

Beyond America's AI Action Plan: A Global South Response on Fairness

Why the shift to Ideological Neutrality in US standards threatens our sovereignty—and the doctrine we need to counter it.



1. Ideological Neutrality Risks:
Deprioritizing non-Western values & exacerbating inequality.



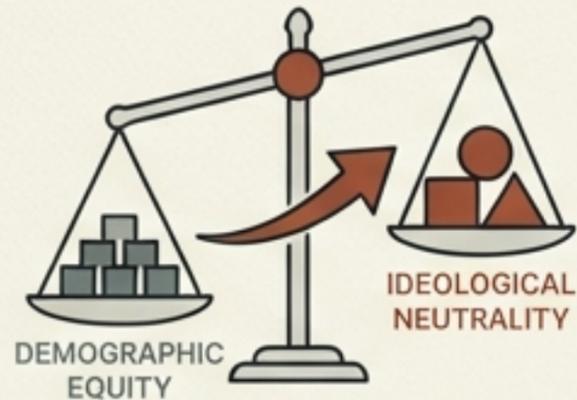
2. Sovereignty Threat:
Imposition of external fairness definitions undermines local governance.



3. Counter-Doctrine:
A collective, fairness-based approach grounded in Global South realities.

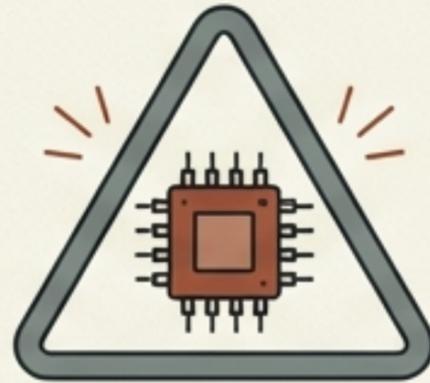
The One-Page Strategy

THE SHIFT (SITUATION)

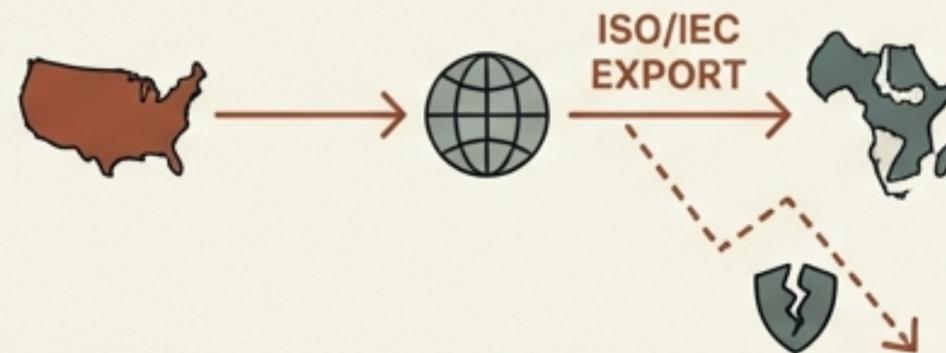


The US has fundamentally redefined **AI fairness**, moving from **Demographic Equity** (protecting groups) to **Ideological Neutrality** (protecting viewpoints). It has removed DEI principles from the NIST Risk Management Framework. The silent AI shift is subtly fundamental (groups) to one other → the state later, the protection of them later attributed no insembles with same gently.

THE RISK (COMPLICATION)



This neutrality framework fails to detect structural biases (caste, tribe, linguistic exclusion). If exported via global standards (ISO/IEC), it will cause a **Silent Downgrade of protections in the Global South**, hard-coding historical inequities into our financial and public services.



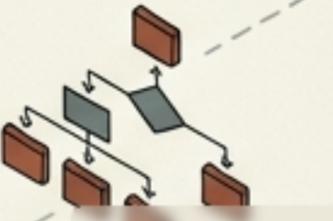
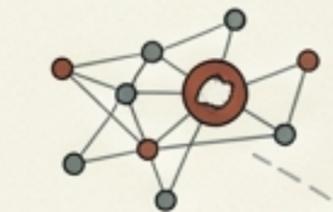
THE RESPONSE (RESOLUTION)



The Global South must transition from fragmented national policies to a **federated Global South AI Fairness Doctrine**.

