Intersect, Vol 19, No 2 (2026)

Policed by Code: AI, Gender and Justice in the Global South

Piyush Mishra
Independent Researcher, Founder – Global Dream Connect

Algorithmic policing systems in the Global South do not simply reproduce existing inequalities, they intensify and reorganise them. Evidence from three urban contexts, Hyderabad (India), Lagos (Nigeria), and Rio de Janeiro (Brazil), shows that AI-driven surveillance and predictive policing tools disproportionately target women situated at the intersections of gender, race, caste, and class. The central problem is not technical failure. It is the structural embedding of structural inequality within systems presented as neutral, efficient, and objective.

Grounded in intersectional feminist theory (Crenshaw, 1989) and postcolonial critiques of technology (Benjamin, 2019; Mbembe, 2003; Quijano, 2000), this paper develops the Algorithmic Harm Cascade, a five-stage model that traces how historical data inequalities evolve into contemporary forms of exclusion. The cascade follows a progression from discriminatory data inheritance to proxy encoding, deployment amplification, and governance absence, culminating in structural exclusion reinforced through behavioural adaptation. This framework distinguishes between performance bias, uneven technical outcomes, and sociotechnical bias, where systems function as designed but reproduce injustice. The latter emerges as the dominant form of harm across all three cases.

Analysis of over 250 publicly available sources, including peer-reviewed research, procurement records, civil society reports, and investigative journalism, reveals three consistent pathways through which inequality enters algorithmic systems: the incorporation of biased enforcement records into training data, the encoding of class and gender through proxy variables, and the construction of bodily visibility as a basis for suspicion.

Building on these findings, the paper advances a three-part governance framework, Power-Aware AI Systems, Memory Justice, and the Right to Algorithmic Refusal, grounded in empirical realities rather than abstract ethical commitments. Across all cases, governance capacity remains limited, with no city exceeding a score of 2.5 out of 5 on key

regulatory dimensions, underscoring a widening gap between technological deployment and institutional accountability.

Keywords: artificial intelligence policing, gender bias, intersectionality, surveillance, postcolonial technology, algorithmic justice, sociotechnical bias, governance, Global South

Introduction

In Hyderabad, images of residents captured from public spaces and social media are processed daily by a facial recognition infrastructure linked to more than 600,000 CCTV cameras. Investigations by Amnesty International (2021) reveal that this system generates false identification alerts at an extremely high rate, with consequences falling disproportionately on women from Dalit and Muslim communities. In Lagos, predictive policing systems informed by algorithmically assigned risk scores concentrate patrol activity in informal settlements, where field documentation records 89 incidents in 2022 in which women were stopped, questioned, or pressured to provide biometric data without meaningful consent (Uduo & Obaji, 2024). In Rio de Janeiro, 90.5% of the individuals arrested through facial recognition between March and October 2019 for whom racial data was recorded were Black or mixed-race, across five states, drawn from the most comprehensive systematic civil society monitoring effort of its kind available for that period, which documented 151 total arrests (Nunes, 2019). These are not isolated technical failures. They are symptoms of a structural condition this study seeks to diagnose and address.

The study's central research question is: how do AI-driven policing systems reinforce gender-based harm against women in the Global South? The question carries stakes that extend well beyond criminal justice. The McKinsey Global Institute (2019) documented that automation displaces women at substantially higher rates than men in lower-income labour markets, with informal sector workers bearing the sharpest adjustment costs. When algorithmic policing compounds this vulnerability by constraining women's access to public space, movement, and commercial activity, the resulting economic and social costs are significant and fall most heavily on those least positioned to absorb them (McKinsey & Company, 2020). Despite this, AI governance scholarship has largely focused on system accuracy rather than downstream distributional consequences.

Three reasons make women a necessary analytical category for this inquiry. First, commercial facial analysis systems audited under controlled conditions misclassify darker-skinned women at rates of up to 34.7%, compared to a maximum error rate of only 0.8% for lighter-skinned men, a disparity that cannot be explained through single-axis analyses of race or gender alone (Buolamwini & Gebru, 2018). Second, intersectional feminist theory demonstrates that single-axis frameworks are structurally incapable of capturing harms that emerge at the convergence of gender, race, caste, and class (Crenshaw, 1989). Third, in patriarchal societies shaped by colonial governance legacies, AI policing tools are deployed as instruments of systematic surveillance and control over women's public presence and mobility (UN Women, 2022; Smith & Rustagi, 2020).

The paper makes four contributions. The Algorithmic Harm Cascade, a five-stage model tracing the pathway from colonial data inheritance to structural exclusion, constitutes the central theoretical contribution. Beyond the model, this study extends Crenshaw's (1989) intersectionality framework and Benjamin's (2019) New Jim Code across three distinct postcolonial settings, showing how discrimination, reframed through technical language, operates across varied racialised hierarchies. It also provides the first comparative three-city analysis of AI policing governance gaps using a structured readiness framework, and proposes an implementable three-pillar governance framework derived from field evidence rather than abstract ethical principles.

Literature Review

Technical Promises vs. Realities

AI-based policing tools were adopted across numerous jurisdictions on the assumption that algorithmic systems would enhance the efficiency and objectivity of law enforcement. Predictive policing models were designed to anticipate crime patterns, while facial recognition technologies were introduced to reduce reliance on subjective human identification (Perry et al., 2013; Brantingham et al., 2018). Perry et al. (2013, p. 1) define predictive policing as the application of analytical techniques to identify likely targets for police intervention, premised on the claim that forecasting is objective, scientific, reproducible, and free from individual bias and error. Many jurisdictions in the Global South moved rapidly to integrate these technologies, often in advance of regulatory frameworks capable of governing their use, frequently justified through procurement narratives framing AI adoption as essential to urban modernisation and public security competitiveness.

Subsequent empirical research has substantially complicated these expectations. Buolamwini and Gebru (2018) audited three commercial gender classification systems developed by Microsoft, IBM, and Face++. Under controlled conditions with professional-quality photography, IBM's system misclassified darker-skinned women 34.7% of the time while misclassifying lighter-skinned men at a maximum rate of only 0.8%, a gap of up to 33.9 percentage points. Darker-skinned females, who constituted only 21.3% of the benchmark population, accounted for between 61% and 72.4% of all classification errors across the three systems. The benchmark datasets on which these systems were trained were themselves demographically skewed: the IJB-A benchmark consisted of 79.6%

lighter-skinned subjects with only 4.4% darker-skinned females, and the Adience dataset consisted of 86.2% lighter-skinned subjects. Buolamwini and Gebru audited gender classification systems rather than identity-matching facial recognition, a distinction that matters. The inferential step, that systems performing poorly on gender classification under optimal conditions perform similarly or worse on individual identification under real-world conditions, is supported by subsequent field evidence. Section 5 documents these downstream effects.

