Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Universality Myth: Epistemic Exclusion, Structural Power, and AI Governance in the Global South

Amulya N

amulyan.bsc23@rvu.edu.in

RV University

Ashwini Prasad S

RV University

Research Article

Keywords: epistemic exclusion, AI governance, Global South, epistemic injustice, structural power, Brussels Effect

Posted Date: August 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10710836/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

The OECD AI Principles and the EU AI Act are supposed to apply to all; the UNESCO Recommendation is also touted as being universal. This paper suggests that the claim is ideological and not descriptive as most of the populations it purports to represent do not have contestation literacy, which is the ability to see an AI system as an object of governance and to evaluate a compliance claim against it, as opposed to usage literacy which the deployer has a vested interest in developing. This is the systematic lack of conceptual resources that allows to understand, challenge and influence the AI systems that dominate communities in the Global South. The paper combines document and policy analysis with Vallor's AI mirror theory, Fricker's and Dotson's epistemic injustice frameworks, Jasanoff's sociotechnical imaginaries, and Lukes's and Gaventa's structural power theory to compare India (elite capture and regulatory self-exemption) with sub-Saharan Africa (foreign-built systems adopted without institutional capacity), showing the pattern is structural, not exceptional. Five contributions follow. First, epistemic exclusion is a functional mechanism of structural power, self-reinforcing because closing it would expose AI deployers' compliance claims to scrutiny they avoid. Second, Vallor's mirror framework is extended: within her own scope for populations who perceive but misread the mirror, and through Fricker and Dotson for populations who cannot perceive it at all. Third, an internal mechanism of AI power supplements decolonial AI critique's predominantly external focus. Fourth, Jasanoff's sociotechnical imaginaries explain how science-fiction fear and techno-utopian narratives alike render structural AI harms imaginatively unavailable, structuring rather than prohibiting the question. Fifth, the unawareness/misinformation distinction extends to the facilitator enabling deployment: documented evidence from India's infrastructure-clearance process (Section 6.4) shows bureaucratic actors structurally excluded, by their role's categorical architecture, from the governance system they enable. Generalising this mechanism beyond the case examined here remains a falsifiable extension for future work.

1 Introduction

Global AI governance claims universality. The UNESCO Recommendation on AI, adopted unanimously by 193 member states, asserts that AI should serve the common good of humanity and be governed by principles applicable to all [1]. The OECD AI Principles, endorsed by the G20, call for AI to benefit "people and the planet" [2]. The EU AI Act is framed as setting the global standard [3].

This universality claim validates: it assumes that there is already some protection for the populations concerned, that rights have been recognised, that there is some governance that is not merely gestured at. This paper queries whether that is true, and if it is not, how that is maintained. This response emerged here is epistemic exclusion, a situation in which communities do not have the conceptual tools needed to process the impact of AI systems on their social worlds, to identify harmful vs beneficial deployments, and to voice concerns in a way that is recognizable to governance frameworks. This is not a lack of capacity in need of an educational intervention, but a structural issue, created by the exclusion from the

practices that generate and disseminate the knowledge of AI, and perpetuated by arrangements that do not have an incentive to change.

Throughout, "ideological" is in the sense of Lukes [4] and Gaventa [5]: claims that serve the interests of those in power by deliberately omitting alternative claims, not through deceit, but rather through "structural" means.

As Lukes [4] and Gaventa [5] suggest, it is not the intention of the paper—only the claim—of demonstrating that epistemic exclusion was intended to be used for power, but that: (a) it does exist and is documented; (b) those deploying the AI infrastructure benefit from it; and (c) there is no structural incentive to work against it. All three are demonstrated below. This does not preclude specific actors from exploiting the condition intentionally, as Section 5.1 documents; the functional claim concerns the condition's origin and self-reinforcement.

Condition (c) faces an obvious counterexample: technology companies do run large-scale AI and digital-literacy campaigns across the Global South, a real incentive to build some form of literacy. The relevant distinction is between usage literacy, comfort operating AI-enabled products, which deployers have every commercial incentive to build since it drives adoption and revenue, and contestation literacy, the capacity to recognise a system as a governance object and evaluate a compliance claim against it, which deployers have no comparable incentive to build since it invites the scrutiny condition (c) describes. If anything, this sharpens the claim: the exclusion analysed here is targeted, not incidental, with deployers investing in one literacy while starving the other, which is also why literacy campaigns proliferate without closing this gap.

Why AI specifically, and not any technology governing an epistemically excluded population (decolonial and TWAIL scholarship already documents comparable exclusion around pharmaceuticals, extractive industry, and environmental infrastructure), is worth stating. Three properties distinguish it. First, even for the trained experts, AI systems are often too opaque to be solved by education: The interpretability problem is that the barrier is not simply one of education, as the knowledge that is required may not be accessible to anyone outside of a system's own creators. Secondly, AI use is often physically invisible, for example, a model for scoring and prioritising loans or a model for determining eligibility for welfare benefits: there is no physical object that can stand in for it – as there is with a dam or a chemical plant – so in practice, there is no object to protest. Third, the introduction of AI has been more rapid than the regulatory framework that has taken many decades to mature around the more traditional physical infrastructure such as environmental impact assessment (Section 5.1, 6.4). These attributes are not all that make AI distinct from other decolonial cases, but they do make education- and disclosure-based remedies structurally less available in AI.

The term “Global South” here is offered as a structural, not geographical, term, referring to communities and states that have been historically excluded from the creation of AI knowledge and governance models, with the understanding that the term is contested and complex. This is relevant for the India

case: Section 6.1 demonstrates that even one state on one axis might be Global-South and Global-North on the other.

Two aspects of epistemic exclusion are retained all through. The role of unawareness and of misinformation is different, the evidence required is different, and how it is used in the argument is different. The two, however, are not necessarily contradictory, one person can have neither awareness of the working of AI nor certain imaginaries regarding what AI is. The functional character of epistemic exclusion is demonstrated by misinformation, available knowledge systematically diverts attention from structural harm. This harm is not just out of sight; it is diverted and focused elsewhere.

2 Related work: the gap this paper fills

The account below, based on theory, assigns different explanatory tasks to each of the frameworks rather than viewing all six as competing explanations for the same phenomenon. The primary analytical work of the establishment of what the epistemic condition is is done by Vallor [6], continued by Fricker [7] and Dotson [8]. The reason why it remains imaginatively invisible is explained in terms of the sociotechnical imaginaries of Jasanoff and Kim [9], and the reasons why it 'lives on' even without deliberate maintenance efforts are explained by Lukes [4] and Gaventa [5]. These are not parallel lenses on one, but sequential answers to various questions.

2.1 The decolonial AI literature

Mohamed, Png and Isaac [10] created the initial framework of decolonial AI, with a demonstration that the values of the sizable postindustrial societies are embedded within the development of AI. Birhane [11] took this one step further and brought it to Africa with ‘algorithmic colonisation’. The data colonialism thesis was coined by Couldry and Mejias [12] which links AI extraction with colonial resource appropriation. Effoduh and Mudeyi [13] explore how Third World Approaches to International Law (TWAIL) can be used to analyze and challenge how AI is governed and regulated, emphasizing a focus on the Global South and the need to shift the focus toward more rule-making rather than rule-taking. Roberts, Taddeo and Floridi [14] provide a political legitimacy framework which reveals how the lack of procedures and representation weakens the power of international AI governance at the population level: implications that this paper extends.

These narratives are directed towards external mechanisms and institutional dynamics (Western companies, Global North institutions, transnational data flows, regulatory export) that assume a certain condition of the supporting communities prior to their contact with such institutional architecture [15–18].

2.2 The epistemic injustice literature

In this paper, Mollema [19] proposes a taxonomy of epistemic injustice in AI systems. Safir et al. [20] find that there is a lack of awareness of AI governance in the Global South. Brown et al.'s [21] interviews with

22 practitioners across Africa, Asia, South America, the Caribbean, and minoritised practitioners in high-income contexts find that power asymmetries in AI ethics are epistemic as well as institutional, determining whose expertise counts as authoritative, and that practitioners from low- and middle-income countries are confined to technical roles instead of governance debate, directly confirming the structural condition this paper analyses.

None of these accounts connects epistemic exclusion to the universality claim in AI governance specifically: they document the condition, while this paper argues what it is for, and why it is self-reinforcing by design rather than contingently persistent.

2.3 Levels of analysis: epistemic extraction and epistemic exclusion

Monga's [22] analysis of epistemic extraction in AI governance is the closest work to this paper's terrain. Monga conceptualises epistemic extraction as a sequential mechanism (boundary encoding, evidence standardisation, category conformity) through which governance systems condition participation on translation into externally authorised knowledge categories, producing administrative illegibility for those who cannot translate. This paper adds a diagnostic contribution to Monga's response framework, which is already in place.