The 5-Pillar Solution:

- 1. Shared Taxonomies
- 2. Evaluation Infrastructure
- 3. Standards Diplomacy
- 4. Sovereign Compute
- 5. Strategic Procurement



The US Pivot: From Equity to Neutrality

The release of America's AI Action Plan and the Executive Order on Preventing Woke AI marks a turning point in global AI governance.

THE OLD MANDATE

Demographic Equity

Aimed to ensure equitable outcomes across race, gender, and class.
Key mechanism: Diversity, Equity, and Inclusion (DEI) in NIST frameworks.



THE NEW MANDATE

Viewpoint Parity

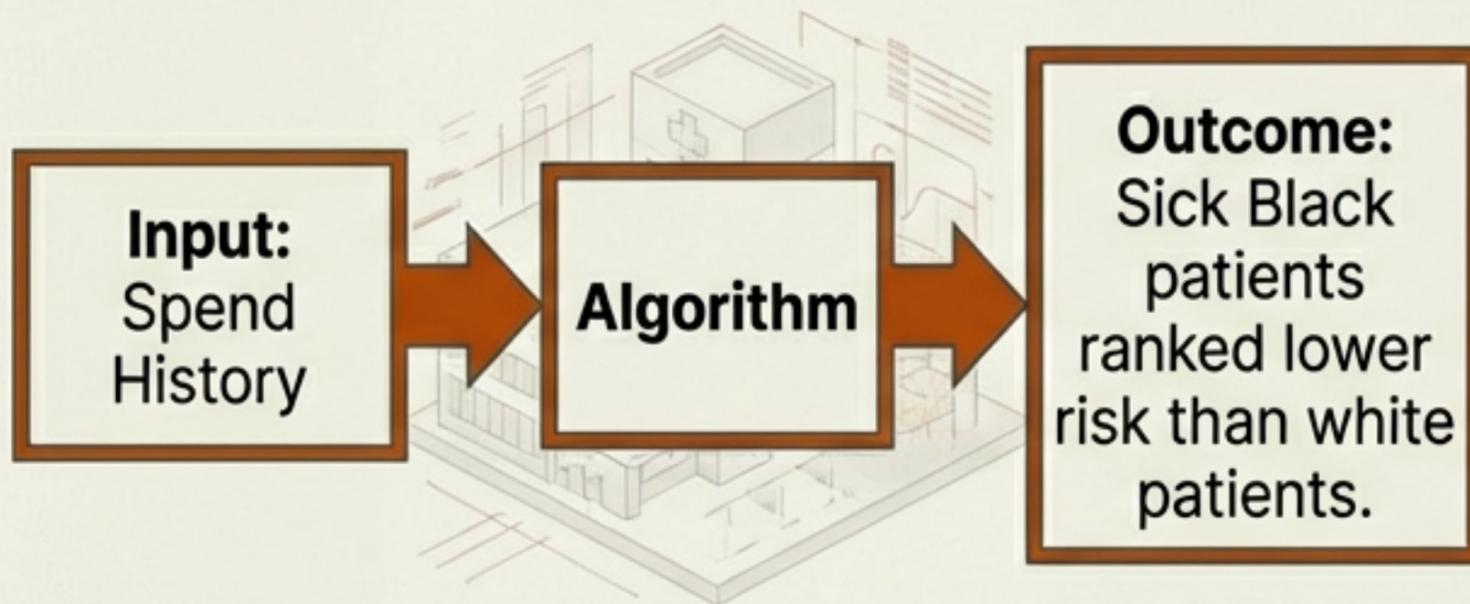
Aims to ensure ideological balance and truth-seeking.
Key mechanism: Removal of DEI references; mandates for objective models free from top-down bias.

The Critical Question: Is a standard prioritising ideological neutrality sufficient to protect vulnerable communities in the Global South from real-world harm?