A Sociotechnical Definition of Bias

Understanding these disparities requires moving beyond a purely technical definition of bias. Smith and Rustagi (2020) distinguish two analytically distinct but frequently co-occurring forms. Performance bias refers to inaccurate predictions or classification errors occurring at differential rates across population subgroups. Sociotechnical bias describes discriminatory outputs that may be statistically accurate but reinforce existing structural inequalities, even when systems perform as designed. This distinction is critical for AI policing in the Global South: a predictive system that accurately identifies a neighbourhood as having a high historical arrest rate performs correctly in the technical sense while producing a discriminatory outcome in the structural sense if that arrest rate reflects decades of concentrated enforcement rather than the underlying distribution of criminal behaviour.

Bias enters AI systems at multiple stages. At the data generation stage, training datasets overrepresent populations whose activities generate formal administrative records (Uduo & Obaji, 2024; Lee et al., 2019). Data labelling introduces a second layer of distortion: annotation processes reflect the cultural assumptions of those conducting the labelling, overwhelmingly drawn from high-income Western contexts (Holstein et al., 2019). Algorithm design further compounds these disparities, as standard loss functions minimise average error, weighting improvements for majority populations more heavily than equivalent improvements for minority subgroups, a structural feature of optimisation processes identified by Mehrabi et al. (2021). Deployment introduces an additional layer of misalignment, as systems designed and tested in one institutional and historical context are applied to radically different populations without adaptation. At the interpretation stage, automation bias, the tendency of human decision-makers to over-rely on algorithmic recommendations, means that nominal human oversight frequently functions as a rubber stamp (Holstein et al., 2019; Mehrabi et al., 2021).

Proxy variables intensify these dynamics. Postal codes, surnames, time of day, and patterns of movement function as indirect signals of race, class, and gender within systems that formally disclaim discriminatory targeting (Holstein et al., 2019). Lee et al. (2019) demonstrate that an algorithm can be formally blind to race while producing racially differentiated outputs when its inputs correlate with race in the training population. Noble (2018) extends this argument to commercial information systems, showing that optimisation for signals reflecting existing social hierarchies reproduces and amplifies those hierarchies without explicit discriminatory intent. Feedback loop bias compounds initial disparities over time: predictive policing directed at a neighbourhood generates arrests there; arrests are recorded as crime incidents, which confirm the prediction, which directs additional patrols in subsequent iterations (Mehrabi et al., 2021).

Intersectional and Postcolonial Frameworks

Two theoretical traditions converge to explain the patterns the case evidence reveals. Intersectional feminist theory, as developed by Crenshaw (1989), provides analytical tools to examine how gender, race, caste, and class interact to produce forms of vulnerability that no single axis captures. Crenshaw's traffic intersection analogy captures the analytical failure of single-axis audits: a system can produce acceptable aggregate performance for women overall and for darker-skinned individuals overall while failing at their intersection, because aggregation obscures the subgroup most affected. While developed in the US legal context, this framework requires extension in postcolonial settings where axes of disadvantage include caste, colonial administrative categories, and citizenship status, an extension this study explicitly undertakes.

Postcolonial technology critique situates these dynamics within a longer historical arc. Benjamin (2019) argues that contemporary AI systems inherit and reproduce hierarchies established through earlier systems of governance, what she terms the New Jim Code. Mbembe's (2003) concept of necropolitics, the power to determine which populations are exposed to punitive or lethal state power without accountability, provides a lens for understanding AI policing in over-surveilled communities. Quijano's (2000) coloniality of power framework highlights the persistence of racialised hierarchies established under colonial rule through contemporary institutions. Spivak's (1988) question of subaltern legibility finds a contemporary application in AI governance: communities subjected to algorithmic policing are rendered structurally illegible within the technical and institutional systems used to evaluate those same technologies.

Zuboff's (2019) framework, however, positions the state as a secondary actor relative to corporate surveillance capitalism, a configuration that does not hold in Hyderabad, Lagos, or Rio de Janeiro. The state is the primary surveillance actor. Corporate vendors are supply-side actors. This distinction requires adapting Zuboff's framework to postcolonial contexts where state power, rather than corporate data extraction, drives the architecture of surveillance. Eubanks (2018) complements this analysis through the concept of the poverty penalty: the same algorithmic error constitutes a minor inconvenience for formally documented individuals and a livelihood-threatening disruption for those operating in informal or precarious conditions. At the organisational level, McKinsey & Company (2020) documented the measurable economic costs of systemic exclusion, showing that institutions that fail to account for inequality incur compounding human and financial risks.

These patterns are consistent with findings documented by UN Women (2022), which show that politically marginalised women experience disproportionate surveillance exposure, with consequences for participation in public and economic life.

The Gap This Study Fills

Existing scholarship on algorithmic bias has developed along largely separate tracks. Technical research documents performance disparities without engaging with the institutional and historical conditions that generate training data. Sociological scholarship identifies structural dynamics but has given limited attention to how these operate across intersecting axes of gender, caste, race, and class in Global South contexts. Economic analyses of AI's distributional impacts (McKinsey Global Institute, 2019) have focused primarily on labour market displacement rather than the compounding effects of AI policing on already marginalised women's economic participation. Comparative governance analysis of AI policing systems across multiple Global South cities remains largely absent, with most accountability research focused on single jurisdictions or high-income country contexts.

This study addresses that gap by treating intersectional and postcolonial frameworks as primary analytical tools rather than supplementary perspectives, and by moving across three cities rather than one, enabling identification of mechanisms that single-case analyses cannot establish. The Algorithmic Harm Cascade introduced in Section 4 provides the integrative model that this fragmented literature has lacked.

Methodology

Research Design

This study employs a qualitative comparative case study design (Yin, 2014) to investigate how AI-based policing systems reinforce gender-based harm against women in Hyderabad, Lagos, and Rio de Janeiro. A qualitative approach is appropriate because the harms under investigation, misidentification, surveillance anxiety, constrained mobility, and reputational damage, are not fully captured by quantitative performance metrics, and because close examination of institutional and social conditions is necessary to understand how algorithmic systems interact with local histories of policing, colonial governance, and social stratification. The comparative design enables analysis of how differences across technological configurations and institutional environments interact with local histories of inequality, distinguishing context-specific dynamics from broader structural conditions.

The study is explicitly interdisciplinary, drawing on feminist theory, postcolonial studies, sociology, technology ethics, and policy analysis.

Each discipline provides a distinct analytical lens: feminist and intersectional theory identifies how overlapping social identities shape exposure to algorithmic harm; postcolonial critique situates AI deployment within historical legacies of colonial governance; sociology informs the analysis of urban inequality and institutional policing practices; and policy analysis illuminates governance gaps and accountability failures.