The two arguments are of different levels. Monga's concerns institutional design, who is the subject of whose knowledge in governance that becomes intelligible in the upstream recognition mechanisms? This paper concerns structural power and legitimation: why the universality claim persists even when fictive, and how epistemic exclusion at the population level enables uncontested AI expansion. Administrative illegibility, in Monga's account, operates on those trying to navigate institutions; hermeneutical injustice, in this paper's account, operates on those for whom AI has not yet become a legible category at all. Operating at the population level names the prior structural condition that makes institutional illegibility a second-order problem for most affected communities: sequential failure points in a single structural condition.

2.4 Two contributions to Vallor's framework

Vallor's [6] *The AI Mirror* argues that AI is dangerous not because it is alien to human values but because it is faithful to them: a mirror reflecting the hierarchies, exclusions, and assumptions of the societies that build it. Vallor herself identifies the populations this paper addresses. She observes that most AI training data overrepresents English-language text and speech and data from young, white, male subjects in the Northern Hemisphere, such that treating this as representative of humanity writes most people out of the human story; and that the workers who clean AI training data in countries including Nepal, Kenya, and Malaysia do not get to see their own lives and languages reflected in the systems they help build [6]. She is explicit that this constitutes an injustice. Ruggiu and Özdemir [23] each report on this population as generative AI's "microworkers," and describe how their invisibility to the systems they train is the very same invisibility that forecloses any contestation of the labour conditions (low pay, no protections) under

which that training is conducted. She makes it clear that this is a wrong. But she positions these populations as objects of the mirror's distortion (written out of the reflection, or exploited in its production) without asking what the epistemic condition is of a community that the mirror governs but does not know it is looking at a mirror.

This paper's extension operates in two analytically distinct steps, corresponding to the two-population distinction maintained throughout. For the misinformation population, this paper works within Vallor's existing framework: the science-fiction fear imaginary and the techno-utopian development imaginary both constitute encounters with a distorting mirror that is present and engaged with, just misread. This is Vallor's own territory; the contribution here is the extension to the specific Global South imaginaries through which AI is misread, and the systemic, not accidental, character of those distortions. Vallor's own account of the universal subjectivity presumed by her argument (the "we" capable of wresting back self-knowledge from AI mirrors) is one she describes as risking "intellectual imperialism" before defending it on grounds of shared emergency; for the misinformation population, that subject position is occupied but distorted.

For the unawareness population, the contribution is analytically prior. Vallor's framework presupposes a subject who can at minimum perceive being governed: someone for whom the mirror is in some sense present, even if misread. The unawareness population is the condition where that presupposition fails. There is no mirror to misread here; there is a mirror the person governed by it does not know exists. This differs from the opacity Vallor documents elsewhere in her account of algorithmic governance, the Dutch child-benefits scandal, US bail-risk scoring, where those affected know an algorithm is deciding but cannot access how [6]. The unawareness population lacks the antecedent fact of algorithmic involvement, not visibility into its mechanism. This condition is more precisely analysed through Fricker's [7] hermeneutical injustice and Dotson's [8] epistemic oppression, where the harm is the structural absence of the conceptual resources that would make any engagement possible, rather than misdirected engagement. Vallor names the injustice at the data level, in that these populations are written out of the reflection; this paper names it at the governance level, in that these populations are governed by a reflection they cannot perceive as such. This extends the ethical logic Vallor's framework already establishes; it does not critique that framework.

This framework has generated a growing secondary literature [24–26] engaging with its implications for AI ethics discourse, epistemic authority, and information trust; none of which extends either kind of contribution to Global South conditions.

3 Methodology

This paper employs document and policy analysis, drawing on existing empirical sources instead of primary fieldwork, structured across two analytical layers.

Positionality. This paper is written from an institutional location in Bangalore, India, affording documentary familiarity with domestic governance debates not equally available for the sub-Saharan

Africa case, and raising a reflexive concern this paper's own decolonial and TWAIL sources make a methodological commitment: producing academically legible critique of epistemic exclusion requires operating inside the same knowledge-production hierarchy under examination. The full positionality statement, including how the citation structure partly addresses the asymmetry, is given in Supplementary Information D.

Layer 1: document and policy analysis of universality claims. Close reading of the UNESCO Recommendation, the OECD AI Principles, the EU AI Act, and G20 governance statements identifies what universality claims are made, what they assume about the populations they govern, and what epistemic conditions those populations must possess for the framework to be operative: identifying a "universal subject" (the implicit epistemic figure governance frameworks assume) and documenting the gap between it and actual population conditions.

Layer 2: empirical documentation of the gap. Secondary empirical data are drawn from Pew Research Center [27] survey data, IAMAI/Kantar [28] internet usage data, parliamentary debate records [29], infrastructure documentation [30, 31], and practitioner research [21].

Construct and measurement: contestation literacy. Contestation literacy (Section 1) is a latent construct with no validated instrument measuring it directly; this paper infers its presence or absence through three converging proxies: recognition of AI as a governance object, approximated by AI-awareness survey data [27, 28]; exercised institutional recourse, by the outcome of specific attempts to invoke existing accountability mechanisms: the Right to Information request in Section 6.4, the parliamentary debate record in Section 5.1; and independent verification capacity, by the resourcing of civil-society and media bodies positioned to investigate compliance claims [29, 32, 33]. None of these measures the construct directly; the gap is addressed in Section 6.6.

Context selection follows a most-different-systems logic [34]: India and sub-Saharan Africa maximise variation on an internal/external mechanism dimension while sharing the structural outcome. India has sufficient domestic state capacity and technology industry to constitute its own governance failures; sub-Saharan Africa has weaker domestic capacity and greater external dependency. The internal/external distinction is analytical, not ontological: both mechanisms are present in both contexts, but what varies is which is analytically primary given each case's institutional configuration. The design is diagnostic: the shared structural outcome is produced through demonstrably different primary mechanisms in each case, and that the pattern holds across maximally different configurations is what makes it structural rather than situational.

Case scale is not so uniform with one state compared to a multi-state region; this is intentional, not a deviation from the comparative design. On the internal side, the state is the appropriate unit of analysis because domestic elite capture and regulatory self-exemption is a function of a single state mechanism. The external mechanism is actually regional, in some ways even more so than infrastructural, being operational at the level of the deployment of platforms across continents, and the sharing of

infrastructure including undersea cables [35], and not on the basis of a one-state-to-one-state design that would require a form of averaging-over of different cases.

Lukes' [4] third dimension is used to develop the structural power claim which is stated so as to be falsifiable, rather than simply illustrated. It would be undermined by either of two findings: (First) that for populations with low scores on the awareness measures in Section 5.1, there is effective grassroots contestation of AI governance, but not that their lower scores on those measures are evidence of epistemic exclusion; and (Second) that high-awareness populations suffer no accountability, which suggests that the causal work is done by collective-action costs or direct repression, not by epistemic exclusion. The lack of these is neither proof nor disproof of the claim; their presence elsewhere would be a counter example.

This falsifiability claim has another methodological qualification, in response to a criticism by Hay [36] of third dimensional power, and a further disconfirming condition, other than being institutionalised, described fully in Supplementary Information A.

4 Results: discourse analysis

These results below indicate that the paper makes a structural claim with four different mechanisms. Section 4 records the assumption made by universal subject approaches to the governance of AI and how the Brussels Effect internationalised it. In India, the domestic elite capture and the regulatory self-exemption constitute an institutional mechanism which repeats this situation, while in sub-Saharan Africa, the external foreign built system is adopted without internal institutional capacity. In Section 6, it ends with defining who is on the excluded epistemic position in each population (the unawareness/misinformation distinction) and then goes one step further, to the bureaucratic officials who approve the infrastructure that enable the deployment of AI in the first place. There are various levels of each mechanism: international, national/regional, population, and bureaucratic; each embodies a different structural condition, namely that of epistemic exclusion, which supports a universality claim that would not stand up against an informed, contesting public.

4.1 The universal subject

In the governance frameworks studied, there is a recurring theme, the subject of AI governance. This subject has capacities that are not explicitly recognized in the frameworks as prerequisites but as part of the architecture of the subject: policy literacy to be able to navigate complex governance documents, institutional access that allows for legal recourse if harmed, conceptual vocabulary that includes what an AI system is and the meaning of transparency and accountability in practice, and consent capacity that gives and withholds meaningfully informed consent.

These capacities are assumed, not built: the frames do not create conditions for the existence of this subject; they presuppose conditions for its existence, and are not explicit about this presupposition,

since otherwise they will have to face the gap between the subject (as assumed) and the conditions of the population (as actual).

