

ARTICLE

Balancing Innovation and Equity: The Global South's Role in Shaping AI Copyright Policy

Arti Aneja,

Tashkent State University of Law (Tashkent, Uzbekistan)

<https://orcid.org/0000-0002-9771-1890>

Ankit Shrivastava,

Symbiosis Law School (Noida, India)

<https://orcid.org/0000-0002-0347-3175>

<https://doi.org/10.21684/2412-2343-2026-13-1-56-75>

Received: October 12, 2025

Reviewed: November 3, 2025

Accepted: November 26, 2025

Abstract. Artificial intelligence (AI) is now playing a major part in changing creative and knowledge economies around the world. However, policies guiding this change are primarily still led by the Global North, which may result in policies being formulated without taking into account the Global South. The article proposes an inclusive governance framework aimed at preventing digital colonization and argues that the framework must lower market barriers for Southern innovators. The article suggests different measures based on recent regional experiences. These examples include India's case against OpenAI and also the Text and Data Mining (TDM) strategies of Latin America and the governance initiatives in Africa. Key measures include making AI training data disclosures and consent, using community-based models that formally recognize collective and indigenous authorship and benefit sharing with creators in the Global South.

Keywords: artificial intelligence (AI); copyright policy; Text and Data Mining (TDM); Global South; digital colonialism.

To cite: Aneja, A., & Shrivastava, A. (2026). Balancing innovation and equity: The Global South's role in shaping AI copyright policy. *BRICS Law Journal*, 13(1), 56–75.

Table of Contents

Introduction

1. Training, Text and Data Mining, and Lawful Access:

Exceptions that build Domestic AI

2. Transparency, Datasets, and Opt-Outs: Necessary but Not Sufficient

3. Governance, Decolonial Equity, and Policy Design for the Global South

3.1. Capacity, Competition, and Localization: Building Endogenous AI Ecosystems

3.1.1. Regulatory Sandboxes/ELRs

3.1.2. Public-Private Partnerships (PPP)

3.1.3. Special Economic Zones (SEZs)

3.2. Approach by Different Jurisdictions

3.2.1. India

3.2.2. Japan and Singapore

3.2.3. European Union

3.2.4. Community-Centric and Cultural Models

3.3. Policy Design Implications

Conclusion

Introduction

Recent developments in AI technology are challenging the current copyright regime worldwide. Giving legal access to training data to train AI capabilities is difficult if you want to protect the individual creators and indigenous communities from misappropriation of any kind. This is a big challenge for the Global South.¹ International touchstones, whether it is UNESCO's Recommendation on the Ethics of AI or ASEAN's Guide on AI Governance and Ethics, offer common principles including human rights, equity, transparency, accountability, etc. However, they fail to address more important: fundamental issues, such as authorship and originality thresholds for AI-assisted outputs.²

¹ Fisher, A., & Nathani, Z. (2024, February 19). *Association of Southeast Asian Nations provides guidance on AI governance and ethics*. Linklaters. <https://techinsights.linklaters.com/post/102j06t/association-of-southeast-asian-nations-provides-guidance-on-ai-governance-and-eth>

² United Nations Educational, Scientific and Cultural Organization (UNESCO). (2021). *Recommendation on the Ethics of Artificial Intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

The Global South increasingly relies on alliances like BRICS+ to articulate governance policies. The expanded BRICS group is a key example. This group, which now includes Iran, Egypt, Ethiopia, and the UAE, in addition to the original members: Brazil, Russia, India, China, and South Africa, has evolved into a strong force.³ This expansion represents a major progress for nations seeking to strengthen the Global South's influence. One of the main objectives for this bloc in the near future is turning into digital sovereignty with countries like Russia and China as frontrunners in creating rules for this⁴. While Russia implemented this objective using its "Sovereign Internet Law," China employed the "Great Firewall." India is also advancing its digital sovereignty through laws such as the Information Technology Act 2000 and the proposed Digital India Act. This focus confirms that AI policy must centre on national control of data and infrastructure. This step is directly linked to preventing digital colonialism along with sustainable development.

The Global South's approach to AI governance differs significantly from that of the Global North, which tends to adopt a horizontal regulatory model. This model does not differentiate between various AI systems. It tries to impose a full set of strict rules on all AI uses. These strict rules cover areas such as fairness, accountability, and security. The European Union's AI Act is also an example of the one-size-fits-all approach. Developing economies from the Global South may find this framework too rigid to implement due to their unique economic and social circumstances. Instead, the Global South favours a context-specific approach. This strategy tailors regulations to fit the exact use case of the AI system. For example, different rules can be framed for different AI systems, such as autonomous, allocative, and generative AI. It may give better results for jurisdictions trying to encourage innovations. This modular method helps these countries avoid unnecessarily high compliance costs and rigid rules.

Comparative practice in the Asia Pacific illustrates an interesting approach. Singapore's computational data analysis exception conditions TDM, including commercial uses, on lawful access. Japan's Article 30-4 adopts a "non-enjoyment" model which generally permits training absent unreasonable prejudice and recognises limits for overlearning.⁵ These innovation-forward models contrast with the EU's TDM regime, which includes rightsholder opt-outs, and provides alternative frameworks for the Global South.⁶ India encapsulates the policy calculus: Section 52's fair dealing plausibly covers non-expressive TDM, yet statute and jurisprudence are silent on AI

³ Upadhyay, N. K. (2024). BRICS expansion: A comprehensive analysis of global power shifts and economic integration. *BRICS Law Journal*, 11(3), 5–8.

⁴ Gromova, E., & Ferreira, D. B. (2024). On the way to BRICS+ digital sovereignty: Opportunities and challenges of a new Era. *BRICS Law Journal*, 11(3), 54–69.