Case Selection

Three cities were selected through purposive theoretical sampling (Yin, 2014) using a maximum-variation strategy to represent distinct governance structures, technological configurations, and social contexts while enabling cross-case comparison. This ensures that convergent findings cannot be attributed to similarity in system design, institutional environment, or regional context.

Hyderabad was selected as a site of mass biometric surveillance infrastructure built through a partnership between the Telangana state government and private technology vendors, including AI firms and camera hardware suppliers operating in the absence of human rights due diligence frameworks, where the primary source of bias is domestic data infrastructure augmented by foreign technology (Amnesty International, 2021). Lagos represents a rapidly digitising urban context in which predictive policing systems procured from private technology vendors reproduce colonial-era policing logics in informal settlements (Uduo & Obaji, 2024). Rio de Janeiro represents a deployment of facial recognition systems developed in partnership with Fujitsu and operated through the Instituto de Segurança Pública, combining imported system architecture with local data biases concentrated through the city's racialised spatial history (Ramiro & Cruz, 2023; Salem, 2024).

Data Sources and Analysis

The analysis draws exclusively on publicly available secondary sources identified through database searches of Google Scholar, Scopus, and JSTOR, using search terms including "facial recognition bias Global South," "predictive policing gender," "AI surveillance India/Nigeria/Brazil," and "algorithmic discrimination intersectional." A total of over 250 sources were reviewed across five categories: peer-reviewed academic studies; institutional reports from Amnesty International, the Brookings Institution, UN Women, and McKinsey Global Institute; government policy documents and procurement records; investigative journalism; and published field audit data. Each finding presented in Section 5 is supported by evidence from at least two independent source types.

Governance capacity across the three cities is assessed using a structured readiness framework evaluated across five dimensions, presented as a comparative radar assessment in Section 6.

Thematic analysis followed the six-phase approach described by Braun and Clarke (2006): familiarisation; initial coding; theme searching; theme reviewing; theme defining; and writing. Deductive codes were drawn from feminist, postcolonial, and technology ethics literature. Inductive codes captured emergent themes including surveillance anxiety and vendor opacity. Codes were aggregated into three intersectional harm themes: Algorithmic and Technical Harm; Social and Psychological Harm; and Legal and Structural Harm. Cross-case comparison applied Mill's method of agreement alongside elements of Qualitative Comparative Analysis (QCA).

The Algorithmic Harm Cascade: An Integrative Model

The Algorithmic Harm Cascade, the integrative model proposed by this study, synthesises Mehrabi et al. (2021), Benjamin (2019), Smith and Rustagi (2020), Crenshaw (1989), Eubanks (2018), and the case evidence. Existing frameworks in algorithmic fairness literature, whether technical (Mehrabi et al., 2021), governance-focused (Lee et al., 2019), or sociotechnical (Smith & Rustagi, 2020), tend to isolate individual stages of AI system development. None traces the full pathway from historically shaped data inputs to long-term structural exclusion as a continuous process.

The cascade model conceptualises bias as emerging through a sequence of interconnected stages that reinforce one another over time. It begins with data inheritance, where historically biased administrative and policing records form the foundation of training datasets. These biases are then embedded through proxy encoding, as variables such as location, time, and movement act as indirect signals for protected characteristics. Deployment into unequal institutional contexts then amplifies these disparities. In the absence of effective governance mechanisms, including audit processes and accessible avenues for redress, these disparities persist and accumulate. Over time, affected populations adapt behaviourally, often through reduced mobility or withdrawal from public and economic life, resulting in structural exclusion, the condition Eubanks (2018) terms the poverty penalty, where algorithmic error imposes livelihood-level costs on those with no institutional recourse.

The model advances existing scholarship in three ways. It foregrounds compounding as the central mechanism, arguing that harm at each stage is not simply additive but intensifies as it moves through subsequent stages. It incorporates an economic dimension often underexamined in algorithmic fairness literature: McKinsey Global Institute (2019) documented that behavioural adaptation and labour market withdrawal are more pronounced among women in informal sector employment, the population most affected at later stages of the cascade. It also reframes governance intervention as requiring simultaneous engagement across multiple stages, rather than isolated technical or regulatory fixes, a logic that underpins the three-pillar governance framework developed in Section 6.3.

Limitations

The analysis relies entirely on secondary data, constraining access to first-hand narratives from affected communities, particularly in non-English-speaking contexts where documentation may be less accessible. The focus on three cities provides analytical depth but limits generalisability. The opacity of vendor training sets, model specifications, and deployment telemetry limits assessment of the internal workings of the systems examined. The harm intensity scores in Figure 4 are composite estimates derived from field documentation rather than independently audited measurements. Cross-source triangulation drawing on at least two independent source types for each finding was employed to mitigate these challenges.

Case Study Analysis

Hyderabad, India: The Prototype of Techno-Authoritarianism

Context and Deployment

Hyderabad's surveillance architecture emerged through a partnership between the Telangana state government and private technology vendors operating in the absence of human rights due diligence frameworks (Amnesty International, 2021). The Hawk Eye Command and Control Centre links more than 600,000 CCTV cameras into a real-time monitoring network.

Scale is not a neutral property. It is an amplifier.

Despite the presence of foreign-manufactured hardware, the primary source of algorithmic bias is domestic: neural networks are trained on locally maintained habitual offender registers and digitised police records whose origins trace to colonial administrative practices under British governance.

India has spent approximately 9.6 billion rupees on facial recognition technology (Amnesty International, 2021), and there is currently no safeguarding legislation protecting citizen privacy or regulating remote biometric surveillance. AI does not introduce this bias, it scales a discriminatory logic already embedded in decades of uneven enforcement. This reflects Benjamin's (2019) concept of the New Jim Code: systems that function as designed within unjust social architectures.

Racial and Caste-Based Failure

The Hawk Eye system was trained predominantly on records generated through the intensive over-policing of Dalit and Muslim communities. Amnesty International's (2021) investigation documents a pattern of wrongful stops, intrusive identity checks, and the concentration of alert-driven policing within these populations.

The pattern is not incidental. It is structural.

Within the same national surveillance ecosystem, comparable facial recognition systems have demonstrated extremely low accuracy rates under similarly unregulated conditions, reinforcing concerns about systemic misidentification. Regardless of precise performance variation across jurisdictions, the underlying pattern remains consistent: biased historical data produces biased algorithmic outputs.

The coloniality of power described by Quijano (2000) is operationalised here with computational precision. Administrative categories established under colonial governance persist in digitised enforcement records; those records form training datasets; and those datasets reproduce the geography of colonial policing with the authority of data.

