with benefit-sharing. A blanket securitisation of AI risks reproducing the exclusionary patterns of nuclear governance, locking Global Majority states out of the frontier indefinitely. Conversely, laissez-faire diffusion risks uncontrolled proliferation of dangerous capabilities.

Non-proliferation in the AI context should therefore be reconceptualised as **a graduated access regime**: one that advances safe and beneficial diffusion, while applying targeted restriction to high-risk capabilities.

6.1 Rules and standards for high-risk capabilities

The first step in operationalising non-proliferation is defining which AI capabilities should be subject to restriction, under what conditions, and by whom. Unlike nuclear material or fissile technology, AI models and datasets are intangible and widely replicable, making safety and security less about physical containment and more about rules of access, licensing, and oversight.

I. Defining high-risk capabilities

- **Frontier models with dual-use potential:** Large-scale AI systems that can enable cyberattacks, design bioweapons, or generate disinformation at scale are prime candidates for control. Establishing criteria based on model size, training data, or demonstrated capability is necessary to distinguish between benign and sensitive systems. To maintain up-to-date safety standards, it is essential to mandate robust red-teaming exercises and clearly define follow-up actions based on their findings.[237]

[237] The UK AI Security Institute's research is an example of frontier AI safety evaluations that have the direct pathway to inform policy responses https://www.aisi.gov.uk/research. Other examples include RAND Corporation's exercise focused on the potential for AI use for biological attacks https://www.rand.org/pubs/research_reports/RRA2977-1.html.

- **Training datasets for high-risk domains:** Datasets related to pathogens, weapons design, or critical infrastructure may warrant graduated access regimes, with licensing requirements for research use.
- **Critical compute resources:** High-end chips and large-scale compute clusters are already subject to export controls. Expanding such regimes to include AI-specific thresholds (e.g. FLOP limits, cluster interconnect speeds) may help prevent uncontrolled scaling of sensitive models.

II. Licensing and authorization regimes

- **Model licensing:** States could mandate licenses for training or deploying high-risk AI systems, requiring applicants to demonstrate compliance with safety, ethical, and security standards. This mirrors regimes for nuclear facilities or pharmaceuticals.
- **User vetting and credentialing:** Access to sensitive models and datasets could be restricted to accredited researchers and institutions, vetted by national or international authorities.[238] The IAEA's system of facility inspections[239] offers a precedent for graduated access to dual-use technologies.
- **Verification measures:** While non-proliferation measures are necessary to limit the spread of high-risk capabilities, their credibility depends on the presence of verifiable safeguards. States, both frontier and Global Majority, should therefore actively advocate for the inclusion of robust verification mechanisms in emerging AI governance frameworks. Such measures are essential to ensure compliance with agreed safety standards, enable cooperation under conditions of limited trust, and guarantee that commitments to both security and equitable diffusion are verifiable in practice.[240]

III. International standards and norms

- **Red-lines agreements:** States could negotiate explicit red lines around certain applications e.g., prohibiting the training or deployment of AI systems for autonomous bioweapons design. Such norms would mirror the Geneva Protocol's ban on chemical weapons, but adapted for digital capabilities.
- **Baseline safety standards:** Internationally agreed standards for risk assessments, safety benchmarks, and secure deployment practices could create common ground for distinguishing responsible from irresponsible use. International certifications grounded in these standards could serve as assurances of AI system safety, helping to build trust

[238] Seger E, Dreksler N, Moulange R, Dardaman E, Schuett J, Wei K, *et al, Open-sourcing highly capable foundation models: An evaluation of risks, benefits, and alternative methods for pursuing open-source objectives*, Centre for the Governance of AI 2023, 32-33, https://cdn.governance.ai/Open-Sourcing_Highly_Capable_Foundation_Models_2023_GovAI.pdf
[239] See https://www.iaea.org/publications/factsheets/iaea-safeguards-overview
[240] See: https://aigi.ox.ac.uk/publications/verification-for-international-ai-governance/; https://www.un.org/scientific-advisory-board/sites/default/files/2025-06/verification_of_frontier_ai.pdf

among states and companies while also offering a mechanism for enforcement and accountability in global AI governance.

- **Differentiated responsibilities:** Frontier states bear primary responsibility for restricting the diffusion of the most sensitive capabilities including through robust regulation to oversee and assess the activities of AI companies. However, restrictions must be paired with capacity-building for Global Majority states to ensure they are not locked out of AI's peaceful applications.