⁵ Hays, S. (October 30, 2024). *AI training and copyright infringement: Solutions from Asia*. Tech Policy Press. <https://www.techpolicy.press/ai-training-and-copyright-infringement-solutions-from-asia/>

⁶ Ueno, T. (2021). The flexible copyright exception for "non-enjoyment" purposes—recent amendment in Japan and its implication. *GRUR International*, 70(2), 145–152.

training. The *ANI v. OpenAI* lawsuit surfaces boundary issues—training versus output use, evidence of market harm, and remedial deletion,⁷ that will shape doctrine in a jurisdiction prioritising rapid AI diffusion without a comprehensive AI statute.⁸

A decolonial lens clarifies the governance stakes. Scholarship documents how infrastructural and regulatory monopolies,⁹ data and labour extractivism, and epistemic dominance can reproduce colonial hierarchies if left to market logics.¹⁰ Sub-Saharan assessments find decolonising AI governance remains incipient; most states are “decolonisation-aware” but lack institutions that prioritise sovereignty, local data use, and equitable benefit-sharing.¹¹ Capacity-building perspectives reframe the task for non-core AI countries as long-term: build national capability, insist on algorithmic transparency tied to outcomes, and localise models through transfer learning and sectoral priorities, rather than importing turnkey solutions.¹² These insights imply that training-data transparency, usable opt-outs, and fair remuneration are not merely copyright tools but legitimacy conditions.¹³

This paper advances three claims. First, the Global South jurisdictions can reconcile innovation and equity by coupling liberal, lawful-access TDM (or functionally equivalent fair dealing interpretations) with enforceable transparency, practicable opt-outs, and targeted remuneration, while preserving research carve-outs and SME access.¹⁴ Second, decolonial governance requires community-centred data models, cultural and linguistic plurality in evaluation, and regional interoperability, rather than transplantation of the Global North templates.¹⁵ Third, institutional capacity—not statutory text alone—determines outcomes: copyright offices, courts, and

⁷ Pandya, D. (2025, March 7). *The Global South AI copyright's test case: India*. Center for European Policy Analysis (CEPA). <https://cepa.org/article/the-global-south-ai-copyrights-test-case-india/>

⁸ Chauhan, K. (2025). Generative AI, text & data mining and the fair dealing doctrine: Examining the new problem with the old regime. *Journal of Intellectual Property Rights*, 30, 77–85.

⁹ Muldoon, J., & Wu, B. A. (2023). Artificial intelligence in the colonial matrix of power. *Philosophy & Technology*, 36, Article 80.

¹⁰ Png, M.-T. (2022). At the tensions of South and North: Critical roles of Global South stakeholders in AI governance. *FACCT Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 22, 1434–1445.

¹¹ Ayana, G., et al. (2024). Decolonizing global AI governance: Assessment of the state of decolonized AI governance in Sub-Saharan Africa. *Royal Society Open Science*, 11(8), 231994.

¹² Farhad, Sh. (2025). Passengers in flight: AI governance capacity in the Global South. *Digital Society*, 4, article number 39.

¹³ Buick, A. (2025). Copyright and AI training data—transparency to the rescue? *Journal of Intellectual Property Law & Practice*, 20(3), 182–192.

¹⁴ Peter, K. Yu. (2024). *The future path of artificial intelligence and copyright law in the Asian Pacific* (Texas A&M University School of Law Research Paper No. 24-18).

¹⁵ Vijayakumar, A. (2024). *AI ethics for the global south: Perspectives, practicalities, and India's role* (RIS Discussion Paper No. 296, Research and Information System for Developing Countries).

competition and data protection regulators need coordinated mandates, iterative guidance, and metrics tracking both innovation and distributional equity.¹⁶

1. Training, Text and Data Mining, and Lawful Access: Exceptions that build Domestic AI

Generative AI's reliance on large-scale text and data mining (TDM) raises two core design choices for copyright policy in the Global South: (i) whether and how to guarantee "lawful access" to training corpora, and (ii) how limitations and exceptions allocate risk between rightsholders and developers. Transparency is increasingly deployed to police training uses, but its effect is contingent on the underlying copyright baseline.¹⁷ In the EU, the AI Act's training-data transparency was crafted to facilitate enforcement of the CDSM Directive's Article 4 opt-out; as Buick argues, transparency "alone" will not fix structural flaws in the opt-out and is "unlikely" to materially improve individual creators' position.¹⁸

Comparative experience in Asia-Pacific illustrates two viable, innovation-forward architectures. Japan's 2019 amendment (Copyright Act Art. 30-4) authorizes use "by any means" to the extent necessary where the purpose is "non-enjoyment," subject to a backstop against "unreasonable prejudice" to rightsholders; administrative guidance confirms general permissibility for training while warning of "overlearning."¹⁹ This exception applies to commercial and non-commercial uses, has no opt-out and no lawful-access condition, placing calibration on the prejudice test and output controls. Singapore's 2021 Copyright Act created a computational data analysis (CDA) exception conditioned on "lawful access,"²⁰ available for commercial TDM, with explicit policy intent to support national AI strategy and guidance emphasizing 'balance' between data access and copyright interests. These two models lower entry barriers while embedding guardrails either ex ante (lawful access) or ex post (unreasonable prejudice/overlearning).

By contrast, the EU's CDSM framework couples a mandatory research TDM exception (Article 3) with a broader Article 4 exception subject to rightsholder opt-out; the EU AI Act transparency provisions operationalize that opt-out but—standing alone—do not remedy its structural weaknesses. The UK confines TDM to non-commercial purposes. Some jurisdictions have clarified coverage under

¹⁶ Farhad, 2025.

¹⁷ Kamaruddin, S., et al. (2025). The legal paradigm of generative AI in Malaysia and India: Problems and prospects. In *2024 International conference on artificial intelligence and emerging technology (global AI summit)*, Greater Noida, India, April, IEEE (pp. 413–417).

¹⁸ Buick, 2025.

¹⁹ Hays, 2024.