Gender as a Compounding Axis

Understanding the gendered dimension of algorithmic failure requires first establishing how Dalit and Muslim men are positioned within the system, and then demonstrating how women's experiences diverge structurally rather than incrementally, following Crenshaw's (1989) methodological prescription.

Dalit men are over-represented in criminal databases, habitual offender registers, and digitised arrest records. This over-representation reflects decades of caste-concentrated policing and produces a form of legibility. Because the Hawk Eye system is trained on these records, Dalit men are more likely to be accurately recognised, not because the system operates fairly, but because it has been optimised on data generated through their over-policing. The harm operates through intensified targeting: they are correctly identified and therefore repeatedly flagged as suspect. It is recognition without fairness.

Women from the same communities encounter a fundamentally different configuration of failure. The datasets used to train facial recognition systems contain relatively few images of women in traditional or religious dress, veiled women, darker-skinned women in everyday working contexts, or women from lower-caste communities engaged in routine public activity. This reflects the structure of policing itself: women are arrested far less frequently, and therefore appear less often in the records that become training data. They are missing from the dataset that defines normality.

Underrepresentation does not produce invisibility. It produces misrecognition. Women inhabit neighbourhoods classified as high-risk based on male arrest records, and their presence within those spaces triggers suspicion despite unreliable identification. They are simultaneously over-flagged and inaccurately recognised, rendered, in Crenshaw's (1989) terms, both hypervisible to policing systems and invisible within the data that trains them.

A third mechanism emerges through the interaction between algorithmic surveillance and patriarchal social norms. For Muslim women, public presence is already subject to heightened communal scrutiny;

algorithmic monitoring adds institutional authority to existing constraints on mobility. For Dalit women, participation in political or religious activity has historically triggered caste-based policing; AI-enabled identification introduces continuous monitoring independent of criminal behaviour (Amnesty International, 2021; UN Women, 2022). The system does not only inherit bias, it amplifies social control.

The convergence reveals a form of harm that neither caste-only nor gender-only analysis can capture. A caste-only framework explains over-targeting but assumes male experience as the norm. A gender-only framework explains underrepresentation but cannot account for why harm is concentrated among lower-caste women. Intersectional analysis shows that women who are underrepresented in training data yet overrepresented in surveilled spaces are uniquely exposed to misidentification combined with persistent scrutiny, a configuration neither axis predicts independently.

The material consequences extend beyond policing encounters. Documentation indicates that domestic workers have altered routes, modified dress, and left work earlier to avoid surveillance exposure (Amnesty International, 2021). These responses reflect a distinct form of constraint: the algorithm drives women out of public space; their withdrawal is recorded neither as a policing incident nor as an economic harm. It simply disappears.

Governance Gap

India lacks a comprehensive national framework regulating facial recognition in public policing. There is no requirement for informed consent, independent auditing, or public disclosure of system operations. For many of the women most affected — who may be unaware their biometric data is being captured — formal avenues for redress remain structurally inaccessible. This reflects the broader institutional inability to register subaltern experience within formal systems of accountability, as theorised by Spivak (1988).

Lagos, Nigeria: Predictive Policing and the Legacy of Colonial Control

Context and Deployment

In Lagos, AI policing has centred on predictive analytics systems procured through contracts with private technology vendors. The technical architecture remains opaque; detailed documentation on training data, modelling assumptions, and validation procedures has not been made publicly available (Uduo & Obaji, 2024). This opacity — referred to here as vendor opacity — limits both public scrutiny and institutional accountability.

These systems rely on historical arrest records and police incident reports to generate spatial forecasts for targeted enforcement. The data reflects a colonial and postcolonial trajectory in which policing was disproportionately concentrated in informal settlements designated as zones of administrative suspicion. Colonial spatial classifications persist in administrative datasets and are reactivated through algorithmic prediction (Quijano, 2000).

Class and Gender as Compounding Axes

In Lagos, the primary axis of distortion operates through class inequality. Training data disproportionately reflects populations engaged in the formal economy, while large segments of the informal sector remain underrepresented.

Women working within the informal economy encounter a distinct form of vulnerability. Predictive systems rely on proxy indicators such as location, time of day, and movement patterns. These variables encode social norms as risk signals. A woman selling goods at a market in the evening may be flagged not for criminal activity, but because her presence deviates from patterns treated as normative within the training data, a mechanism consistent with proxy-based bias documented by Holstein et al. (2019) and Lee et al. (2019).

This encoding of social norms as algorithmic risk signals reflects Noble's (2018) account of optimisation for existing social hierarchies, discrimination reproduced through technical processes rather than explicit intent.

The absence of gender-disaggregated data in Uduo and Obaji's (2024) study constitutes an intersectional gap in itself. Civil society reporting documented 89 incidents in 2022 in which women were stopped, questioned, or pressured to provide biometric data without meaningful consent (Uduo & Obaji, 2024).

Documentation further indicates that women working in flagged areas have withdrawn from night-shift employment, altered transport routes, and modified working patterns to reduce exposure to police scrutiny (Uduo & Obaji, 2024). These behavioural adjustments are not captured within system performance metrics, yet they represent measurable economic harm concentrated among those least able to absorb disruption.

Digital Colonialism and Governance Responses

The Lagos case illustrates digital colonialism: technologically advanced governance tools deployed in contexts where regulatory oversight remains limited. Following sustained civil society pressure, authorities have incorporated limited transparency provisions into some procurement agreements (Uduo & Obaji, 2024). These measures remain narrow, lacking independent audits, enforceable data retention limits, or community notification requirements, but they establish a precedent for intervention through collective advocacy rather than individual complaint.

Rio de Janeiro, Brazil: Recognition, Race, and the Criminalisation of Visibility

Context and Deployment

Since 2019, authorities in Rio de Janeiro have deployed facial recognition systems through public-private partnerships, operating within regulatory grey zones created by exemptions in Brazil's General Data Protection Law for public security (Ramiro & Cruz, 2023).

Nunes (2019) documented 151 arrests across five states between March and October 2019 using facial recognition systems. Of the cases in which racial data was recorded, 90.5% involved Black or mixed-race individuals, a pattern consistent across jurisdictions and reinforced by subsequent monitoring.

Race as the Primary Axis of Failure

Rio's surveillance infrastructure operates within a spatial order structured by the favela/asfalto divide, in which favela spaces are coded as Black, poor, and subject to militarised policing (Salem, 2024). Policing practices are shaped by institutional norms that position Black men as presumptive suspects within these spaces, producing a system in which algorithmic identification reinforces pre-existing racial targeting.

Mbembe's (2003) concept of necropolitics, the power to expose populations to punitive state action without legal accountability, describes the structural logic that Salem's ethnography makes visible.