Rules and Standards for High-Risk Capabilities	
Defining high-risk capabilities	1. Frontier models with dual-use potential 2. Training datasets for high-risk domains 3. Critical compute resources
Licensing and authorization regimes	1. Model licensing 2. User vetting and credentialing 3. Verification measures
International standards and norms	1. Red-lines agreements 2. Baseline safety standards 3. Differentiated responsibilities

Table 8: Summary of the non-proliferation and safety components of AI benefit-sharing.

Conclusion

This paper began with a straightforward observation: despite widespread rhetorical commitment to "sharing the benefits of AI," the mechanisms required to translate principle into practice remain underdeveloped, fragmented, or absent altogether. We are witnessing a technological transformation that could either reduce global inequalities or entrench them permanently, and the window for meaningful intervention is narrowing rapidly.

The paper's central argument is that AI benefit-sharing requires integration across three distinct governance traditions: redistribution of economic gains, technology transfer and capacity-building, and non-proliferation and safety controls.

The analysis has implications for ongoing governance debates. Current approaches to AI safety and AI access are proceeding largely on separate tracks – the former concentrated in frontier states and companies, the latter addressed primarily through corporate expansion strategies or fragmented development assistance. This separation is both analytically flawed and practically dangerous. Safety and access are interdependent: exclusion from governance processes undermines legitimacy and compliance, while uncontrolled diffusion increases catastrophic risks. The framework proposed here demonstrates how these objectives can be pursued simultaneously through mechanisms that are targeted, graduated, and legitimate.

We have argued that AI benefit-sharing cannot be reduced to *ex post* redistribution. It must integrate three interdependent dimensions:

1. **Redistribution** ensures that the economic gains from AI are not captured exclusively by a narrow elite but are channeled toward broad-based social welfare through taxation, sovereign funds, procurement requirements, and fiscal governance.
2. **Technology transfer and capacity-building** enables states – particularly in the Global Majority—to develop their own AI capabilities, reducing dependency and fostering self-reliance through infrastructure investment, knowledge exchange, and legal frameworks that support absorption rather than extraction.
3. **Non-proliferation and safety** manages the dual-use risks of advanced AI systems through graduated access regimes, licensing requirements, verification mechanisms, and differentiated responsibilities that ensure diffusion occurs under conditions of safety rather than uncontrolled proliferation.

These three pillars are not alternatives but complementary requirements. Redistribution without capacity-building risks perpetuating dependency; capacity-building without safety mechanisms risks catastrophic misuse; and safety regimes that ignore equity concerns risk reproducing the exclusionary patterns of past arms control regimes. Effective benefit-sharing requires all three, implemented in ways that are mutually reinforcing rather than contradictory.

Practical Pathways Forward

The mechanisms outlined in Chapters 4–6 provide a toolkit rather than a prescription. Different contexts will require different combinations of instruments, adapted to local institutional capacity, geopolitical positioning, and development priorities. However, several cross-cutting principles emerge:

- **States remain central actors**, even in a domain dominated by private firms. Through regulation, procurement, taxation, and infrastructure control, states retain leverage to shape outcomes – if they choose to exercise it.
- **Special Economic Zones** offer a practical institutional template for bundling fiscal incentives, regulatory experimentation, and governance oversight in bounded geographies, allowing states to pilot benefit-sharing mechanisms before scaling them nationally.
- **International coordination** is essential but must be designed to preserve sovereignty rather than undermine it. Regional hubs, South-South cooperation, and multilateral financing mechanisms can support diffusion without imposing one-size-fits-all governance frameworks.
- **Verification and transparency** are critical for managing dual-use risks while maintaining legitimacy. Non-proliferation regimes that operate opaquely or exclude affected parties will lack credibility and invite resistance.

Readiness as a Precondition Toward a Readiness Checklist

Taken together, these prerequisites can be translated into a **practical readiness checklist** for policymakers:

- Does the state have fiscal systems capable of taxing AI-related value?
- Does it have distribution channels to allocate resources equitably?
- Does it possess research and infrastructural capacity to absorb technology transfers?
- Does it have institutions capable of implementing safety and non-proliferation rules?

This checklist offers a diagnostic tool. It does not prescribe identical pathways for all states, but highlights the **minimum conditions without which benefit-sharing will remain rhetorical**.

Image 4: Summary of the selected AI benefit-sharing mechanisms relevant for different stakeholders.

Centre for AI Security and Access is fiscally sponsored by Rethink Priorities.