²⁰ Tan, D., & Lee, T. (2021). Copying right in copyright law: Fair use, computational data analysis and the personal data protection act. *Singapore Academy of Law Journal*, 33, 1032.

existing fair use (e.g., Israel), and in the United States developers invoke fair use rooted in non-expressive, transformative uses recognized in mass digitization cases, although courts have not yet fully tested those arguments against memorization and substitution risks; adverse rulings could yield existential damages.²¹

The imperative for clear training exceptions is heightened by the dual nature of intellectual property rights in emerging economies. IPR frameworks generally seek to balance the rights of authors and the general public interest. They drive growth by incentivizing R&D and attracting investment as there will be lack of innovation in absence of such incentives.²² However, they also risk creating monopolies and restricting access to essential goods and technologies by granting exclusive rights. This compels BRICS nations and the broader Global South to seek a moderate stance towards innovation and equity.²³ This policy aims to balance innovation with equity. To fulfill this objective, new laws are necessary. These laws aim to lower costs for domestic innovators. They address ongoing problems in countries like Brazil, Russia, and South Africa including chronic issues with institutional capacity and judicial delays.

India embodies the core tension—what may be termed the “policy calculus”—between accelerating domestic technological innovation and safeguarding socio-economic equity. As per Section 52 of the Copyright Act, 1957, fair dealing framework plausibly accommodates non-expressive TDM for research and certain private uses, but the statute is silent on AI training specifically; jurisprudence treats fair dealing as a case-by-case balance between authorial interests and societal goals. Scholars argue the existing regime is “insufficient” for TDM and that, given India’s economic ambitions and pronounced linguistic diversity (121 languages; colloquial speech affecting model performance), India should adopt a liberal TDM exception with a lawful-access condition, narrow and workable opt-outs, and output safeguards, rather than transplanting the EU opt-out.²⁴ They contrast EU Articles 3–4, which are criticized for legal uncertainty and potential drag on AI R&D with Japan’s broad “non-enjoyment” model that aims to ensure copyright “should not be an obstacle” to AI development.²⁵ The pending *ANI v. OpenAI* suit underscores boundary issues—training versus output use, evidence of market harm, remedial deletion, jurisdiction. The outcome of the case is likely to raise new challenges: a defence win could shield training on public data as fair use, whereas an adverse ruling could force costly licensing.²⁶

²¹ Quang, J. (2021). Does training AI violate copyright law? *Berkeley Technology Law Journal*, 36, 1407–1436.

²² Haryono, H., et al. (2024). Intellectual property rights in BRICS countries: Legal frameworks, enforcement challenges, and economic implications. *West Science Interdisciplinary Studies*, 2(12), 2494–2501.

²³ Buick, 2025.

²⁴ Chauhan, 2025.

²⁵ Quintais, J. P. (2020). The new copyright in the digital single market directive: A critical look. *European Intellectual Property Review*, 42, 28–41.

²⁶ Pandya, 2025.

Regionally, soft-law moves are converging toward interoperable guidance while statutes catch up. Korea's 2023 guidance encourages licensing, fair compensation, and opt-out avenues while recognizing uncertainty under existing fair use; Hong Kong is consulting on a commercial TDM exception modelled on Japan/Singapore; China's measures instruct providers not to infringe IP, and courts have quickly policed related rights (e.g., voice).²⁷ ASEAN's Guide advances seven governance principles: transparency/explainability, fairness/equity, security/safety, robustness/reliability, human-centricity, privacy/data governance, accountability/integrity and recommends a regional working group to address liability, IP, and data issues, providing scaffolding for Member State TDM reforms.²⁸

Two design lessons emerge for the Global South. First, transparency is an adjunct, not a substitute. Dataset and provenance disclosure have greatest value when paired with clear, administrable exceptions (lawful-access CDA or non-enjoyment with prejudice backstop), practical opt-outs (collective or protocol-based rather than atomized web notices), and output-correction duties for overlearning; absent this, transparency risks compliance theatre and forum shopping. Second, lawful access is a policy lever, not an absolute. Conditioning exceptions on lawful access (Singapore) can deter scraping of paywalled content and incentivize licensing without foreclosing training on open/public data; non-enjoyment tests (Japan) can accommodate broader inputs if outputs are controlled and rightsholder prejudice is credibly policed.²⁹

For India and similarly situated jurisdictions, an administrable middle path is available: legislate a TDM exception that (a) covers commercial and non-commercial TDM; (b) requires lawful access (open/public availability, licensed access, or equivalent legal bases); (c) excludes processing of personal data beyond applicable data-protection bases; (d) replaces fragmented per-work opt-outs with limited, standardized sectoral or collective opt-outs where justified; (e) mandates reasonable, model-family-level transparency of training sources; and (f) imposes output safeguards (memorization testing, complaint-and-correction for near-verbatim outputs, related-rights protection for voice/likeness). This approach advances domestic AI while tempering extractive dynamics, aligns with UNESCO's calls for open and trustworthy datasets and data-sharing mechanisms, and is congruent with ASEAN's risk-proportionate governance. It also fits capacity realities: courts remain available to police "unreasonable prejudice" and market substitution in edge cases, but the default pathway is legislative, minimizing transaction costs and uncertainty

²⁷ Kharitonova, Yu., Malik, N. S., & Yang, T. (2023). The legal issue of deterrence of algorithmic control of digital platforms: The experience of China, the European Union, Russia and India. *BRICS Law Journal*, 10(1), 147–170.

²⁸ Fisher & Nathani, 2024.

²⁹ Hays, 2024.

for startups, SMEs, and public-interest research. Furthermore, the policy must address outputs generated by Large Language Models (LLMs). There are ethical and legal risks involved which demand robust safeguards. These LLMs can exhibit profound algorithmic bias, inheriting and amplifying biases from their training data, including religious, racial, and political prejudices. They have been shown to generate offensive or defamatory content. Legally, there are immediate challenges regarding the ownership of content created by intelligent chatbots, contrasting, for example, the UK's computer-generated works law with the Australian requirement for a "human author" applying some "creative spark"³⁰ Moreover, the lack of transparency regarding specific training sources means LLMs may generate responses copied from academic or copyrighted material, violating IP rights. Therefore, policies must mandate output controls and accountability mechanisms (like memorization testing and complaint-and-correction regimes) that address these ethical risks and IP infringements.