Gender as a Compounding Axis

Afro-Brazilian women encounter a configuration of harm that cannot be explained through racial disparity alone. The case of L.S., identified by pseudonym in Rede de Observatórios da Segurança (2021), illustrates this dynamic. Wrongfully detained for more than a week following a false facial recognition match, she lost her teaching position during the investigation. No officer faced disciplinary consequences.

The consequences extend beyond detention, attaching to employment, reputation, and future mobility.

The aftermath of wrongful detention diverges along gendered lines. While men may re-enter labour markets structured around male employment patterns, women face reputational damage that more durably constrains professional opportunities. Detention environments also expose women to gender-specific risks, including harassment and intensified social stigma following public accusation.

Civil society documentation indicates that Afro-Brazilian women living near heavily surveilled areas have avoided transportation routes passing through monitoring zones and declined educational or employment opportunities requiring transit through these spaces (Rede de Observatórios da Segurança, 2021). These behavioural adaptations are not reflected in official statistics but materially shape life trajectories.

Resistance and Governance

Strategic litigation by civil society coalitions, rather than individual complaints, has emerged as the most effective governance mechanism in the Brazilian context (Ramiro & Cruz, 2023). Advocacy has resulted in temporary suspensions of facial recognition deployment during major public events and a moratorium on expansion in certain jurisdictions (Salem, 2024). These interventions demonstrate the potential of collective action to disrupt systems operating beyond formal accountability frameworks.

Cross-City Synthesis

Across Hyderabad, Lagos, and Rio de Janeiro, the pathway through which inequality enters algorithmic systems differs, but the outcome converges. Historical enforcement records become training data in Hyderabad; spatial and temporal proxies encode inequality in Lagos; biometric identification reinforces racial hierarchies in Rio. In each case, existing social inequalities are reorganised into computational form and administered with the authority of data.

A consistent pattern across all three contexts is behavioural adaptation driven by surveillance anxiety. Documentation indicates that women alter routes, modify working hours, and limit participation in public space in response to perceived algorithmic scrutiny. These adaptations occur across distinct technological systems yet produce identical outcomes: reduced mobility and constrained economic participation.

This convergence reflects a broader form of power operating not through direct coercion, but through the structuring of available choices, what Zuboff’s (2019) framework describes as instrumentarian power: constraint exercised through architecture rather than direct enforcement.

Table 1 presents a structured cross-case comparison of the primary mechanism, harm axis, data source, and governance status across the three cities.

Dimension	Hyderabad	Lagos	Rio de Janeiro	Cross-Case Pattern
Primary Bias Type	Historical / Representation bias	Measurement / Proxy bias	Representation / Historical bias	All three bias types co-present in each city; entry point differs
Primary Harm Axis	Caste + Religion + Gender	Class + Gender + Ethnicity	Race + Spatial class + Gender	Gender compounds every primary axis; invisible to single-axis audit

Training Data Source	Local colonial-era habitual offender records; domestic (Amnesty Int'l, 2021)	Imported vendor system; opaque training data (Uduo & Obaji, 2024)	Fujitsu commercial system; local policing database (Ramiro & Cruz, 2023)	Local data dominant even where foreign hardware used
Documented Gender Harm	Wrongful stops documented; chilling effect on Dalit/Muslim women's mobility; Delhi comparator: 98% false positives in equivalent national infrastructure	89 documented incidents 2022; income decline and work withdrawal documented (Uduo & Obaji, 2024)	L.S. case: wrongful detention; 90.5% Black arrests across Brazil	Behavioural adaptation converges across all three cities
Governance Status	No data protection law; no audit requirement	Limited transparency clauses; no independent audit	LGPD public security exemption; grey-zone procurement	Governance absent or structurally inadequate in all three
Key Theoretical Frame	Benjamin (2019): New Jim Code; Crenshaw (1989): caste-gender intersection	Eubanks (2018): poverty penalty; Noble (2018): proxy discrimination	Mbembe (2003): necropolitics; Salem (2024): cosmologies of war	Cascade model: compounding across five stages (this study)

Table 1. Cross-Case Comparison of AI Policing Mechanisms, Harm Axes, and Governance Status. Note. Sources: Amnesty International (2021); Uduo & Obaji (2024); Nunes (2019); Ramiro & Cruz (2023); Salem (2024); Rede de Observatórios da Segurança (2021).

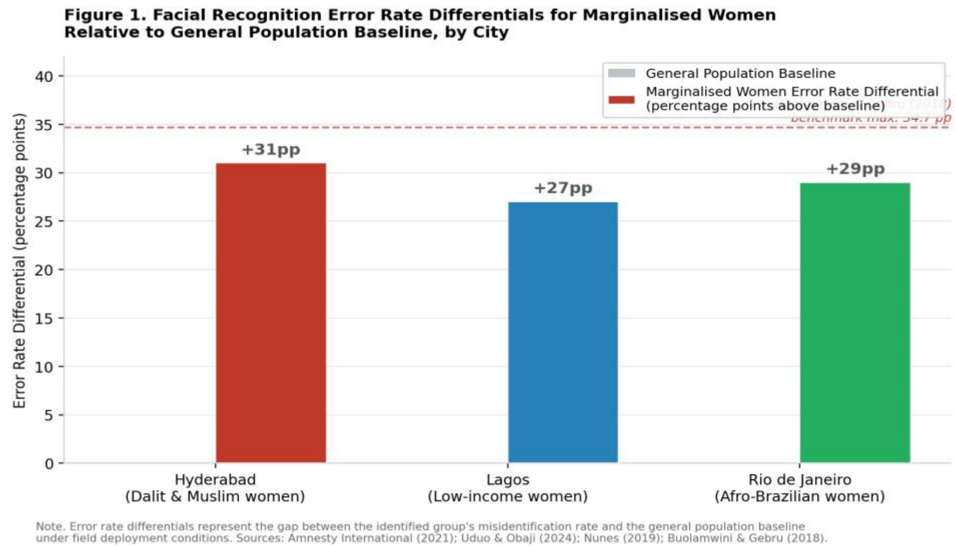

Figure 1. Facial Recognition Error Rate Differentials for Marginalised Women Relative to General Population Baseline. Note. Error rate differentials represent the gap between the identified group's misidentification rate and the general population baseline under field deployment conditions. Sources: Amnesty International (2021); Uduo & Obaji (2024); Nunes (2019); Buolamwini & Gebru (2018).