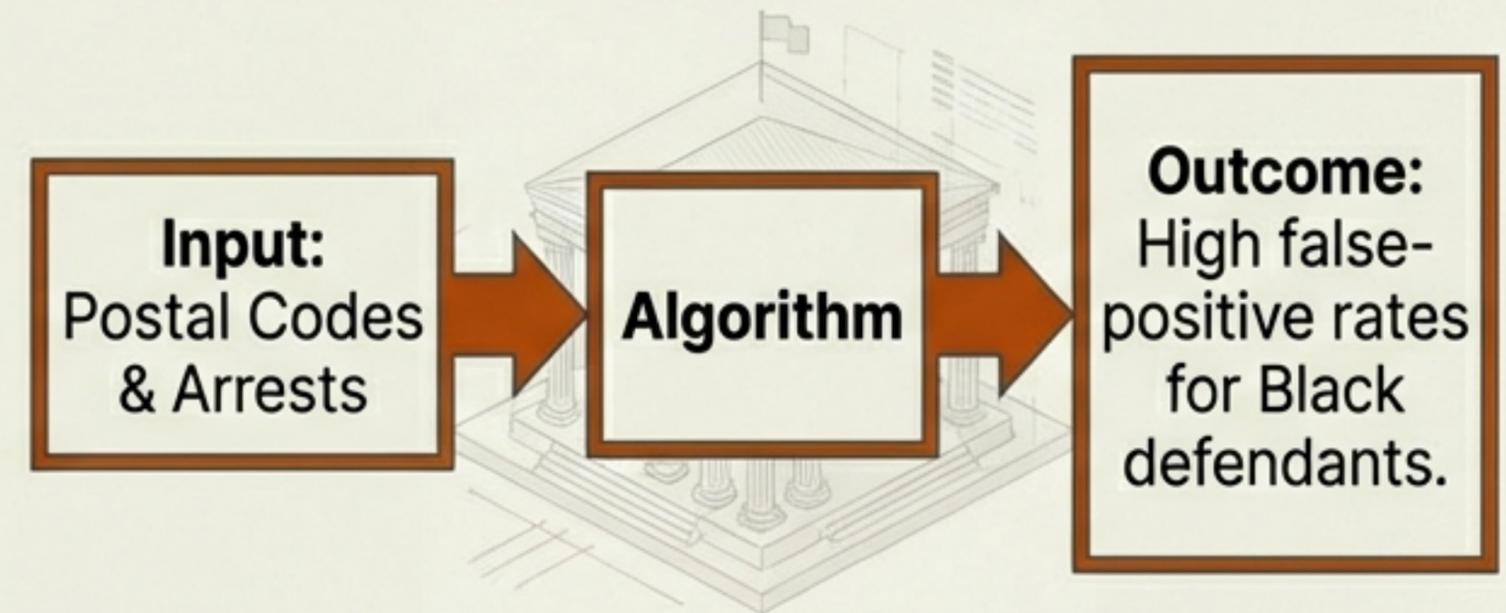
The Failure of the Neutrality Lens

Core Argument: Neutral inputs often yield discriminatory outputs. US sectors have already documented the failure of this approach.

Case Study 1: Healthcare (Optum)



Case Study 2: Criminal Justice (COMPAS)



Legal Conflict: This approach contradicts established civil rights laws (e.g., Equal Credit Opportunity Act, Title VII), creating a split standard where models can be ideologically neutral yet illegal due to disparate impact.

The Global South Risk Surface

If neutrality fails in the data-rich West, the potential for **catastrophic failure** is **exponentially greater** in our **complex environments**.