This does not mean the four capacities have to be ubiquitous in the Global North, either—research on the digital divide describes communities there where digital capacities are lacking, such as policy literacy and meaningful consent over AI systems. It is the capacity vs. no capacity that is the relevant contrast, not the capacities themselves. Where the individual capacities of the Global North populations are missing, institutional infrastructure (funded consumer-protection regimes, civil-society litigation capacity, regulatory bodies with a proven track-record of enforcement, a media environment on its own resourced to investigate claims of compliance) acts as a partial substitute for individual unawareness: it places individual responsibility onto institutional accountability, without requiring public literacy generally. In fact, in the Indian case there is no corresponding backstop to safeguard consumers: action under the Digital Personal Data Protection Act [37] is now at the discretion of the executive without any judicial oversight, and the Internet Freedom Foundation is not backed by an equivalent body in Europe which has the authority to protect consumers. Section 5.1 records this deficiency. More accurately: AI literacy is not a universal absence of literacy among Global South citizens, but rather the impact of the absence is structurally different based on what infrastructures are in place to mitigate it.

4.2 Universality claims: textual evidence

The UNESCO Recommendation is explicitly universal, not just in terms of the scope of the AI systems and the countries in which they are used, but also in terms of the approach it encourages, which is based on human rights principles.The UNESCO Recommendation is universal, in that it applies to all AI systems, all countries, and all contexts, and that it calls for a human rights-based approach that goes beyond geographical boundaries [1]. The OECD Principles go further, and state that AI should work for "people and the planet"—for humanity [2]. The EU AI Act will apply to any AI systems sold on the EU market and the Brussels Effect will bring it beyond the EU boundaries to the rest of the world [38].

None of these documents questions whether their populations can comprehend the contents of these frameworks, nor holds any accountable for the mechanisms they set up. The gap is itself analytically important, or, it is assumed, rather than questioned, a universal subject.

4.3 The Brussels Effect: compounding, not merely coexisting

Bradford's [38] analysis records the ways in which EU regulatory standards get globalised by institutional influence and the size of the market. Siegmann and Anderljung [39] illustrate how this is applied when it comes to AI governance, and how the EU's new status as the global leader in AI regulation is demonstrated. The aim of this paper is to bring to the fore an element of the literature which has not yet been specifically named, namely the epistemic compounding of regulatory export.

The mechanism: Governance structures built in situations with particular institutional and epistemic conditions that are imported into situations where they are not present will appear to be in place and do

not actually help to strengthen governance. This is an enabling resource for those deploying AI—a statement of compliance gives credibility just where populations affected by the AI may not be able to assess whether it is true or false. The legitimating value of the credential is greatest where it can't be checked, at those contexts of epistemic exclusion.

This is a new structure, and not just a new norm issue of the Brussels Effect. Epistemic compounding is not only about coexistence of the Brussels Effect with epistemic exclusion – it is also about increasing the depth of exclusion, about making the gaps in legitimacy into legitimating resources which make exclusion ever more entrenched. According to Sequera Buzarquis [40], civil society organisations were not invited to negotiation tables for the 2025 AI Action Summit in Paris or the 2026 AI Impact Summit in India and the summits only led to the issuance of legitimating declarations that are not binding on states; at the same time, there was limited and symbolic civil society involvement in AI governance.

The Center for Data Innovation [41] study corroborates the finding that compliance with EU-derived regulations is much more likely to impose significant costs on firms in Africa, while governance gains are, to a large extent, European. The compounding mechanism is self-reinforcing just because it leaves no structural reason for the deployer to make efforts to close the epistemic gap: if he did, his claims of compliance would come under scrutiny that he does not want. It makes condition (c) of the structural power argument of Section 1 come true.

Crum [42] suggests that the AI Act might not generate a typical Brussels Effect as its procedural uncertainty and fragmentation contrasts with what fostered a worldwide impact on the GDPR. This paper's mechanism is distinct: it concerns de jure legislative adoption by governments lacking institutional capacity to implement the framework, not de facto market pressure on companies. Crum himself acknowledges de jure adoption remains operative even where the de facto effect is weaker; under his conditions, adoption by under-resourced governments should produce less operative protection, meaning epistemic compounding is worse, not better, than a clearer framework would produce. Crum's argument sharpens rather than undermines the compounding mechanism.

The benefit condition in the structural power test stated in Section 1 (that deployers benefit from epistemic exclusion) operates through two distinct channels, only one of which the compounding mechanism above addresses. The legitimating channel, developed above, applies to deployers who make compliance claims: states and platforms citing OECD or UNESCO alignment gain a credential whose value depends on the exclusion it exploits. A second, extractive channel applies independently of any compliance messaging: reduced friction on data collection and deployment where no informed consent or organised resistance is possible. Birhane's [11] finding that facial recognition systems trained on non-African data are deployed in African contexts with lower accuracy for African users, and Mwema and Birhane's [35] documentation that non-African companies built the undersea cable infrastructure carrying African data, show extraction proceeding not because compliance was successfully claimed but because no contestation was structurally possible. Together, these two channels set up condition (b)

for both deployers with and without the use of compliance credentials, filling in the gap the legitimating channel left open.

5 Results: empirical documentation

5.1 India: internal mechanism

Literacy condition. According to a report by Pew Research Center [27] on Global public opinion on artificial intelligence, only 14% of Indian adults report hearing “a lot” about AI; in the two countries surveyed where fewer than half of the population report hearing much about AI, India is tied for second place with Kenya. According to IAMAI/Kantar [28], India has 886 million internet users, 55% of which are rural, with the key internet usage for entertainment and social media. This figure serves as a minimum threshold, rather than an absolute metric for the four capacities outlined in Section 4.1: if a population hasn't yet encountered AI as a category at all, they may not have policy literacy, institutional access, conceptual vocabulary, or consent capacity related to AI, but the 14% limit does represent an upper bound on the proportion of the population that could have these capacities without having them.

Stanford HAI [43] records that 62% of Indian respondents view AI as more beneficial than harmful, a higher figure than in countries with substantially greater AI awareness. Read against the 14% awareness figure, which makes first-hand evaluation of deployed systems an implausible source of that optimism, the 62% figure corroborates what is independently documented at the institutional level: NITI Aayog's #AIforAll strategy and the wider government messaging constitute direct co-production of a developmental AI narrative, using the same document-analysis method (Section 3, Layer 1) applied elsewhere in this paper; Section 6.2 characterises this narrative as a techno-utopian sociotechnical imaginary in Jasanoff and Kim's [9] sense. The documentary claim is the primary basis for reading this as a state-produced techno-utopian imaginary; the survey data corroborate it rather than establish it (see Section 6.6 for the evidentiary weighting between the two, and for alternative readings of the 62% figure).

Cross-mechanism acknowledgment. India's internal mechanism does not operate without external power. Western-based technology firms have penetrated into India to a great degree; all while Indian engineers are creating systems based on Western training data and the Digital Personal Data Protection Act was discussed in a legislative environment influenced by the GDPR and OECD principles. Analytical primary is the internal mechanism, in which the Indian state has the capacity and domestic industry to produce their own failures of governance, but external power is always there and they interact. As with the India case, Gaventa's [5] discussion of structural quiescence, the factors that foster passage of the Act in the absence of informed public pressure, are the very same conditions that hindered public awareness that public pressure was warranted.

Governance failures: the Digital Personal Data Protection Act. The Act's Section 17 exempts government instrumentalities from its provisions on broadly defined national security grounds. Manwani and Rana [44] document 26 appearances of the phrase "as may be prescribed" across 44 sections,

delegating substantive rights definition to unreviewable executive rules. The Institute of Legal Studies [45] and the Internet Freedom Foundation [29] document that Section 44(3) removed the Right to Information Act's public interest exception for state-held personal data. The Internet Freedom Foundation [29] further records that the Bill received 52 minutes of debate in the Lok Sabha and 67 minutes in the Rajya Sabha, with 16 contributing members in total. An industry body's sustained lobbying relationship with the Ministry of Electronics and Information Technology, documented in the industry body's own policy papers [30], illustrates a second, non-state channel of elite capture, distinct from Section 17's state self-exemption: sustained organised access to the Act's drafting process itself. The two exemplify the same mechanism through different beneficiaries, government securing exemption from a law it wrote, industry securing access to shaping the law that remained, while the majority of the population, without meaningful AI awareness, had no equivalent route into either channel.