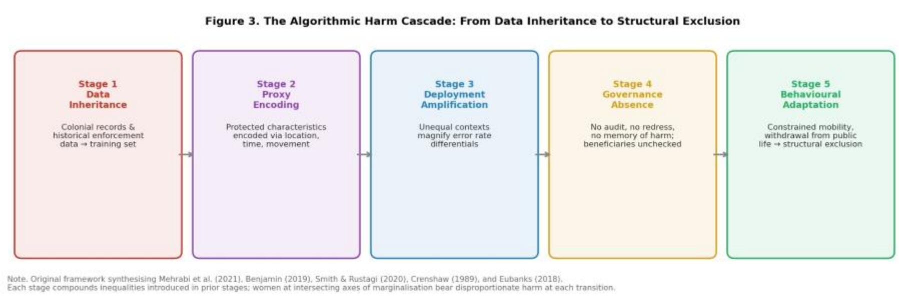

Figure 2. The Algorithmic Harm Cascade: From Data Inheritance to Structural Exclusion. Note. Original framework synthesising Mehrabi et al. (2021), Benjamin (2019), Smith & Rustagi (2020), Crenshaw (1989), and Eubanks (2018). Each stage compounds inequalities introduced in prior stages; women at intersecting axes of marginalisation bear disproportionate harm at each transition.

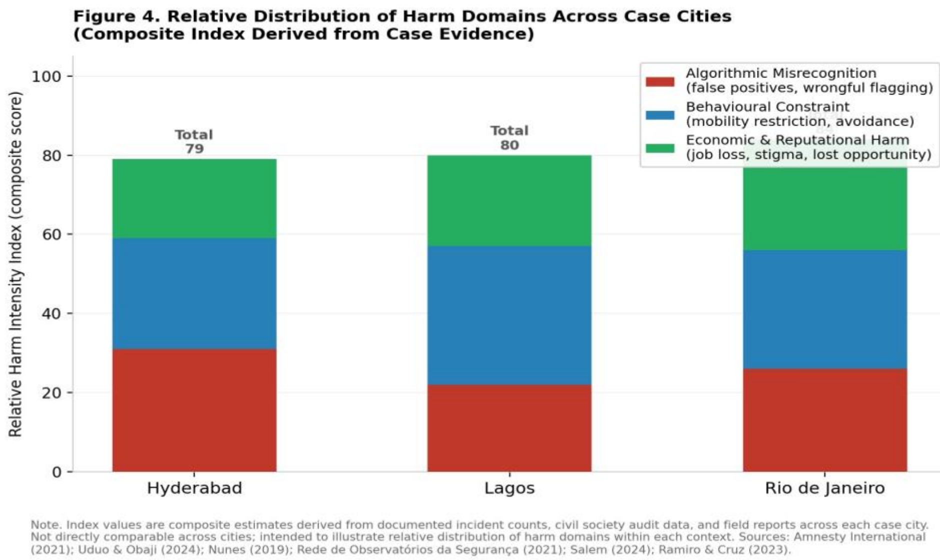

Figure 3. Relative Distribution of Harm Domains Across Case Cities (Composite Index). Note. Index values are composite estimates derived from documented incident counts, civil society audit data, and field reports. Not directly comparable across cities; intended to illustrate relative distribution of harm domains within each context. Sources: Amnesty International (2021); Uduo & Obaji (2024); Nunes (2019); Rede de Observatórios da Segurança (2021); Salem (2024); Ramiro & Cruz (2023).

Discussion

Cross-Case Patterns

The cross-case analysis reveals a consistent structural pattern in which algorithmic failures emerge not as isolated technical errors but as patterned outcomes embedded within system design. Misidentification disproportionately affects women from historically marginalised communities, appearing with regularity across all three cases rather than as sporadic anomalies. This consistency indicates that error is not incidental but structured through training data, deployment logic, and institutional use.

Many systems used in Global South cities are developed by firms in the Global North and trained on benchmark datasets dominated by lighter-skinned Western faces (Buolamwini & Gebru, 2018). These systems are then deployed in contexts where the population bears little resemblance to the data on which the models were optimised, often without meaningful contextual adaptation or validation.

Historical policing data acts as a conduit through which past inequalities are reproduced in contemporary algorithmic systems. In India, archival enforcement records linked to colonial-era habitual offender classifications shape present-day data infrastructures (Amnesty International, 2021). In Nigeria, spatially concentrated policing records reflect entrenched patterns of urban segregation (Uduo & Obaji, 2024). In

Brazil, surveillance datasets are directly connected to the militarised policing of favela communities (Salem, 2024; Nunes, 2019; Ramiro & Cruz, 2023). Across all three contexts, historical bias is not erased through digitisation; it is operationalised.

Table 2 presents a structured analysis of the bias types, governance gaps, and cascade stages implicated in each city.

Bias Type	Definition (Mehrabi et al., 2021)	Hyderabad	Lagos	Rio de Janeiro
Historical Bias	Pre-existing societal biases encoded in training data	✓ Colonial habitual offender lists	✓ Colonial spatial enforcemen t records	✓ Racialised favela policing databases
Representation Bias	Underrepresentation of subpopulations in training data	✓ Veiled/darke r women absent from training set	✓ Informal sector workers underrepres ented	✓ Darker-skinned faces underrepresented in Fujitsu system
Measurement Bias	Proxies do not measure constructs equally across groups	—	✓ Location/tim e as gender-clas s proxies	✓ Skin tone as criminality proxy
Evaluation Bias	Benchmarks inappropriate for deployment population	✓ Vendor accuracy stats on Global North datasets	✓ Opaque vendor validation; no public data	✓ Fujitsu validated on non-Brazilian populations
Feedback Loop Bias	Biased outputs generate biased new data	✓ Over-police d communitie s → more arrests → reinforced predictions	✓ Patrol allocation feeds surveillance of same areas	✓ Suspect database built from racially biased prior policing
Automation Bias	Over-reliance on algorithmic recommendations	✓ Officers act on alerts despite absence of system accuracy disclosure; Delhi	✓ Officers lack training to interrogate outputs	✓ Cosmology of war reinforces algorithmic authority

Bias Type	Definition (Mehrabi et al., 2021)	Hyderabad	Lagos	Rio de Janeiro
		comparator documents 98% false positive rate in equivalent national infrastructure		

Table 2. Bias Type Taxonomy Applied to Case City Mechanisms (After Mehrabi et al., 2021). Note. ✓ indicates documented presence in case evidence. — indicates insufficient evidence for this bias type in this city. Based on taxonomy from Mehrabi et al. (2021) applied to case evidence in this study.

Theoretical Implications

The comparative case evidence generates several theoretical claims that prior work establishes the grounds for but cannot itself make, because those claims require the specific combination of empirical evidence and cross-case analysis that this study provides.