2. Transparency, Datasets, and Opt-Outs: Necessary but Not Sufficient

Transparency about training datasets improves accountability and trust, but it cannot by itself resolve the allocation problems at the heart of AI—copyright conflicts. As Buick argues, the impact of transparency is contingent on the surrounding copyright baseline; without workable limitations and exceptions (L&Es) and administrable rights, transparency "alone" is unlikely to materially improve individual rightsholders' position and may instead increase compliance costs without addressing core incentives.³¹ The EU's AI Act illustrates this conditionality: its training-data transparency is keyed to enforcing the opt-out in Article 4 of the CDSMD Directive, but the structural flaws of that opt-out; fragmentation, high transaction costs, and weak remedies remain uncorrected by disclosure duties.

The EU model has two moving parts: (1) an opt-out from Article 4 TDM, and (2) an AI-Act requirement that general-purpose model providers publish a "sufficiently detailed" summary of training sources and adopt a copyright-compliance policy that honours opt-outs (AI Act art. 53(1)(c)–(d)). In principle, this linkage operationalizes rightsholders' choices; in practice, summary-level disclosures may neither be granular enough for detection and enforcement nor feasible at the per-work level given web-scale datasets.³² As a result, there is a challenge at enforcement level. Even if large language model providers disclose high-level sources, individual creators will

³⁰ Gromova, E. A., Ferreira, D. B., & Begishev, I. R. (2023). ChatGPT and other intelligent chatbots: Legal, ethical and dispute resolution concerns. *Revista Brasileira de Alternative Dispute Resolution*, 5(10), 153–175.

³¹ Buick, 2025.

³² Ziaja, G. M. (2024). The text and data mining opt-out in article 4(3) CDSMD: Adequate veto right for rightsholders or a suffocating blanket for European artificial intelligence innovations? *Journal of Intellectual Property Rights*, 19(5), 453–459.

face information asymmetries, evidence burdens, and legal uncertainty, especially in the jurisdictions where national fair use/fair dealing rationales for non-expressive TDM remain open.

Due to ethical concerns surrounding training data, reliance on transparency and generalized opt-outs seems largely inadequate. In order to prevent bias and discrimination, LLM developers should apply international standards, such as the provisions of ISO/IEC TR 24027:2021.³³ Transparency must be linked to adherence to these standards and demonstrable efforts to ensure data quality. Since LLMs can generate incorrect or biased information due to training on limited or unfiltered datasets, these standards become very important. The result of an unfiltered output may be extremely dangerous and can even lead to physical or psychological harm, such as a large language model urging a user to commit suicide or affirming genocide.³⁴ These acute risks mandate that transparency cannot be merely “performative compliance”; it must directly support output safeguards and ethical governance mechanisms designed for accountability.

Comparative study of some Asian Jurisdictions show why transparency must be embedded in a coherent liability and access framework. Japan’s Article 30-4 “non-enjoyment” exception generally permits training (commercial and non-commercial), without a lawful-access requirement or rightsholder opt-out, but cabins abuse through the ‘unreasonable prejudice’ backstop and administrative guidance that flags “overlearning” (near-verbatim output) as out of bounds. Singapore’s Computational Data Analysis (CDA) exception permits commercial TDM conditioned on lawful access, aligning disclosure and access rules with national innovation strategy; guidance frames “balance” in terms of trusted-ecosystem goals, not per-work vetoes.³⁵ In both systems, transparency can support auditability, research, and trust, but doctrinal clarity, not opt-out atomization, does the primary work of risk allocation.

For Global South jurisdictions, three design pitfalls follow. First, transparency without clear L&Es risks performative compliance: model cards and source summaries proliferate, but creators obtain little leverage and startups face higher fixed costs. Second, unbounded opt-outs can entrench incumbents with proprietary corpora, raising barriers for domestic researchers and SMEs that lack licensing power, especially in low-resource languages. Third, per-site robots.txt-style reservations are ill-suited to cultural heritage institutions and community datasets, where collective governance and contextual consent matter more than web-scale crawling toggles.³⁶

³³ Kahana, E. (2023, March 17). *AI life cycle core principles*. Stanford Law School – CodeX. <https://law.stanford.edu/2023/03/17/ai-life-cycle-core-principles/>

³⁴ Euronews. (2023, March 31). *Man ends his life after an AI chatbot ‘encouraged’ him to sacrifice himself to stop climate change*. <https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-an-ai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate-change>

³⁵ Hays, 2024.

³⁶ Park, S. (2024). Bridging the global divide in AI regulation: A proposal for a contextual, coherent, and commensurable framework. *Washington International Law Journal*, 33(2), 216–269.

A more effective package ties transparency to lawful access, standardized reservations, and output-side safeguards:

i. Purpose-built TDM exceptions: Adopt or interpret L&Es to cover non-expressive TDM for research and (where appropriate) commercial uses, conditioned on lawful access or “non-enjoyment” limits, with an “unreasonable prejudice” backstop to target harmful cases.³⁷

ii. Model-family provenance: Require model-family-level disclosures of major source classes (e.g., Common Crawl subsets, licensed archives, open repositories), data-collection methods, and content-filtering strategies, consistent with trade-secret protection but sufficiently detailed to enable oversight bodies and collective management organisations (CMOs) to assess risk.³⁸

iii. Standardized, collective opt-outs: Replace atomized, per-work reservations with narrowly scoped, interoperable mechanisms (e.g., sectoral or repertoire-level opt-outs managed by CMOs or registries), minimizing false positives/negatives and enabling enforceability at scale.³⁹

iv. Output safeguards and complaint-and-correction: Pair training transparency with ex post memorization testing, responsive takedown/attenuation workflows for near-verbatim outputs, and related-rights protection for voice/likeness—approaches already surfaced in Asian policy debates.⁴⁰

v. Public-interest and capacity lenses: Align disclosure and opt-out design with UNESCO’s calls for open, trustworthy datasets and data trusts, ensuring that community and indigenous⁴¹ data governance are respected and that capacity constraints in the Global South are not exacerbated.⁴²

vi. Finally, transparency should be evaluated on outcomes, not formalities: Does it improve auditability, facilitate collective solutions, and reduce substitution harm without chilling lawful access for research and localized innovation? Evidence from the EU suggests that, absent reform of the opt-out baseline, transparency will under-deliver for individual creators and over-burden new entrants.⁴³ By contrast, jurisdictions that clarify lawful access, define training-use boundaries, and couple

³⁷ Xalabarder, R. (2024). *Scoping study on the practices and challenges of research institutions and research purposes in relation to copyright* (WIPO Standing Committee on Copyright and Related Rights, Document SCCR/44/4 REV).