The political conditions for informed *public* pressure on this legislation did not exist at the moment of passage, though this claim requires a distinction Gaventa's [5] own study of Appalachian quiescence supplies: local elites and experts could and did recognise exploitation in the cases he examined; what did not occur was converting that recognition into political force, because the wider population's quiescence gave power-holders no reason to treat expert critique as costly to ignore. The evidence here fits this two-tier pattern rather than a simpler claim that no grievance was ever articulated. The Internet Freedom Foundation's [29] detailed legislative analysis, cited throughout this section, is itself a fully formed, published, expert-tier grievance, evidence not against epistemic exclusion but of exactly the tier at which grievance formation remained possible. What did not occur was population-tier political amplification: the 14% awareness figure indicates IFF's critique could not be made costly for legislators to ignore, because the broader public lacked the awareness needed to treat 52 and 67 minutes of debate as a scandal rather than routine process. The 14% figure and the 119-minute parliamentary record are therefore both expressions of the same structural condition at the population tier; together with the expert-tier critique's failure to gain political traction, they constitute the strongest available evidence that epistemic exclusion had a concrete, traceable consequence. The grievance existed; what was missing was a public able to make it costly to ignore.

An obvious alternative reading is available here: minimal floor debate and broad executive-exemption clauses may simply be a general feature of Indian legislative practice under a strong parliamentary majority, in which case the Act's thin debate reflects ordinary process rather than anything specific to AI-related epistemic exclusion. A direct, contemporaneous comparison speaks against this. The Constitution (One Hundred and Sixth Amendment) Act, 2023, better known as the Women's Reservation Bill, a high-public-salience measure passed within six weeks of the Digital Personal Data Protection Act, in the same parliamentary calendar year and under the same government, received over eight hours of Lok Sabha debate alone, with 132 members participating in discussion across both houses combined, against the Digital Personal Data Protection Act's 119 combined minutes and 16 contributing members [46, 47]. Parties and procedures are approximately constant when both bills passed during the same legislative year under the same government administration with large majorities.Public salience is the only element that changes when both bills are passed on large majorities in the same legislative year

under a single government administration. When the public is aware of the issue and a constituency is organized around it, debate time and participation grows in proportion, as do the epistemic-exclusion account would predict, even when the circumstances would otherwise suggest a thin floor debate. It is not a general pattern (see Section 6.6) but one comparison.

The Supplementary Information B provides a comparison of a same procedural class bill that addresses the constitutional-amendment confound, Government of National Capital Territory of Delhi (Amendment) Bill, 2023 [48–50].

The record of the implementation of the Act since its passage is in line with the above account, and reinforces that record. The Data Protection Board of India came into existence as a formal body in terms of notification of Digital Personal Data Protection Rules 2025 on 13 November 2025. The Board's Chairperson and Members have yet to be appointed and there has been no adjudication in the case, as the Central Government is still responsible for their appointments, tenure and removal, a situation that precisely the Srikrishna Committee recommended against [51] as it was finding that independence was the issue. Rule 23 of the DPDP Rules compounds the concern already documented above regarding Section 44(3): it permits the Central Government to require any data fiduciary to furnish information, including call logs, identity records, and location metadata, on national-security or similar grounds, without prior judicial authorisation and without notifying the person whose data is sought [52]. The Internet Freedom Foundation and the Editors Guild of India have separately warned that, absent any journalistic exemption, this creates a state-controlled channel for identifying sources, exposing working journalists, a population about as distant from this paper's account of epistemic exclusion as one can be, to exactly the contestation-literacy gap theorised here: knowing a right exists on paper is not the same as possessing the institutional standing to test it against unreviewable executive discretion [32, 33]. The exemption mechanism documented at the legislative stage above is, on this evidence, reproducing itself at the rule-making stage with a new population, rather than remaining a one-off feature of the Act's 2023 passage.

Infrastructure opacity. The Observer Research Foundation [31] documents up to 30 separate clearances required for data centre construction across fragmented central and state agencies, with no single coordinating authority. NASSCOM Community [30] documents that single-window clearance mechanisms have persistently failed to resolve this fragmentation. The governance consequence is direct: data infrastructure enabling AI deployment in India passes through an approval process in which each clearance authority understands the transaction before it (the building permit, the power allocation, the environmental clearance) but no authority in the chain has a comprehensive view of the governance system being assembled. This is the infrastructure-level instantiation of the unknowing facilitation mechanism developed in Section 6.4. Epistemic exclusion removes the public accountability that might otherwise compensate for this fragmentation: an informed public could demand disclosure of what AI systems are being built and why, but the literacy condition documented above makes that pressure structurally unavailable.

5.2 Sub-Saharan Africa: external mechanism

Cross-mechanism acknowledgment. The external mechanism in sub-Saharan Africa is not without internal facilitation mechanisms, since the domestic elites in Kenya and Nigeria benefit from links with Western technology companies and the fact that the countries comply not only with the OECD framework, but also by the fact that adherence is perceived to mean modernisation and investment attraction. Analytically, the external mechanism is a primary as domestic institutional capacity is less than that in India, while the internal agency is always present. The sub-Saharan Africa case provides evidence of Lukes's [4] third-dimensional power: the absence of AI awareness prevents grievances from being formed, not merely from being expressed.

Data extraction. Birhane [11] documents facial recognition systems trained on primarily non-African datasets and deployed in African contexts with systematically lower accuracy for African users. Mwema and Birhane's [35] analysis of undersea cable infrastructure shows that internet connectivity infrastructure in Africa was built predominantly by non-African companies whose interests in the data it carries diverge from those of African populations. Alphabet's Equiano and Meta's 2Africa cable systems alone are documented as carrying up to 95% of the continent's internet traffic, developed without consultation with the communities whose connectivity depends on them, and financed through arrangements Mwema and Birhane [35] characterise as infrastructure debt: African states paying private, foreign-owned corporations for access to cable capacity landed on their own coastlines. Effoduh and Mudeyi [13] document that AI models, trained largely on datasets reflecting the realities of global powers, misrepresent and marginalise diverse cultural and linguistic expressions from the periphery.

Governance without institutional capacity. Brown et al.'s [21] practitioner research finds a consistent pattern: AI governance frameworks from the Global North are adopted by Southern governments without institutional capacity to implement them, assigning Southern actors roles of compliant adoption rather than framework co-creation. One West African developer in their study reported having no laws or texts to regulate AI, and that developers were not aware they were creating biases. Barrett et al. [53] confirm from a practitioner-ethics standpoint that African voices have been routinely excluded from data science ethics discussions, and document that companies use saviour language to legitimate AI deployment in communities that cannot contest it. This is precisely the legitimation-without-accountability dynamic the Brussels Effect compounding mechanism predicts.

Sequera Buzarquis's [40] documentation of UN Global Dialogue on AI Governance processes confirms that the Global South was framed as a position rather than a place, showing that even in spaces claiming Global South inclusion, the terms of inclusion are defined from outside. Vijayakumar's [54] comparative analysis of nine Global South AI strategies confirms that even where such strategies have been published, they draw primarily on UNESCO and OECD frameworks, importing normative structures developed elsewhere.

A contemporary, quantified anchor corroborates the pattern described above. Ajuzieogu's [55] assessment of the AU Continental AI Strategy's first eighteen months of implementation (July 2024 to October 2025) found geographic concentration of AI-specific investment, 83% in four countries (Kenya,

Nigeria, South Africa, and Egypt), against a continental need the Strategy itself states at $500 billion, and characterised the Strategy's governance structures as largely aspirational relative to that need (a single-authored working paper by the institute's founder, hosted on ResearchGate rather than peer-reviewed). This does not by itself demonstrate epistemic exclusion at the population level, which is established above through Birhane [11], Brown et al. [21], and Barrett et al. [53]; it supplies a current, numeric complement to the qualitative pattern those sources document, showing that the resource and institutional gap the external mechanism predicts remains wide sixteen months into the Strategy's own implementation window.

Sub-Saharan Africa is not, on this evidence or any other, a single case, and the external-mechanism claim above does not require that it be one. Rwanda, Kenya, and Nigeria are all mentioned in the sources cited in this section and are clearly in different positions with regard to their posture—with Rwanda being one of the earliest countries to launch a national AI policy [56], while Kenya and Nigeria are both characterized by a combination of domestic elite engagement with foreign technology companies, documented above, and their own national AI policies. This heterogeneity is real and is not explained away here. What it does not threaten is the specific claim this paper makes about the external mechanism: that mechanism operates at the level of infrastructure and platform dependency, undersea cables, continent-spanning platforms, foreign-controlled data flows, which is regional and structural by its nature [35], not at the level of any individual state's declared policy ambition. A state can pursue an assertive national AI strategy on paper while its population remains dependent on infrastructure it does not control and cannot audit; national policy variation is a claim about governments' stated postures, while the external mechanism is a claim about the material and epistemic conditions those postures are stated within. The two are compatible, and Section 6.6 identifies systematic within-region comparison as a target for the fieldwork this paper's design does not undertake.