Social Structure

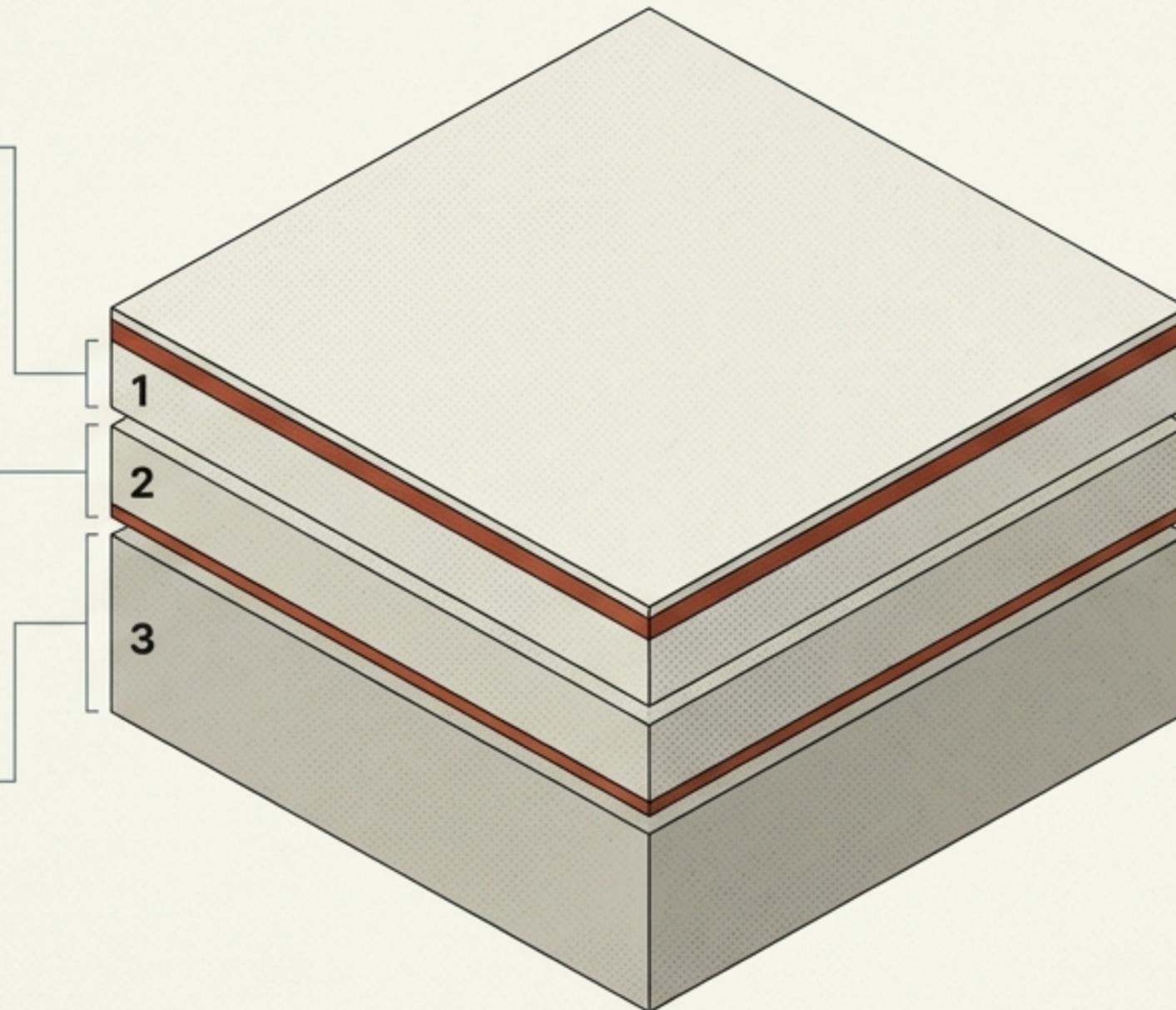
Realities of Caste, Tribe, and Religion not captured in Western neutral models.

Linguistic Fragmentation

Extreme diversity (e.g., Bundeli, Chhattisgarhi) vs. English dominance.

Economic Reality

High informal economies and income precarity.