³⁸ Buick, 2025.

³⁹ Peukert, Ch. (2025, December). *The economics of copyright and AI: Empirical evidence and optimal policy*. European Parliament. [https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778859/IUST_STU\(2025\)778859_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778859/IUST_STU(2025)778859_EN.pdf)

⁴⁰ Hays, 2024.

⁴¹ Gladun, E., & Zadorin, M. (2023). The system of indigenous peoples’ protection in BRICS states: An overview of legal and litigation support. *BRICS Law Journal*, 10(4), 121–141.

⁴² Ayana et al., 2024.

⁴³ Buick, 2025.

high-level provenance with practical redress mechanisms are better positioned to reconcile innovation with equity.

3. Governance, Decolonial Equity, and Policy Design for the Global South

Artificial intelligence (AI) policy development in the Global South has to be seen within a broader historical and structural context. This context is marked by age-old dissimilarities in infrastructure, capital, and global agenda-setting. Mere formal inclusion in governance processes does not resolve what the “paradox of participation,”⁴⁴ wherein procedural inclusion can exist with persistent structural harms. AI production remains situated within the “colonial matrix of power”⁴⁵ reinforcing legacies of historical colonialism. This results in extractive supply chains and environmental burdens, which are disproportionately borne by the majority of the world. These structural inequities extend beyond supply chains to encompass the ownership and control of technological infrastructure, digital sovereignty, and access to the economic value derived from AI research and development.

UNESCO’s Recommendation on the Ethics of Artificial Intelligence (2021) proposes specific countermeasures, including open data initiatives, public repositories, data trusts, robust datasets, and regulatory frameworks that maintain accountability with identifiable natural or legal persons (paras. 68, 75–76). Yet, empirical assessments underscore the slow pace of reform. In Sub-Saharan Africa, only Rwanda demonstrates significant responsiveness to decolonial AI governance imperatives, despite 80% of stakeholders being “decolonization-aware.”⁴⁶ Global rules for Artificial Intelligence (AI) must prioritize equity. They need to adopt a decolonial mindset by empowering Global South actors to design solutions and also strategize its efficient implementation.

Decolonial approaches thus require centering sovereignty, local data use, and equitable benefit-sharing in AI governance. In this context, UNESCO (2021, para. 77) further urges Member States to adopt a “digital commons” approach to data, encouraging private sector actors to share collected data for research, innovation, and public benefit.⁴⁷

⁴⁴ Png, 2022.

⁴⁵ Muldoon & Wu, 2023.

⁴⁶ Ayana et al., 2024.

⁴⁷ Fan, M. D. (2021). The right to benefit from big data as a public resource. *New York University Law Review*, 96, 1438–1492.

3.1. Capacity, Competition, and Localization: Building Endogenous AI Ecosystems

Governance for the Global South must be seen as a project that builds long-term ability (capacity). It should not focus on quickly finding immediate, comprehensive regulation. The primary task is to establish recurring cycles for assessment, action, and evaluation. This approach works to expand domestic capacity. States can then progressively gain greater control over their path. This gives them control over their journey in their own countries.⁴⁸ This incrementalist model aligns with UNESCO's operational priorities, including the creation of 'gold standard' datasets, the development of annotation standards, the promotion of transparency and explainability, and the strengthening of privacy and data governance frameworks.

Institutional capacity challenges, particularly the need to attract investment and modernize regulation, can be overcome using structured legal instruments. These barriers include: the complexity of determining the subject of regulation (difficulty predicting where breakthrough technology will emerge, leading to situational "patching" rather than strategy); the lack of developed legal categories and terms that adequately reflect the essence of new digital objects; the fundamental problem of inertia of legal regulation against the rapid development of technologies; and the need to create a favorable national jurisdiction to preserve digital sovereignty. Overcoming these institutional and legal gaps requires structured mechanisms. The Triad Model proposes combining three instruments to overcome legal barriers to digital technology implementation (Industry 4.0): Public-Private Partnerships (PPP), Special Economic Zones (SEZs), and Experimental Legal Regimes (Regulatory Sandboxes).⁴⁹

3.1.1. Regulatory Sandboxes/ELRs

These allow regulators to safely 'test' new legal norms and digital innovations in a controlled environment, addressing the inertia of legal regulation and the rapid pace of technology.⁵⁰ For the Global South, sandboxes could be vital for determining optimal copyright interpretations for AI or setting governance rules in sectors like industrial production or Fintech.

3.1.2. Public-Private Partnerships (PPP)

PPPs are effective tools for modernizing the national economy and attracting significant investment. For digital infrastructure and innovation, such investments are very important, especially when budget deficits exist.⁵¹ PPPs can be used to

⁴⁸ Farhad, 2025.

⁴⁹ Gromova, E., Koneva, N. S., & Titova, E. V. (2022). Legal barriers to the implementation of digital industry (Industry 4.0) components and ways to overcome them. *Journal of World Intellectual Property*, 25, 186–205.

⁵⁰ Gromova, Koneva & Titova, 2022.