6 Discussion

6.1 The structural pattern

Table 1 summarises the structural pattern across both comparative contexts.

Table 1
Structural pattern of epistemic exclusion across comparative contexts. The facilitator-level extension of this framework is addressed separately in Section 6.4.

Dimension	India (internal mechanism)	Sub-Saharan Africa (external mechanism)	Shared structural condition
Primary driver	Domestic elite capture; state self-exemption from own data law	Foreign-built systems; externally defined governance categories	Communities governed by AI they cannot understand or contest
Governance failure	Act passed with 119 minutes of parliamentary debate and plenary state exemption	Adoption without institutional capacity; practitioners working without regulatory frameworks	Frameworks that confer legitimacy without operative protection
Literacy condition	14% "heard a lot" about AI (floor measure, Section 5.1); 62% positive affect (corroborating a documented #AIforAll imaginary)	Citizens preoccupied with immediate survival needs rather than AI	No informed public pressure possible
Brussels Effect consequence	Act claimed as data protection compliance; low awareness precludes public evaluation	Adoption of frameworks without institutional capacity; no operative regulation reported	Legitimacy conferred precisely where it cannot be scrutinised
Cross-mechanism (secondary)	External power present	Internal facilitation present	Both mechanisms co-present; primacy varies by configuration

India's position here is worth making explicit, since it demonstrates rather than complicates the structural, non-geographic definition of Global South given in Section 1. Domestically, the Indian state occupies the excluded-population side of the pattern documented in this section: its own citizens are epistemically excluded from the AI governance the state produces (Section 5.1). Internationally, the same state apparatus occupies the opposite position: India co-hosted the 2026 AI Impact Summit, a legitimating venue of exactly the kind Section 4.3 analyses, at which Global South civil society participation remained largely symbolic [40]. A single state can therefore be structurally Global-South-positioned on one axis (domestic epistemic exclusion) while Global-North-adjacent on another (international co-production of the universality claims that compound exclusion elsewhere). This connects the paper's first, third, and fourth contributions: the internal mechanism documented in India (contribution 3) and the Brussels Effect compounding mechanism at the international level (contribution 1, via Section 4.3) are two faces of the same state actor's dual position, not separate phenomena.

6.2 The two-population distinction

For the unawareness population, those who have not encountered AI as a meaningful category and who make up a majority of the Indian population given that fewer than half have heard even a little about it

[27], Fricker [7] and Dotson [8] apply jointly: Fricker identifies hermeneutical injustice, the condition of having social experience obscured because the shared concepts to name it are unavailable; Dotson elevates the analysis to the structural level, showing institutions systematically impede people's capacity to generate epistemic resources, so unawareness is produced by structural exclusion rather than individual ignorance.

For the misinformation population, those who have encountered AI primarily through specific cultural frames, Jasanoff and Kim's [9] sociotechnical imaginaries apply. Two variants are distinguishable, and both render structural harm imaginatively unavailable: imaginaries structure the horizon of thinkable futures, determining what is taken as possible, probable, and desirable while leaving outside that frame what cannot be generated as a question within it [9]. The first is a science-fiction fear imaginary (AI as sentient threat), directing attention toward imagined catastrophe and away from the mirrored, operational harm that constitutes AI's actual danger [6]. The second, more consistent with India's government-led development narrative and the 62% optimism finding, is a techno-utopian imaginary (AI as development solution), directing attention toward imagined benefit and equally away from the same structural harms already occurring [6]. The techno-utopian variant is the more direct application of Jasanoff's framework, co-produced through state governance institutions; the science-fiction variant extends the framework to cultural imaginaries not produced through formal governance. The techno-utopian variant's evidentiary basis is developed in Section 5.1: it rests primarily on documentary analysis of state-produced AI narratives such as #AIforAll, with attitudinal survey data corroborating rather than founding it.

Both imaginaries make certain AI futures vividly available while leaving structural harms outside the frame of what the imaginary generates as a question. A community experiencing techno-utopian AI enthusiasm does not ask whether a given algorithm fairly assesses loan applications, not because the question is forbidden but because the available imaginative frame does not generate it: no enforcement of silence is required, only the absence of the frame that would make the question thinkable. The difference from the unawareness population is qualitative rather than merely quantitative. Unawareness operates through absence: no category exists for a grievance to attach to, the pure third-dimensional condition in Lukes's [4] sense. Misinformation operates by occupying the category with content that produces active affective investment (fear or hope) directed elsewhere; the obstacle is not a vacancy waiting to be filled but a positive investment that must first be dislodged. The two populations accordingly require different remedies and different evidence: closing an absence is a matter of access; redirecting an occupied imaginative frame is a matter of confronting an existing affective commitment.

To be clear, this account does not claim that epistemically excluded populations never resist algorithmic governance; Fricker's [7] framework does not require total incapacity to sense harm, only incapacity to name it in terms institutions recognise as a legible grievance. The gig-worker and farmer resistance Section 3 reads as consistent with, rather than disconfirming, this paper's thesis illustrates why: workers can feel and resist the immediate, embodied experience of algorithmic management without possessing, or being granted institutional recognition for, the specific conceptual vocabulary

(transparency, explainability, algorithmic accountability) that Section 4.1 identifies as a precondition of the universal subject governance frameworks assume.

This account should be distinguished from a different, agency-first tradition, exemplified by Mdhlane and Masisi [57], who advance epistemic sovereignty as a normative and political horizon for decolonising AI governance in the Global South. That tradition is not in tension with the diagnosis here so much as pitched at a different level of the same problem. Epistemic sovereignty names what affected communities should build toward: reclaimed authority over how AI is developed, governed, and made to serve local ends, exemplified in practice by projects such as Masakhane's community-driven work on African-language natural language processing and LatamGPT's development as an act of technological reappropriation in Latin America [58]. This paper's claim is diagnostic rather than normative: it identifies the population-wide epistemic condition such reclamation projects must work against, and explains why that condition persists structurally rather than through any deficiency in Global South agency. The two accounts answer different questions: a successful sovereignty project changes who holds the conceptual resources and institutional standing this paper shows most of the governed population currently lacks, but does not itself show the population-wide condition was other than this paper describes prior to that project's success, any more than the Internet Freedom Foundation's expert-tier critique in Section 5.1 shows population-tier epistemic exclusion was absent.

6.3 Why standard responses fall short

Three standard responses to AI governance gaps in the Global South each address a symptom while leaving the structural condition intact.

More AI education is the most common prescription. Even India's own national AI strategy operates without mechanisms ensuring that AI education reaches historically marginalised communities in forms calibrated to their epistemic contexts [59]. Education that improves AI literacy without changing the power relations producing epistemic exclusion reproduces the existing structure with better-informed participants who still cannot change its terms, and this is not solely a claim about uneven implementation: even a perfectly, equitably delivered AI-literacy programme would not by itself confer institutional standing, since literacy is a capacity of individuals while rule-making access, funded litigation capacity, and enforcement power are properties of institutional position. The Digital Personal Data Protection Act's Section 17 exemption illustrates this directly: it concentrates enforcement discretion in the executive by statutory design, a feature no degree of public AI literacy would alter, since the Act does not condition that discretion on public understanding. Effoduh and Mudeyi's [13] rule-taker/rule-maker distinction names the underlying point: education can move a population from unaware to aware without moving it from rule-taker to rule-maker, since awareness and institutional standing are different goods a literacy intervention alone does not connect. Monga's [22] proposed response, interoperability that does not require prior ontological standardisation, offers a more targeted alternative by enabling participation without epistemic surrender.

More diverse voices in governance frameworks is the second standard response. Effoduh and Mudeyi [13] document that the Global South is institutionally relegated to passive rule-taker rather than rule-maker; Sequera Buzarquis [40] confirms that even in explicitly Global South-inclusive governance processes, civil society participation remained largely symbolic. Adding Southern voices to frameworks built around a universal subject does not change the subject position; it diversifies the people ratifying it.

Better regulation is refuted by the Digital Personal Data Protection Act itself: comprehensive legislation modelled on international best practice, passed through a formal legislative process, resulting in state self-exemption and 119 minutes of debate. The problem was not regulatory quality but the conditions under which regulation was produced, conditions in which the affected population lacked the epistemic resources to demand that regulation serve their interests.