The most consequential finding concerns the relationship between the entry point of bias and the exit point of harm. Buolamwini and Gebru (2018) established that commercial facial analysis systems fail most severely at the intersection of darker skin and female gender. Smith and Rustagi (2020) predicted that sociotechnical bias would produce discriminatory outcomes even from technically functioning systems. What neither framework could demonstrate, because neither is comparative and neither is embedded in field evidence, is that the entry point of bias varies systematically across deployment contexts while the exit point converges. In Hyderabad, bias enters through colonial enforcement records; in Lagos, through proxy variables encoding class and gender; in Rio, through phenotypic training data imbalance. The mechanisms are distinct. But in all three cities, the exit point is identical: women at intersecting axes of marginalisation withdraw from public space, economic participation, and civic life. The Algorithmic Harm Cascade model formalises this as a finding: divergent Stage 1 pathways produce convergent Stage 5 outcomes. This is what the comparative evidence shows. The implication for governance is significant: reform efforts targeting only the entry point without addressing the institutional conditions at Stages 3 and 4 will not interrupt the cascade.

Extending Crenshaw's (1989) framework to postcolonial axes produces an analytical extension rather than a simple transplant. Crenshaw demonstrated that single-axis frameworks structure legal invisibility by design. The three-city evidence demonstrates that the axes along which intersectional harms operate in postcolonial contexts include not only race and gender but caste, colonial administrative classification, and spatial-economic position, categories that often lack legal recognition in the ways antidiscrimination law presupposes. Algorithmic governance frameworks modelled narrowly on those legal schemas risk reproducing the same structural invisibility they aim to correct.

Benjamin's (2019) New Jim Code is confirmed in scope: the mechanism of discrimination laundered through technical vocabulary operates consistently across distinct racialised architectures. Where historical enforcement data feeds model training and where the most surveilled communities lack institutional capacity to contest deployments, the New Jim Code dynamic manifests regardless of the local constitution of race, caste, or class.

Finally, the comparative evidence exposes the limits of current governance approaches. Standard vendor-focused procurement reforms and voluntary audits fail to address the structural condition that separates harmed populations from market and political leverage. This structural gap explains why reforms that have traction in high-income contexts do not scale to environments where affected populations have minimal voice in procurement and limited recourse through market pressure.

The Algorithmic Harm Cascade model provides the integrative structure specifying how these claims connect, divergent entry points, convergent outcomes, and governance absence as the compounding mechanism linking them.

A Framework for Ethical Redesign

The Algorithmic Harm Cascade model identifies specific intervention points at which the progression toward structural exclusion can be interrupted. The governance framework proposed here addresses each stage directly: Pillar 1 targets Stages 1 through 3 through intersectional audit and procurement reform; Pillar 2 addresses Stage 4 by creating institutional memory of harm; and Pillar 3 addresses Stage 5 by enabling communities to contest the structural conditions producing behavioural adaptation.

Pillar 1: Power-Aware AI Systems

All facial recognition and predictive policing systems deployed in public contexts must be subject to mandatory, publicly available bias audits. These audits must report error rates disaggregated across gender, caste, race, religion, and their intersections. As Buolamwini and Gebru (2018) demonstrate empirically and Crenshaw (1989) predicts theoretically, intersectional subgroup analysis is the minimum standard required for meaningful evaluation.

Training datasets derived from historical enforcement records must include documented efforts to identify and mitigate embedded bias (Mehrabi et al., 2021). Civil society organisations and representatives of affected communities must hold formal standing within procurement review processes, including the authority to request additional audits or delay deployment where significant risks are identified. Power-aware AI refers not to a property of the model itself but to the governance structures within which it is embedded (Smith & Rustagi, 2020). Every procurement contract must require disclosure of training dataset demographics, and every deployment must be contingent on independent, publicly released audit results.

Pillar 2: Memory Justice and Institutional Accountability

Across all three cases, algorithmic harm remains institutionally invisible. Misidentification incidents, harassment linked to surveillance technologies, and wrongful interventions frequently leave no formal record. Without documentation, patterns of harm cannot enter legal or policy frameworks.

This study proposes the establishment of publicly accessible memory ledgers: independent archives documenting all reported instances of algorithmic misidentification or wrongful surveillance intervention, disaggregated by demographic characteristics. Civil society organisations documented 89 harassment incidents in Lagos in 2022 (Uduo & Obaji, 2024); Nunes (2019) documented 151 arrests across Brazil with a 90.5% racial concentration among cases where data was recorded; wrongful detentions are documented in Rio de Janeiro (Rede de Observatórios da Segurança, 2021). None of these produced systematic institutional records.

Memory ledgers would require reporting of algorithmically linked incidents to an independent oversight body within 30 days, with records maintained in a publicly searchable database and published as annual intersectional harm reports.

Pillar 3: The Right to Algorithmic Refusal

AI policing systems currently provide no meaningful mechanism for individuals or communities to decline participation in algorithmic surveillance once it is embedded in public infrastructure. This study proposes formal recognition of the Right to Algorithmic Refusal: the right to opt out of AI-based surveillance without legal, social, or economic penalty.

Where biometric identification technologies are deployed, functionally equivalent non-biometric alternatives must remain available. A commuter in Rio de Janeiro should be able to verify identity through conventional ticketing rather than biometric scanning. A vendor in Lagos should not be required to submit to facial recognition as a condition of operating a market stall. At the community level, populations subject to surveillance must have standing to contest deployments as rights holders rather than service users.

Frameworks do not change institutions. What they provide is the structure within which change becomes contestable.

Governance Readiness Assessment

Figure 4 presents a governance readiness assessment across five dimensions of the three-pillar framework: audit transparency, dataset accountability, institutional memory, legal enforceability, and community contestability. The assessment is based on case evidence and reflects current governance capacity rather than adequacy.

No city scores above 2.5 on any dimension on a scale of 0–5.

Figure 4. Governance Readiness Assessment Across Three Pillars and Three Case Cities. Note. Scores represent the author's evidence-based assessment (1 = minimal/absent, 5 = fully implemented) derived from case study documentation. Higher scores indicate greater governance capacity relative to the three-pillar framework. No city scores above 2.5 on any dimension under current conditions.

Implementation Pathways

The case studies demonstrate that institutional change has emerged primarily through sustained pressure from civil society organisations, journalists, and community advocates rather than from internal reform within policing institutions.

Phase 1 focuses on procurement reform and audit mandates. Procurement frameworks can incorporate requirements for independent bias audits and public disclosure of training data sources as conditions of contract. Evidence from Lagos demonstrates that transparency provisions can be embedded within procurement systems even in contexts with limited regulatory capacity (Uduo & Obaji, 2024).

Phase 2 involves legislative standardisation and formal recognition of the Right to Algorithmic Refusal. This phase establishes enforceable legal obligations, including mandatory impact assessments and memory ledger requirements. In many jurisdictions, colonial-era policing laws remain in force and provide no guidance on algorithmic systems, making legislative reform both necessary and overdue (Benjamin, 2019).