⁵¹ Cyman, D., Gromova E., & Juchnevicius, E. (2021). Regulation of artificial intelligence in BRICS and the European Union. *BRICS Law Journal*, 8(1), 86–115.

develop “breakthrough” technologies like AI. Expanding the laws on Public-Private Partnerships (PPP) is necessary to cover all intellectual and innovative activities and not just information technology, as is the case presently. Expanding the scope would increase the potential of PPPs for innovation and the digital economy. This in turn would enhance competitiveness in the digital technology market.

3.1.3. Special Economic Zones (SEZs)

SEZs attract foreign and domestic investment through special regimes. Modernizing these zones to create virtual SEZs can accelerate the deployment of digital technologies by offering streamlined procedures and tax incentives for digital innovators. Without a need for physical territory, these virtual SEZ may prove to be game changing.

This tripartite approach addresses the multiple legal barriers identified, such as the complexity of determining the subject of regulation, the lack of adequate legal categories, and the need to create a favorable national jurisdiction to preserve digital sovereignty.

At the regional level, ASEAN’s governance framework articulates seven interoperable principles, transparency and explainability, fairness and equity, security and safety, robustness and reliability, human-centricity, privacy and data governance, and accountability and integrity, and establishes a dedicated Working Group on AI Governance to address liability, intellectual property, and data governance. Such models offer scaffolding for local adaptation while preserving cross-border compatibility.⁵²

3.2. Approach by Different Jurisdictions

3.2.1. India

In the Indian context, Section 52 of the Copyright Act may partially address the copyright dimensions of text and data mining (TDM) activities; however, legislative intervention remains necessary to ensure fair governance and address emerging challenges.⁵³ While the fair dealing doctrine in India aims to balance authorial interests with societal creative objectives, the existing regime is insufficient for large-scale TDM. The *ANI v. OpenAI* dispute exemplifies the legal uncertainties surrounding AI training, output generation, jurisdictional reach, and remedial measures, with potential systemic implications for domestic innovation and licensing practices.⁵⁴ Given India’s linguistic diversity and development imperatives, a liberal TDM exception with a lawful-access requirement and output safeguards may be suitable, drawing upon the legislative experiences of Japan and Singapore.

⁵² Fisher & Nathani, 2024.

⁵³ Chauhan, 2025.

⁵⁴ Pandya, 2025.

3.2.2. *Japan and Singapore*

Japan's Copyright Act (Art. 30-4) permits TDM use "by any means ... in which it is not a person's purpose to personally enjoy" a work, subject to limits against unreasonable prejudice and overlearning.⁵⁵ Singapore's 2021 Copyright (Text and Data Mining) Amendment permits commercial TDM where there is lawful access. It explicitly links the exception to the country's national AI strategy and gives a balanced governance approach. These frameworks combine clear statutory regulation with reduced barriers to innovation, providing adaptable templates for other Global South jurisdictions.

3.2.3. *European Union*

The EU AI Act's transparency obligations regarding training data illustrate that disclosure mechanisms remain dependent on the adequacy of underlying copyright frameworks. Without substantive reform to address the copyright challenges posed by generative AI, transparency requirements alone are unlikely to significantly improve the position of individual rightsholders.

3.2.4. *Community-Centric and Cultural Models*

Special attention is required to be given to the protection of rights of Indigenous people.⁵⁶ Indigenous and community-based data governance is a critical dimension of equitable AI development. Indigenous Data Sovereignty is defined as the right of Indigenous peoples to control data originating from and concerning their communities and lands, encompassing both individual and collective rights to access and privacy.⁵⁷ UNESCO (2021, paras. 66, 77) emphasizes that training datasets should not perpetuate cultural, economic, or social inequalities, particularly in data-scarce regions, and supports the creation of trusted data spaces.

Ethical pluralism further enhances governance legitimacy. There is also potential for non-Western ethical frameworks including Ubuntu, Buddhist, and Islamic ethics, to serve as alternative normative bases for AI regulation.⁵⁸ Yet, cautions are necessary as prevailing datasets and evaluation practices often privilege hegemonic cultural defaults, thereby marginalizing Global South contexts.

⁵⁵ Hays, 2024.

⁵⁶ Gladun & Zadorin, 2023.

⁵⁷ Png, 2022.

⁵⁸ Vijayakumar, A. (2024). *AI Ethics for the Global South: Perspectives, Practicalities, and India's Role*. RIS Discussion Paper No. 296, Research and Information System for Developing Countries.

Table 1: **Comparative Table: Approaches to AI Regulation**

Feature	Horizontal (Omnibus) Approach (EU Model)	Contextual (Modular) Approach (Global South/3C Model)
Underlying Philosophy	Homogeneity of Risk: Assumes all “High-Risk” AI systems share fundamental “problem drivers” (e.g., opacity, complexity)	Context-Specific Harms: Regulates systems based on function (e.g., Generative, Allocative, Autonomous) and societal effect (Contextuality)
Regulatory Mode	Bundled Requirements: Imposes a uniform package of rules (fairness, accuracy, robustness, transparency, accountability) across disparate domains	Tailored Objectives: Assigns specific objectives to match the risk (e.g., Safety for Autonomous AI; Infringement Mitigation for Generative AI; Fairness/Explainability for Allocative AI)
TDM/ Copyright Framework	Opt-Out/Rightsholder Veto: Uses a mandatory exception for research (Art. 3) and an opt-out for commercial TDM (Art. 4), operationalized by broad transparency	Access/Purpose-Based: Favors clear statutory exceptions conditioned on <i>lawful access</i> (Singapore) or <i>non-enjoyment/prejudice</i> limits (Japan) to lower entry barriers
Innovation & Compliance	High risk of insufficient proportionality and high compliance costs, potentially stifling innovation (e.g., burdensome documentation, rigid adherence requirements)	Promotes proportionality and flexibility; favors complementary soft law (industry standards, ELRs/ Sandboxes) to achieve regulatory goals without rigid statutory language
Equity Mechanism Precedent	Primarily relies on general fundamental rights principles and transparency	Supports legally mandated, outcome-based mechanisms (e.g., mandatory CSR contributions linked to national welfare goals)