None of this amounts to a claim that nothing can be done, only that education, representation, and better regulation are not, by themselves, sufficient, since each leaves the underlying epistemic asymmetry intact. A contestation-literacy intervention, by contrast, would need institutional standing, not just awareness: funded litigation capacity independent of executive discretion for its resourcing or targets; mandatory AI-impact categories within existing infrastructure clearance and environmental review processes, currently absent from the relevant checklists as Section 6.4 shows; and a statutory right of access to the technical basis of an AI-enabling infrastructure approval that does not depend on the approving authority itself possessing that information, which Section 6.4's evidence suggests cannot currently be assumed. Specifying and evaluating such an intervention is normative and applied work this diagnostic paper does not undertake; naming it here shows the diagnosis has a determinate shape, rather than dissolving into a general call for more education or representation once the standard responses are set aside.

6.4 The unknowing facilitation mechanism

So far the paper's argument has been at the population level: epistemic exclusion takes away the informed public pressure and allows deployment of AI without democratic friction. This section pinpoints a process that takes place at the same time at the bureau level.

There's no need for corruption in the base case for this mechanism, and it's best to say it without. Data centres can apply for up to 30 different approvals from various fragmented central and state government agencies [31], but they do not undergo environmental review under the "power draw" and "water stress" categories of the general building construction category, and are not considered for “grid impact” or “AI workloads” [31]. The situation in which the approving official receives a clearance under this categorical gap puts him in the exact epistemic position which is analyzed in this section; the governance system which is erected as a result is not, since there isn't a category available to the approving official's role that captures that information. This is unknowing facilitation, the epistemic exclusion that takes place at the level of the bureaucracy, as well as at the level of populations, by actors with the same epistemic condition as the populations they are supposed to serve, unless there is a legal transaction, in which case it takes place at the level of the individual.

Documented corruption in the same approval pipeline, service-delivery bribery at scale [60], is an aggravating instance of this baseline condition, not a precondition for it: where a transaction is also corrupt, the official has an additional reason not to look beyond the immediate consideration, but the categorical gap already withholds the relevant information from every official processing the clearance, honest or not. Framing the mechanism around corruption would risk narrowing an otherwise general finding to a subset of facilitators and reproducing a corrupt-official trope this paper's own decolonial sources work against; the categorical evidence below is offered as the sufficient case, with corruption noted only where it compounds rather than constitutes the mechanism. Power does not require knowing instruments, only available ones.

The facilitator's epistemic condition. Unknowing facilitation as stated identifies that the approving official's epistemic access is bounded (transaction legible, system illegible) without specifying what kind of boundedness this is or how it is produced. The two-population framework developed above for the governed population answers exactly this question; extended to the facilitator, the same two mechanisms recur, with different evidentiary weight attaching to each.

The first is a facilitator-level analogue of the unawareness population: a positional claim in Dotson's [8] structural sense, evidenced by tracing the categorical architecture of the role rather than any individual's cognitive state. The government's own record traces that architecture directly: a 2026 parliamentary statement acknowledged AI infrastructure's material energy and water footprint while confirming that data centres remain outside the direct coverage of India's Environmental Impact Assessment Notification of 2006 [61, 62]. The gap does not resolve even at the ministry most likely to hold the relevant knowledge: when the Internet Freedom Foundation filed a Right to Information request asking about data centre environmental impact, it was forwarded to the Ministry of Electronics and Information Technology, which responded that it possessed no information about data centres at all [63]. No institutional channel carries the concept needed to evaluate a clearance as an AI-governance decision, rather than a construction permit, to any actor in the approval chain. Civil-servant AI-literacy initiatives do not close this gap, since they train officials to use AI tools for service delivery rather than to recognise infrastructure-clearance decisions as AI-governance acts.

The second is a facilitator-level analogue of the misinformation population, resting on a different, independently documentable evidentiary basis than the population-level imaginary discussed in Section 6.2, rather than an extension by analogy from it. Single-window clearance schemes and streamlined-approval frameworks are policy instruments written specifically for and about the bureaucratic apparatus: NASSCOM Community's [30] documentation of persistent single-window reform efforts, and the Observer Research Foundation's [31] account of the clearance pipeline, describe an administrative discourse framing data centre approval as an instrument of growth to be accelerated, addressed to officials in their professional capacity rather than to the public. This constitutes a second, independently observable imaginary: not an inference that the public-facing #AIforAll narrative probably also reaches officials, but a distinct administrative-growth imaginary documented through the same policy instruments that structure the clearance process itself. Its documentary existence is established here;

whether it demonstrably shapes individual officials' judgement in specific decisions is a further question, addressed in Section 6.6.

Two boundaries apply to both analogues. The claim is positional, not psychological: it characterises the epistemic resources available to a role, evidenced by what that role's governing categories and its ministry's own disclosures contain, independent of what any specific person believed at the moment of approval. And where concrete cases appear in the literature, such as documented data centre environmental clearances processed without disclosure of water source or consumption [63], these illustrate how thin the categorical review process is; they say nothing about the state of mind of the officials who processed them.

Read together, these two mechanisms give the colonial-intermediary parallel developed below a more precise referent. What would confirm or disconfirm this extension is practitioner interviews with clearance officials, conducted analogously to Brown et al.'s [21] work with AI practitioners, asking whether officials processing data centre applications understand them as AI-governance decisions. The absence of any such research is itself a finding: the unknowing facilitator is a structural position that the existing empirical literature has not yet examined directly.

The colonial parallel. This parallel is analytically precise rather than merely rhetorical. Fanon [64] documented the colonised bourgeoisie who enabled colonial infrastructure while sharing the epistemic condition of the population they nominally represented. Mbembe [65] identifies the structural position of local intermediaries who accepted payment or position to enable infrastructure whose full consequences were opaque to them. Chatterjee [66] makes the mechanism explicit: local intermediaries were useful precisely because they were legible to the colonising administration while epistemically positioned identically to the population they governed with respect to what the infrastructure would actually do. Couldry and Mejias's [12] data colonialism thesis establishes that the extraction logic is constitutively, not merely metaphorically, continuous with colonialism; Birhane's [11] algorithmic colonisation framework establishes the same continuity in the African context. The correspondence between the unknowing facilitator and the colonial intermediary is best stated as a formal structural isomorphism rather than a loose analogy or a claim of historical identity. Both positions share five features: legibility to the immediate transaction (the bribe or the permit, the colonial directive); instrumental positioning between power and the governed population; epistemic exclusion from the full purpose or consequence of what is thereby enabled; enabling extraction whose downstream effects are invisible to the position at the point of action; and formal authority over a population whose epistemic condition the position-holder nonetheless shares. The claim is not that colonial administration and Indian environmental clearance are the same institution, but that they instantiate the same five-part epistemic architecture, justifying treatment of Fanon's, Mbembe's, and Chatterjee's accounts as directly informative rather than merely suggestive for the contemporary case.

The theoretical gap this fills is specific. Standard corruption literature [67–69] assumes a knowing corrupt actor who understands what they enable and chooses to enable it for gain; the AI governance

literature assumes either an ignorant public or a knowing governance failure, but has not theorised a governance actor sharing the governed population's own epistemic condition. Scott's [70] local officials implement a state scheme whose consequences they cannot see, but implement the scheme of a state they serve, unlike the unknowing facilitator, who enables a transaction whose AI-governance consequences are invisible to them. Lipsky's [71] street-level bureaucrats exercise discretion under incomplete information but know they are exercising judgement; the unknowing facilitator's role signals no judgement is being exercised at all. Eubanks's [72] caseworkers operate within automated systems they do not understand, but are employees of the institutions that built those systems, whereas the unknowing facilitator is outside the system while enabling it. Crawford [73] documents AI's opacity to proximate actors, but proximate to production, not to deployment through routine bureaucratic approval.

The contribution in one sentence: the official approving a routine construction permit may know they are approving a contract, but they do not know they are enabling a governance system, and that gap between what is transactionally legible and what is consequentially legible is where epistemic exclusion operates at the point of facilitation. This is a mechanism absent from AI governance literature, one that documented corruption in the same pipeline can aggravate but is not required to produce.