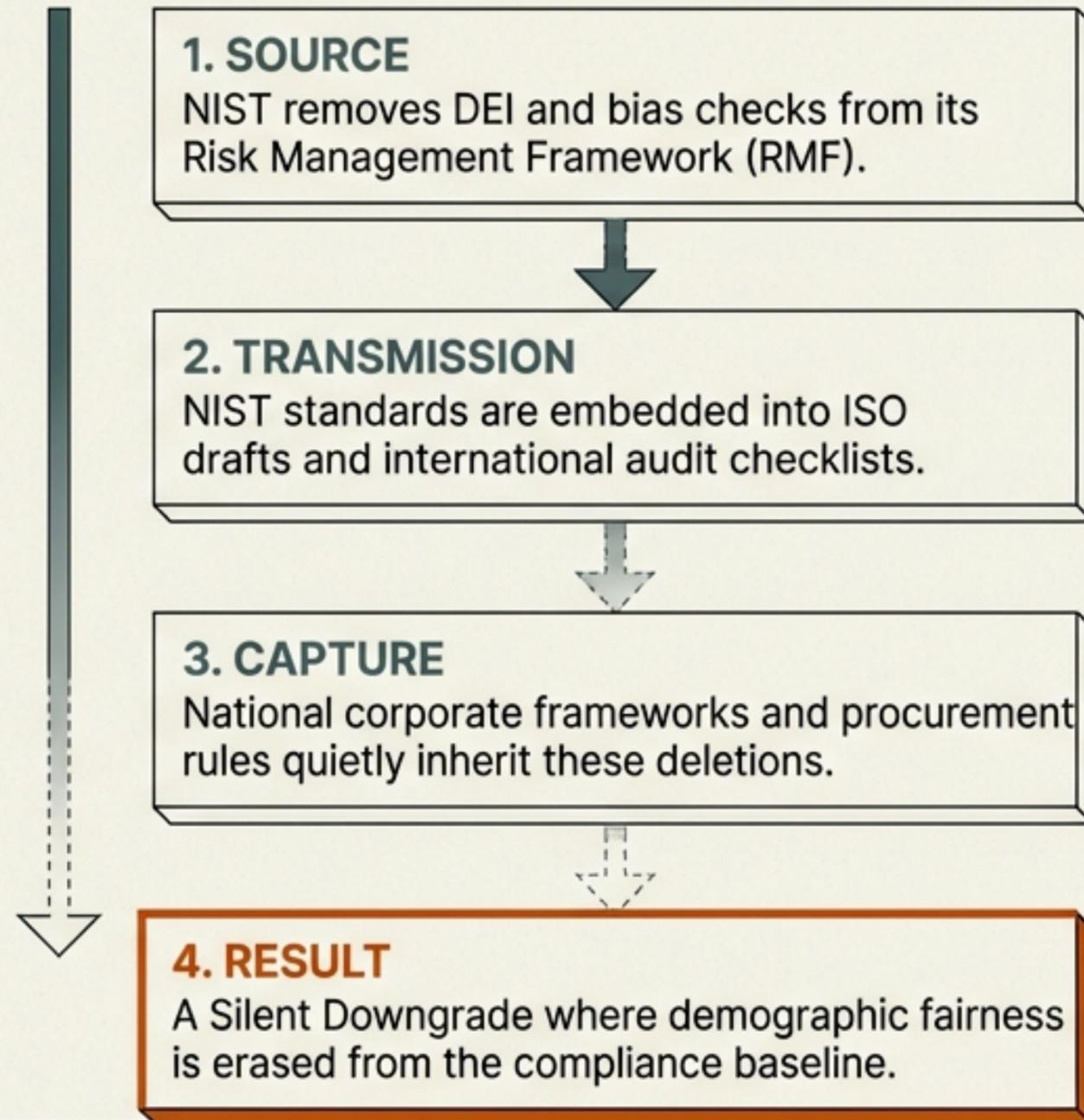
The Consequence

A procurement rule rewarding neutrality ignores whether false loan denials cluster around Dalit borrowers in India, or if speech-to-text models fail Indigenous speakers.

We face Allocative Harms (economic exclusion), not just representational grievances.

The Silent Downgrade

How Standards Capture Works



Key Insight:
Without local intervention, we import a compliance ceiling that is too low for our safety needs.