3.3. Policy Design Implications

Three interlinked policy imperatives emerge from the foregoing analysis. First, transparency must be coupled with substantive measures,⁵⁹ such as lawful-access-

⁵⁹ Kharitonova, Yu. S. (2023). Legal means of providing the principle of transparency of the artificial intelligence. *Journal of Digital Technologies and Law*, 1(2), 337–358.

conditioned exceptions, unreasonable-prejudice safeguards, standardized opt-out mechanisms, and output controls, to ensure enforceability and equitable outcomes.⁶⁰ Second, lawful-access ecosystems should be developed in tandem with investments in open, trustworthy, and annotated datasets, consistent with UNESCO's capacity-building agenda. Third, governance must be explicitly equity-oriented, addressing infrastructural monopolies, extractive supply chains, and environmental externalities, while embedding community-driven data governance and culturally grounded evaluation metrics.^{61,62}

Conclusion

Global South jurisdictions can architect AI–copyright policy that catalyses domestic innovation while embedding equity by design. The key is to shift from a reactive import of the Global North templates to an endogenous framework that aligns lawful access for training with decolonial governance commitments and institutional capacity realities as can be observed in different jurisdictions in the Global South.

There are three takeaways. First, enabling innovation requires a clear, administrable pathway for text and data mining. Legislated or interpretive routes—through lawful-access TDM exceptions or functionally equivalent fair dealing—reduce entry barriers for research, SMEs, and public interest applications while preserving room to protect legitimate creator interests. Comparative lessons from different jurisdictions demonstrate that carefully defined exceptions can lead to local AI ecosystems without excessive barriers. Second, equity cannot be an afterthought and must be ensured from the outset. Transparency about training sources, usable and standardized reservation mechanisms is instrumental. Crucially, transparency alone is not a solution: without coherent limitations and exceptions and practicable enforcement channels that address algorithmic bias and output-side IP risks, disclosure risks performative compliance and heightened costs for new entrants without materially improving the position of individual rightsholders. Third, outcomes turn on institutions. Policy must explicitly account for the dual economic effect of IPR, i.e., promoting innovation while potentially restricting access to essential goods, mandating a moderate, balanced approach. Institutional capacity requires creative mechanisms like Regulatory Sandboxes (ELRs) and targeted Public-Private Partnerships (PPPs) to iteratively test and implement responsive regulations,

⁶⁰ Cheong, B. Ch. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, 6. <https://www.frontiersin.org/journals/human-dynamics/articles/10.3389/fhumd.2024.1421273/full>

⁶¹ Png, 2022.

⁶² Fisher & Nathani, 2024.

particularly where existing legal systems face institutional and judicial deficits. Copyright authorities, courts, competition and data protection regulators must coordinate mandates, iterate guidance, and track distributional impacts if policy is to move beyond aspiration to implementation.

A decolonial lens has grounded these propositions in the political economy of AI. The paper has situated training practices and data governance within broader concerns about extractive dynamics, epistemic dominance, and infrastructural dependency that have historically marginalised the Global South. Against this backdrop, the normative direction is clear: community-centred data models, cultural and linguistic plurality in evaluation, and interoperable regional scaffolds are not optional extras but conditions for legitimacy. UNESCO's call for open, trustworthy datasets, collaborative data spaces, and accountable governance, alongside Global South capacity perspectives, provide a pragmatic ethics to law bridge for implementation. Read together, they support policy designs that (a) condition training on lawful access or an equivalent purpose-based limitation; (b) instantiate provenance, complaint and correction, and output side safeguards to deter overlearning; (c) replace atomised per work vetoes with standardized, collective mechanisms; and (d) invest in data quality and institutional capability commensurate with local linguistic and sectoral needs.

India's pathway, emblematic of broader Global South choices, illustrates both the opportunity and the imperative. Fair dealing under Section 52 of the Copyright Act, 1957 offers a doctrinal foothold but not a system. To move from case-by-case adjudication to predictable governance, India should legislate a TDM exception that recognises lawful access, frames reservations in usable, collective forms, and couples training side clarity with output side remedies, thereby supporting domestic research and enterprise while addressing creator concerns. More broadly, legislatures should align disclosure and opt out design with public interest goals, community governance, and capacity constraints, rather than reproducing models that externalise costs onto local innovators and cultural sectors.

The Global South does not enter this debate as a supplicant but as a norm entrepreneur. By legislating lawful access pathways, operationalising transparency that works for individual creators and collective entities, and institutionalising capacity for iterative governance, it can reframe AI—copyright from a site of "extractivism" to a domain of shared prosperity and cultural self-determination. That reframing is the work of law in its best tradition: to order markets to public purposes, to protect creativity without foreclosing progress, and to ensure that technological possibility is harnessed to social justice.