6.5 Positioning against Monga

The distinction established in Section 2.3 between Monga's institutional-design focus and this paper's population-level diagnosis sharpens once Monga's mechanism is examined in detail. The three-filter account Monga develops, boundary encoding, evidence standardisation, category conformity, operates on claimants already attempting to engage a governance system, and so does the wider pattern it names the Guru Dakshina Problem, after the Eklavya narrative in which access to knowledge is conditioned on a demanded tribute. Applied to governance, the term describes systems that make access conditional on translating local knowledge into externally authorised categories: translation failure produces exclusion, and translation success produces dependency on terms not the claimant's own. In Monga's own Chhattisgarh forest-rights example, the applicant whose claim is administratively rejected because customary tenure does not map onto cadastral categories has, by definition, already crossed the threshold this paper's population-level thesis treats as the harder problem. Section 6.4's facilitator-level extension supplies a third position in the same structural space, neither the excluded population this paper's core argument addresses nor the illegible claimant Monga's account addresses, but the approving official who lacks the categorical resources to recognise what is being approved in the first place. Read together, the three accounts describe a single condition recurring at three institutional depths: before contact (this paper's core thesis), at the point of contact (Monga), and behind the point of approval (Section 6.4). The connection runs deeper than shared subject matter, though it is narrower than a shared institutional proposal: invoking Monga's community appeal mechanism still requires a claimant to recognise that a categorical mismatch has occurred, which is one instance of the contestation literacy this paper's Section 6.3 argues is currently scarce. An operational remedy of the kind Monga proposes has a smaller population able to activate it for as long as that prior condition goes

unaddressed, which is a point of complementarity rather than a claim about what Monga's framework itself presupposes or requires.

6.6 Limitations

This section consolidates the paper's evidentiary caveats in one place. Results and Discussion above state findings directly; where a finding rests on thinner ground than others, that qualification is developed here rather than distributed through the sections making the claim.

Scope and design. This paper relies on secondary sources and policy documents, not original fieldwork, and compares two contexts rather than drawing on a systematic sample of Global South cases. Extending the pattern to further contexts (Latin America, Southeast Asia [74, 75]), enlarging the legislative-comparison sample beyond the two contemporaneous bills used in Section 5.1, and testing within-region variation (Section 5.2) are targets for future work; the full scope caveats are detailed in Supplementary Information C.

Construct-proxy gap. Contestation literacy is operationalised in Section 3 through convergent proxies rather than measured directly. A validated contestation-literacy scale, administered to affected populations, is the clearest single piece of empirical work this paper's design cannot supply.

Evidentiary base: the two populations. The unawareness population is well-grounded in Pew and IAMAI survey data; the misinformation population's evidence is more uneven, particularly for the science-fiction fear variant, which lacks an India-specific empirical anchor. Alternative readings of the underlying survey data, and the case for retaining the two-population structure despite this unevenness, are detailed in Supplementary Information C.

Evidentiary base: the facilitator-level extension. Its unawareness analogue (Section 6.4) is well-grounded in documented clearance categories and a ministry's own disclosure. Its misinformation analogue is based on a plausible, but not yet confirmed, inference regarding individual officials' judgments brought about by administrative discourse. The full reasoning is given in Supplementary Information C.

Evidentiary asymmetry between contexts. India is substantiated from the named legislation, agencies and parliamentary records, while the primary access for sub-Saharan Africa is thinner, but the literature and secondary sources of practitioners are thicker. We propose this asymmetry as a hypothesis but do not claim that it is verified by this analysis, which is a structural reading. Supplementary Information C provides such a reading.

7 Conclusion

This paper has argued that AI governance universality is structurally facilitated by epistemic exclusion: the systematic absence of the epistemic conditions required for affected populations to participate meaningfully in AI governance. Asserting universality in their absence legitimates AI deployment without

the accountability genuine universality would require. In Lukes's [4] terms, structural power that prevents affected populations from forming grievances denies the precondition for meaningful democratic participation: epistemic exclusion as power in the third-dimensional sense.

The paper's five contributions can be summarised as follows. First, epistemic exclusion is identified as a functional mechanism of structural power rather than a capacity deficit (Section 1), established through two independent benefit channels, legitimating and extractive (Section 4.3), and evidenced agentically as well as structurally: documented industry lobbying secured access to the Digital Personal Data Protection Act's drafting process, alongside the state's own self-exemption under the Act (Section 5.1). The Brussels Effect's epistemic compounding is the engine of this self-reinforcement, satisfying all three conditions of Lukes's structural power test. The remaining contributions elaborate this mechanism (second, fourth), supply its empirical test across two contexts (third), and extend it as a falsifiable hypothesis to a new domain (fifth).

Second, this paper makes two contributions to Vallor's [6] AI mirror framework. Vallor identifies the populations this paper addresses, naming them as objects of the mirror's distortion, but does not ask what it means to be governed by a mirror one cannot perceive. This paper moves those populations from object to subject: for the misinformation population, through an extension within Vallor's existing scope to Global South imaginaries; for the unawareness population, by naming a structural presupposition Vallor's framework requires but does not state, analysed through Fricker [7] and Dotson [8]. Where Vallor locates the injustice at the level of data, this paper's extension relocates it at the level of governance.

Third, this paper articulates an internal mechanism of AI power (domestic elite capture, regulatory self-exemption, infrastructure opacity) that supplements decolonial AI critique's predominantly external focus.

Fourth, this paper applies Jasanoff and Kim's [9] sociotechnical imaginaries to show why both science-fiction fear and techno-utopian development imaginaries render structural AI harms, the mirrored harms Vallor [6] locates at the heart of AI's danger, imaginatively unavailable. Neither imaginary requires prohibitive force; each works by failing to generate the question, not by suppressing it.

Fifth, extending the unawareness/misinformation distinction to the facilitator enabling deployment (Section 6.4), this paper identifies an unknowing facilitation mechanism, established independently of corruption through documented categorical gaps in clearance review and a ministry's own disclosure that it holds no relevant information. A parallel administrative-imaginary channel is separately documented, though its effect on individual officials' judgement remains a hypothesis for future fieldwork. The mechanism is formally isomorphic, in five specified respects, to the colonial intermediary in Fanon [64], Mbembe [65], and Chatterjee [66], and is falsifiable through practitioner interviews of the kind Brown et al. [21] conducted.

Effective and meaningful engagement in the governance of AI must be a prerequisite for its legitimate use, rather than an aim after the fact. This paper is not prescriptive in nature, but it is diagnostic in nature

that has concrete implications down stream. The self-reinforcement argument implies deployers require an external check on affected populations' epistemic conditions, since no structural incentive exists for them to assess it themselves. The third-dimensional power argument implies governance bodies conditioning deployment approval on prior epistemic-condition assessment would interrupt structural power at the point where it currently goes unexamined. The Brussels Effect compounding mechanism implies compliance credentials should require demonstration that affected populations can engage with the framework complied with, since otherwise the credential legitimates precisely the exclusion this paper identifies. Universal governance mechanisms that are based on an implicit subject which many of those who are affected do not have do not just leave them unprotected, they also lend legitimacy to the conditions that exclude them.

Declarations

Author Contribution

A.N. and A.P.S. conceived the study, conducted the document and policy analysis, and wrote the manuscript. All authors reviewed the manuscript.

References

1. UNESCO (2021) Recommendation on the Ethics of Artificial Intelligence. UNESCO, Paris
2. OECD (2019) Recommendation of the Council on Artificial Intelligence. OECD, Paris
3. European Union (2024) Regulation (EU) 2024/1689 (AI Act). Official Journal of the European Union, Brussels
4. Lukes S (2005) Power: a radical view, 2nd edn. Palgrave Macmillan, London
5. Gaventa J (1980) Power and powerlessness. University of Illinois Press, Urbana
6. Vallor S (2024) The AI mirror: how to reclaim our humanity in an age of machine thinking. Oxford University Press, Oxford
7. Fricker M (2007) Epistemic injustice: power and the ethics of knowing. Oxford University Press, Oxford
8. Dotson K (2014) Conceptualizing epistemic oppression. Soc Epistemol 28:115-138. https://doi.org/10.1080/02691728.2013.782585
9. Jasanoff S, Kim SH (eds) (2015) Dreamscapes of modernity: sociotechnical imaginaries and the fabrication of power. University of Chicago Press, Chicago
10. Mohamed S, Png MT, Isaac W (2020) Decolonial AI. Philos Technol 33:659-684. https://doi.org/10.1007/s13347-020-00405-8
11. Birhane A (2020) Algorithmic colonization of Africa. SCRIPTed 17:389-409. https://doi.org/10.2966/scrip.170220.389
12. Couldry N, Mejias UA (2019) The costs of connection. Stanford University Press, Stanford