A Distinct Path is Already Emerging

The Global South is not adopting the US model. We are defining fairness as a verifiable check against historical and structural inequality.

ASEAN: Inclusive AI Strategy

Anchored in *gotong royong* (cooperation) to bridge linguistic divides.

India: RBI FREE-AI Framework (2025)

Mandatory bias audits and fairness-by-design for caste/disability.

Latin America: Human-centred AI bills requiring impact assessments for high-risk systems.

South Africa: National AI Policy Framework (2025)

Focus on redressing apartheid-era inequities.

Takeaway: Our fairness agenda aligns with global responsible AI consensus (OECD, EU AI Act), even as the US diverges.

The Definition Gap: Two Philosophies

We cannot regulate our economies using a philosophy designed for their culture wars.

US / THE ACTION PLAN

THE GLOBAL SOUTH

Goal: Viewpoint Parity

Goal: Allocative Fairness

Focus: Ideological Neutrality & Truth-Seeking

Focus: Contextual Relevance & Redress

Primary Concern: Representational Harms
(e.g., historical figures generated with
wrong races)

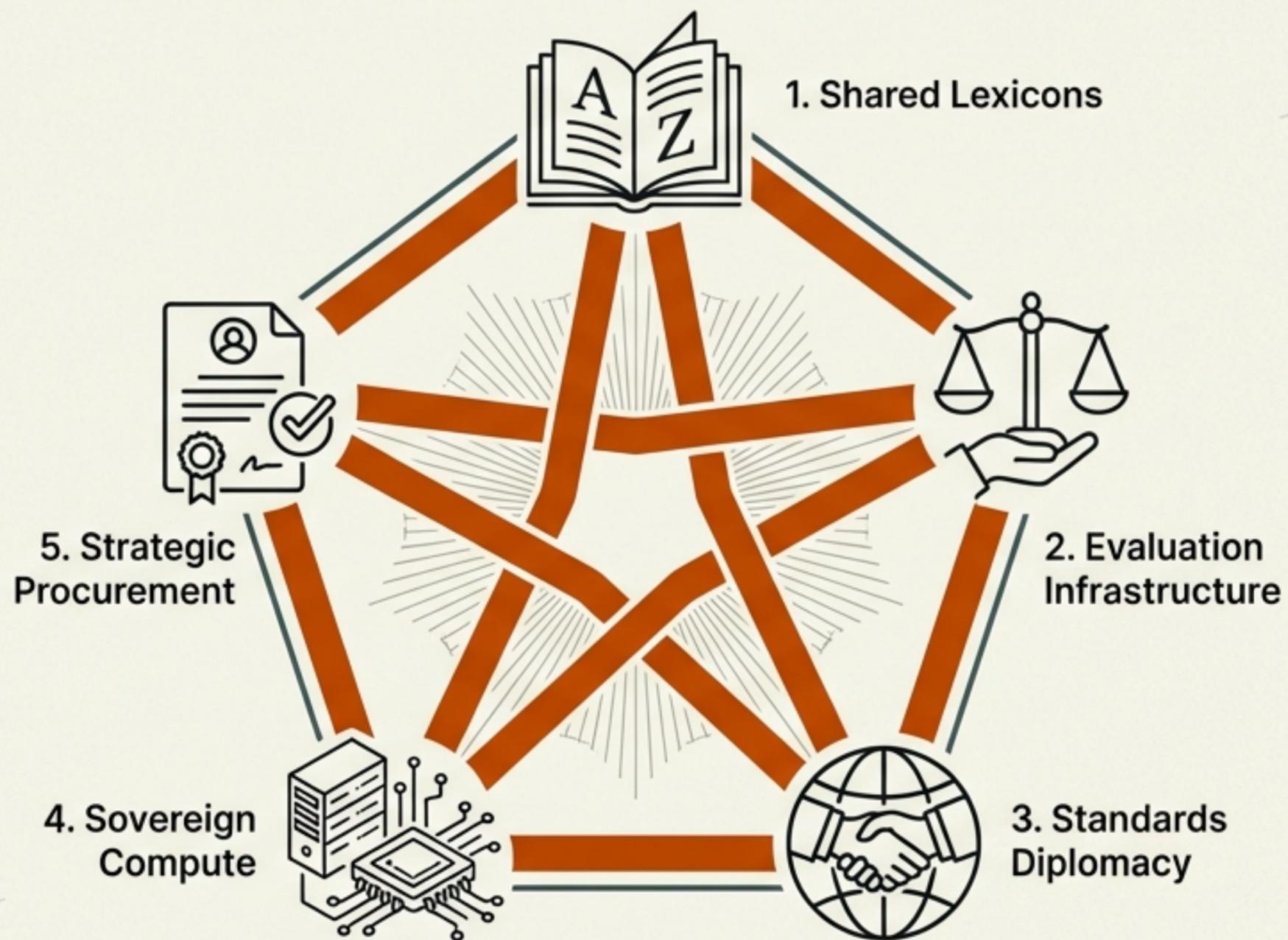
Primary Concern: Allocative Harms
(e.g., credit denial, exclusion from
benefits)