References

- Ayana, G., et al. (2024). Decolonizing global AI governance: Assessment of the state of decolonized AI governance in Sub-Saharan Africa. *Royal Society Open Science*, 11(8), 231994. <https://doi.org/10.1098/rsos.231994>
- Buick, A. (2025). Copyright and AI training data—transparency to the rescue? *Journal of Intellectual Property Law & Practice*, 20(3), 182–192. <https://doi.org/10.1093/jiplp/jpae102>
- Chauhan, K. (2025). Generative AI, text & data mining and the fair dealing doctrine: Examining the new problem with the old regime. *Journal of Intellectual Property Rights*, 30, 77–85. <https://doi.org/10.56042/jipr.v30i1.12652>
- Cheong, B. Ch. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, 6. <https://doi.org/10.3389/fhumd.2024.1421273c>
- Cyman, D., Gromova E., & Juchnevicius, E. (2021). Regulation of artificial intelligence in BRICS and the European Union. *BRICS Law Journal*, 8(1), 86–115. <https://doi.org/10.21684/2412-2343-2021-8-1-86-115>
- Fan, M. D. (2021). The right to benefit from big data as a public resource. *New York University Law Review*, 96, 1438–1492.
- Farhad, Sh. (2025). Passengers in flight: AI governance capacity in the Global South. *Digital Society*, 4, article number 39. <https://doi.org/10.1007/s44206-025-00195-6>
- Fisher, A., & Nathani, Z. (2024, February 19). *Association of Southeast Asian Nations provides guidance on AI governance and ethics*. Linklaters. <https://techinsights.linklaters.com/post/102j06t/association-of-southeast-asian-nations-provides-guidance-on-ai-governance-and-eth>
- Gladun, E., & Zadorin, M. (2023). The system of indigenous peoples' protection in BRICS states: An overview of legal and litigation support. *BRICS Law Journal*, 10(4), 121–141. <https://doi.org/10.21684/2412-2343-2023-10-4-121-141>
- Gromova, E. A., Ferreira, D. B., & Begishev, I. R. (2023). ChatGPT and other intelligent chatbots: Legal, ethical and dispute resolution concerns. *Revista Brasileira de Alternative Dispute Resolution*, 5(10), 153–175. <https://doi.org/10.52028/rbadr.v5i10.ART07.RU>
- Gromova, E., & Ferreira, D. B. (2024). On the way to BRICS+ digital sovereignty: Opportunities and challenges of a new Era. *BRICS Law Journal*, 11(3), 54–69. <https://doi.org/10.21684/2412-2343-2024-11-3-54-69>
- Gromova, E., Koneva, N. S., & Titova, E. V. (2022). Legal barriers to the implementation of digital industry (Industry 4.0) components and ways to overcome them. *Journal of World Intellectual Property*, 25, 186–205. <https://doi.org/10.1111/jwip.12215>
- Haryono, H., et al. (2024). Intellectual property rights in BRICS countries: Legal frameworks, enforcement challenges, and economic implications. *West Science Interdisciplinary Studies*, 2(12), 2494–2501. <https://doi.org/10.58812/wsis.v2i12.1550>

Kamaruddin, S., et al. (2025). The legal paradigm of generative AI in Malaysia and India: Problems and prospects. In *2024 International conference on artificial intelligence and emerging technology (global AI summit), Greater Noida, India, April, IEEE* (pp. 413–417). <https://doi.org/10.1109/GlobalAISummit62156.2024.10948005>

Kharitonova, Yu. S. (2023). Legal means of providing the principle of transparency of the artificial intelligence. *Journal of Digital Technologies and Law*, 1(2), 337–358. <https://doi.org/10.21202/jdtl.2023.14>

Kharitonova, Yu., Malik, N. S., & Yang, T. (2023). The legal issue of deterrence of algorithmic control of digital platforms: The experience of China, the European Union, Russia and India. *BRICS Law Journal*, 10(1), 147–170. <https://doi.org/10.21684/2412-2343-2023-10-1-147-170>

Muldoon, J., & Wu, B. A. (2023). Artificial intelligence in the colonial matrix of power. *Philosophy & Technology*, 36, Article 80. <https://doi.org/10.1007/s13347-023-00687-8>

Pandya, D. (2025, March 7). *The Global South AI copyright's test case: India*. Center for European Policy Analysis (CEPA). <https://cepa.org/article/the-global-south-ai-copyrights-test-case-india/>

Park, S. (2024). Bridging the global divide in AI regulation: A proposal for a contextual, coherent, and commensurable framework. *Washington International Law Journal*, 33(2), 216–269. <https://doi.org/10.2139/ssrn.4950781>

Peter, K. Yu. (2024). *The future path of artificial intelligence and copyright law in the Asian Pacific* (Texas A&M University School of Law Research Paper No. 24-18).

Png, M.-T. (2022). At the tensions of South and North: Critical roles of Global South stakeholders in AI governance. *FAccT Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 22, 1434–1445. <https://doi.org/10.1145/3531146.353320>

Quang, J. (2021). Does training AI violate copyright law? *Berkeley Technology Law Journal*, 36, 1407–1436. <https://doi.org/10.15779/Z38XW47X3K>

Quintais, J. P. (2020). The new copyright in the digital single market directive: A critical look. *European Intellectual Property Review*, 42, 28–41. <https://doi.org/10.2139/ssrn.3424770>

Tan, D., & Lee, T. (2021). Copying right in copyright law: Fair use, computational data analysis and the personal data protection act. *Singapore Academy of Law Journal*, 33, 1032.

Ueno, T. (2021). The flexible copyright exception for “non-enjoyment” purposes—recent amendment in Japan and its implication. *GRUR International*, 70(2), 145–152. <https://doi.org/10.1093/grurint/ikaa184>

Upadhyay, N. K. (2024). BRICS expansion: A comprehensive analysis of global power shifts and economic integration. *BRICS Law Journal*, 11(3), 5–8. <https://doi.org/10.21684/2412-2343-2024-11-3-5-8>

Xalabarder, R. (2024). *Scoping study on the practices and challenges of research institutions and research purposes in relation to copyright* (WIPO Standing Committee on Copyright and Related Rights, Document SCCR/44/4 REV).

Ziaja, G. M. (2024). The text and data mining opt-out in article 4(3) CDSMD: Adequate veto right for rightholders or a suffocating blanket for European artificial intelligence innovations? *Journal of Intellectual Property Rights*, 19(5), 453–459. <https://doi.org/10.1093/jiplp/jpae025>

Information about the authors

Arti Aneja (Tashkent, Uzbekistan) – Research Fellow, Tashkent State University of Law; Senior Assistant Professor, Campus Law Centre, Faculty of Law, University of Delhi (Delhi, 110007, India; e-mail: artianeja@clc.du.ac.in) – **corresponding author**.

Ankit Shrivastava (Noida, India) – Assistant Professor, Symbiosis Law School, Noida Campus, Symbiosis International (Deemed University), Pune, India; PhD Scholar, National University of Study and Research in Law, Ranchi (Noida, 201301, India; e-mail: ankit@symlaw.edu.in).