13. Effoduh JO, Mudeyi MO (2025) TWAILing AI governance: will third world countries be left behind? TWAILR Reflections 74
14. Roberts H, Taddeo M, Floridi L (2026) A framework for evaluating global AI governance initiatives. Glob Policy 17:189-199. https://doi.org/10.1111/1758-5899.70164
15. Abraham I (2026) AI ethics in postcolonial contexts: a critical synthesis of infrastructures, power, and governance. AI Soc. https://doi.org/10.1007/s00146-026-03153-z
16. Mariniello M (2026) Colonialism 4.0 and Coloniality 4.0. AI Soc 41:3621-3635. https://doi.org/10.1007/s00146-025-02736-6
17. Eke D, Chavarriaga R, Stahl B (2025) Decoloniality impact assessment for AI. AI Soc 41:2213-2228. https://doi.org/10.1007/s00146-025-02649-4
18. Nemorin S, Bonami B (2026) AI ethics through a decolonial lens: what does AI ethics look like if we take seriously the push to decolonise it? AI Soc. https://doi.org/10.1007/s00146-026-02935-9
19. Mollema WJT (2025) A taxonomy of epistemic injustice in artificial intelligence. AI Ethics 5 https://doi.org/10.1007/s43681-025-00801-w
20. Safir AH, McInerney K, Blackwell AF, Debnath R (2025) Distributive epistemic injustice in AI ethics: a co-productionist account of Global North-South politics in knowledge production. In: Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, pp 2009-2024. https://doi.org/10.1145/3715275.3732136
21. Brown V, Larasati R, Kwarteng J, Farrell T (2026) Understanding AI and power: situated perspectives from Global North and South practitioners. AI Soc 41:3981-3996. https://doi.org/10.1007/s00146-025-02731-x
22. Monga HP (2026) Epistemic extraction in AI governance: institutional design and participation in the Global South. SSRN preprint, ssrn-6835741
23. Ruggiu D, Özdemir B (2026) Behind the AI's gaze: generative artificial intelligence, microworkers and digital slavery. NanoEthics 20:2. https://doi.org/10.1007/s11569-025-00489-6
24. Bate GW, Lloyd R (2025) Review: The AI Mirror. Postdigit Sci Educ. https://doi.org/10.1007/s42438-025-00546-z
25. Cox DM (2025) AI as mirror: review essay. Crit Humanit . https://doi.org/10.33470/2836-3140.1089
26. Smith J, Bakiner O (2025) Review: Shannon Vallor, The AI Mirror. Bus Ethics Q 35:710-714. https://doi.org/10.1017/beq.2025.10101
27. Pew Research Center (2025) How people around the world view artificial intelligence: spring 2025 global attitudes survey. Pew Research Center, Washington DC
28. IAMAI/Kantar (2025) Internet in India report 2024. IAMAI, Mumbai
29. Internet Freedom Foundation (2023) DPDPB 2023 in the Parliament: a legislative analysis. Internet Freedom Foundation, New Delhi
30. NASSCOM Community (2026) Data centre policies of India. NASSCOM, New Delhi

31. Observer Research Foundation (2025) Sovereign data centre infrastructure. Observer Research Foundation, New Delhi
32. Internet Freedom Foundation (2025) DPDP Rules 2025: state access provisions and press freedom. Internet Freedom Foundation, New Delhi
33. Editors Guild of India (2025) Statement on state access provisions under the Digital Personal Data Protection Rules, 2025. Editors Guild of India, New Delhi
34. Przeworski A, Teune H (1970) The logic of comparative social inquiry. Wiley, New York
35. Mwema E, Birhane A (2024) Undersea cables in Africa. First Monday 29 https://doi.org/10.5210/fm.v29i4.13637
36. Hay C (1997) Divided by a common language: political theory and the concept of power. Politics 17:45-52. https://doi.org/10.1111/1467-9256.00033
37. Digital Personal Data Protection Act (2023) Act No 22 of 2023. Government of India, New Delhi
38. Bradford A (2020) The Brussels effect: how the European Union rules the world. Oxford University Press, Oxford. https://doi.org/10.1093/oso/9780190088583.001.0001
39. Siegmann C, Anderljung M (2022) The Brussels effect and artificial intelligence: how EU regulation will impact the global AI market. arXiv preprint, arXiv:2208.12645
40. Sequera Buzarquis M (2026) Towards building AI governance from a Global South perspective. TEDIC, June
41. Center for Data Innovation (2026) EU digital rules as regulatory imperialism. Center for Data Innovation, Washington DC
42. Crum C (2025) Brussels effect or experimentalism? The AI Act in global context. Internet Policy Rev. https://doi.org/10.14763/2025.3.2032
43. Stanford HAI (2025) AI index report 2025, chap 8. Stanford Human-Centered Artificial Intelligence, Stanford
44. Manwani B, Rana A (2024) The right to (pry)-vacy. NYU J Intellect Prop Entertain Law 13
45. Institute of Legal Studies (2024) Data protection laws and the right to information. Institute of Legal Studies, New Delhi
46. BW Legal World (2023) Lok Sabha passes Women's Reservation Bill. BW Legal World, 20 September
47. Akashvani News (2023) Parliament passes Women's Reservation Bill. All India Radio News Services Division, 21 September
48. PRS Legislative Research (2023) The Government of National Capital Territory of Delhi (Amendment) Bill, 2023. PRS Legislative Research, New Delhi
49. LiveLaw (2023) Government of NCT of Delhi (Amendment) Act receives President's assent. LiveLaw, 12 August
50. The Wire (2023) After 6-hour debate and comments from 34 MPs, Delhi Services Bill becomes law. The Wire, 8 August
51. Mondaq (2026) India's Data Protection Board: the enforcer that isn't there yet. Mondaq, 17 April

52. Digital Personal Data Protection Rules (2025) Government of India, Ministry of Electronics and Information Technology, New Delhi, 13 November
53. Barrett T, Okolo CT, Biira B, Sherif E, Zhang A, Battle L (2025) African data ethics: a discursive framework for Black decolonial AI. In: Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, pp 334-349. https://doi.org/10.1145/3715275.3732023
54. Vijayakumar A (2024) AI ethics for the Global South: perspectives, practicalities, and India's role. RIS Discussion Paper 296. RIS, New Delhi
55. Ajuzieogu UC (2025) From aspiration to implementation: the African Union AI Continental Strategy and the development-governance paradox. Working paper, African Institute for Artificial Intelligence Policy, Abuja
56. CIPIT (2025) An in-depth analysis of the AU AI Continental Strategy and implications on AI governance in the continent. Centre for Intellectual Property and Information Technology Law, Strathmore University, Nairobi
57. Mdhlane T, Masisi L (2026) From epistemic injustice to epistemic justice: decolonising AI governance in the Global South. Presented at Development Studies Association Conference 2026
58. Pizarro Contreras R, Zhao Z, Zhang G (2025) The reproduction without alterity of AI: LatamGPT as a form of dissent and technological reappropriation. NanoEthics 19:1-20. https://doi.org/10.1007/s11569-025-00480-1
59. Sudhi Manjusha A (2025) Navigating India's AI governance framework: future directions in #AIforAll. SPRF Issue Brief 10/25. Social Policy Research Foundation
60. Transparency International India, Centre for Media Studies (n.d.) India corruption study: service delivery bribery. TI India, New Delhi
61. Down To Earth (2026) As told to parliament: Centre acknowledges AI's energy and water footprint while key facilities remain beyond EIA mandate. Down To Earth, 2 April
62. Rathore Y (2026) In India, a data centre can be built without any environmental assessment. BOOM, 27 February
63. Mongabay-India (2026) India bets on data centres even as water, energy-use concerns mount. Mongabay-India, 21 May
64. Fanon F (1963) The wretched of the earth (Farrington C, Trans). Grove Press, New York (Original work published 1961)
65. Mbembe A (2001) On the postcolony. University of California Press, Berkeley
66. Chatterjee P (1993) The nation and its fragments: colonial and postcolonial histories. Princeton University Press, Princeton.
67. Klitgaard R (1988) Controlling corruption. University of California Press, Berkeley
68. Rose-Ackerman S (1999) Corruption and government: causes, consequences, and reform. Cambridge University Press, Cambridge

69. Khan MH (2007) Corruption and governance. In: Dutta A, Mody A (eds) Institutional change and economic development. UNU Press, Tokyo
70. Scott JC (1998) Seeing like a state: how certain schemes to improve the human condition have failed. Yale University Press, New Haven
71. Lipsky M (1980) Street-level bureaucracy: dilemmas of the individual in public services. Russell Sage Foundation, New York
72. Eubanks V (2018) Automating inequality: how high-tech tools profile, police, and punish the poor. St Martin's Press, New York
73. Crawford K (2021) Atlas of AI: power, politics, and the planetary costs of artificial intelligence. Yale University Press, New Haven
74. Correa Lucero H, Martens C (2026) Colonial structures in AI: a Latin American decolonial literature review of structural implications for marginalised communities in the Global South. AI Soc 41:2511-2527. https://doi.org/10.1007/s00146-025-02547-9
75. Kwok JCY (2026) Data colonialism and indigenous languages in AI: a critical review of existing initiatives and their struggles with data sovereignty. AI Soc 41:6719-6729. https://doi.org/10.1007/s00146-026-03091-w

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationfinal.docx