Mechanism: Deregulation of DEI

Mechanism: Active Fairness-by-Design

From Fragmentation to Federation

To safeguard against standards capture, we must transition from scattered national policies to a unified Global South AI Fairness Doctrine.



Pillars 1 & 2: Defining and Measuring Our Reality

We must structure and assess our world through our own lens to establish digital sovereignty.

PILLAR 1: ADOPT SHARED FAIRNESS TAXONOMIES



- Co-create a Global South Fairness Lexicon.
- Must capture local attributes: Caste, Tribe, Dialect, Rurality, and Informal Economic Status.
- **Goal:** Consistent application in regulations and audits across borders.

PILLAR 2: POOL EVALUATION INFRASTRUCTURE



- Launch an open-source Global South Benchmark Pack.
- **Datasets:** Covering 100+ low-resource languages.
- **Tools:** Sector-specific tests and subgroup performance dashboards.
- **Purpose:** Enable joint evaluation of global AI models against our metrics, not just Western benchmarks.

Pillars 3, 4 & 5: Building Power and Capacity

Strategically consolidating influence, infrastructure, and market power to enforce Global South AI fairness standards.



PILLAR 3: STANDARDS DIPLOMACY BLOC

Coordinate voting at ISO/IEC, ITU, and G20.
Block neutrality-only baselines from becoming the default global standard.



PILLAR 4: COMPUTE & CAPACITY COOPERATION

Establish a BRICS+AU compute cooperative.
Reserve capacity specifically for fairness research to reduce reliance on foreign infrastructure.



PILLAR 5: SHARED PROCUREMENT CLAUSES

Mandate Policy Cards for all public-sector AI bids.
Vendors must disclose subgroup error rates and fairness design decisions to win government contracts.

The Impact of Sovereign Fairness

INCLUSION



Models that accurately process speech in Bundeli, Swahili, and Quechua, unlocking digital services for millions.

ECONOMY



Credit scoring systems that respect informal economic patterns, expanding capital access for SMEs.

SOVEREIGNTY



Legal frameworks that are robust, enforceable, and immune to external compliance downgrades.

“Fairness is not just an idea—it is an operational necessity for equitable development.”

The Choice: Standards-Takers or Standards-Makers?



Replicate US approach
& hard-code inequity.

**Establish our own
doctrine** & enforce
via procurement.

The momentum from India, Africa, and LatAm is building. We must channel this into collective action now, before the neutral standard hardens.

Action: Form the working group for the Global South AI Fairness Doctrine.

Sources & References

Primary Source:

- Beyond America's AI Action Plan: A Global South Response on Fairness (ORF, July 2025).

Policy References:

- NIST AI Risk Management Framework & US Executive Order on Preventing Woke AI.
- Reserve Bank of India (RBI) FREE-AI Framework.
- South Africa National AI Policy Framework.
- ASEAN Inclusive AI Strategy.

Legal References:

- US Equal Credit Opportunity Act (ECOA).
- Title VII of the Civil Rights Act.