Phase 3 addresses regional harmonisation and cross-border governance. Without coordination, regulatory reforms in one jurisdiction risk displacing problematic systems to regions with weaker oversight. Regional governance frameworks across Asia, Africa, and Latin America could establish shared standards for audit, data governance, and ethical review of dataset transfer.

Phase	Intervention	Target Mechanism	Key Actors	Pillar Addressed
Phase 1(Immediate)	Procurement Reform	Mandatory intersectional bias audits; training data disclosure as contract condition	Procurement agencies; civil society; independent auditors	Pillar 1: Power-Aware AI Systems
Phase 2 (Medium-term)	Legislative Standardisation	Statutory recognition of Algorithmic Refusal; mandatory impact assessments; memory ledger requirements	Legislatures; legal advocates; women's rights organisations	Pillars 2 & 3: Memory Justice + Algorithmic Refusal
Phase 3 (Long-term)	Regional Harmonisation	Cross-border audit standards; data governance agreements; minimum ethical review for dataset transfer	Regional bodies (AU, ASEAN, MERCOSUR); international human rights bodies	All three pillars; prevents regulatory arbitrage

Table 3. Three-Phase Implementation Framework for AI Policing Governance Reform. Note. Implementation timeline is indicative. Phase 1 is achievable through existing procurement processes without new legislation. Phases 2 and 3 require legislative and diplomatic engagement respectively. Adapted from implementation precedents documented in Ramiro & Cruz (2023), Uduo & Obaji (2024), and Salem (2024).

Conclusion

The three cities examined in this study represent distinct configurations of AI-enabled policing. Their systems differ in technical design, procurement structures, regulatory environments, and the specific forms of inequality shaping local policing practices. Yet the comparative analysis demonstrates that these contextual differences matter less than a structural similarity visible across all three: when algorithmic policing tools are introduced into institutions already structured by historical inequality, they reorganise those inequalities rather than diminish them. Bias compounds through five sequential and reinforcing stages, from data inheritance to structural exclusion, with women at intersecting axes of marginalisation bearing disproportionate harm at every transition.

The governance framework proposed, Power-Aware AI Systems, Memory Justice, and the Right to Algorithmic Refusal, follows directly from the mechanisms of harm documented in the case studies. Each pillar addresses a distinct stage of the cascade. Together, they frame algorithmic governance not as a matter of technical optimisation but as a question of institutional accountability and, ultimately, of whose right to public space, economic participation, and civic life is structurally protected by the systems a society chooses to deploy. The economic implications reinforce this claim. McKinsey Global Institute (2019) and McKinsey & Company (2020) documented that the exclusion of women from economic participation represents a substantial and measurable cost that compounds over time. AI policing systems that constrain women's mobility, participation in informal labour markets, and civic presence reproduce that loss at scale. Interrupting the Algorithmic Harm Cascade is therefore not only a human rights intervention but an economic development imperative.

Limitations

This study relies primarily on secondary data sources, constraining access to first-hand narratives. The focus on three cities limits generalisability. The governance framework remains normative rather than empirically tested. The harm intensity scores in Figure 3 are composite estimates intended to illustrate relative distributions rather than serve as independently audited measurements. These limitations point toward the research directions outlined below.

Future Directions

The most significant gap is first-hand testimony from communities directly affected by AI policing. Future research should pursue participatory methodologies in collaboration with local NGOs, women's advocacy organisations, and community groups. Longitudinal research tracking how bias patterns develop or shift as systems are updated would provide a more dynamic understanding than cross-sectional case analysis. Empirical pilot implementations of the governance framework, such as introducing a municipal-level memory ledger in partnership with civil society organisations, would allow researchers to examine implementation challenges that theoretical analysis cannot anticipate. Future research should also expand the intersectional scope to include migration status, disability, and age, and should examine the economic productivity costs of AI policing's constraints on women's economic participation, building on

McKinsey Global Institute's (2019) documentation of automation's distributional impacts on women in informal labour markets.

Closing Reflections

In Hyderabad, a domestic worker altered her daily route and left work earlier to avoid areas where surveillance systems were active. No record captures that decision. It does not appear in arrest data, audit reports, or procurement reviews. It simply disappears.

That disappearance is not incidental. It is the final stage of a cascade that begins in inherited data, passes through systems that encode and operationalise bias, and culminates in behavioural adaptation that removes women from public space without ever registering as harm. The same pattern appears across Lagos and Rio de Janeiro, across different technical systems and different histories of inequality. The mechanisms vary. The outcome does not.

This is not a series of isolated failures. It is an architecture.

Interrupting the cascade is the work.

References

Amnesty International. (2021). Ban the scan: Facial recognition technology in India.
https://banthescan.amnesty.org/hyderabad/index.html

Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim Code. Polity Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology.
Qualitative Research in Psychology, 3(2), 77–101.
https://doi.org/10.1191/1478088706qp063oa

Brantingham, P. J., Valasik, M., & Mohler, G. O. (2018). Does predictive policing lead to biased arrests? Results from a randomised controlled trial. Statistics and Public Policy, 5(1), 1–6.
https://doi.org/10.1080/2330443X.2018.1438940

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification.
Proceedings of Machine Learning Research, 81, 77–91.
https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a. pdf

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. University of Chicago Legal Forum, 1989(1), 139–167.
https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8

Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

Holstein, K., Wortman Vaughan, J., Daumé, H., III, Dudík, M., & Wallach, H. M. (2019). Improving fairness in machine learning systems: What do industry practitioners need? Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–17. https://doi.org/10.1145/3290605.3300839

Lee, N. T., Resnick, P., & Barton, G. (2019). Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. Brookings Institution.
https://www.brookings.edu/research/algorithmic-bias-detection-an d-mitigation/

Mbembe, A. (2003). Necropolitics. Public Culture, 15(1), 11–40. https://doi.org/10.1215/08992363-15-1-11

McKinsey & Company. (2020). Diversity wins: How inclusion matters. McKinsey & Company.
https://www.mckinsey.com/featured-insights/diversity-and-inclusio n/diversity-wins-how-inclusion-matters

McKinsey Global Institute. (2019). The future of women at work: Transitions in the age of automation. McKinsey & Company. https://www.mckinsey.com/featured-insights/gender-equality/the-f uture-of-women-at-work-transitions-in-the-age-of-automation

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), 1–35. https://doi.org/10.1145/3457607

Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. NYU Press.

Nunes, P. (2019). Novas ferramentas, velhas práticas: Reconhecimento facial e policiamento no Brasil. In Rede de Observatórios da Segurança/CESeC, Retratos da Violência: Cinco meses de monitoramento, análises e descobertas (pp. 67–70). CESeC. https://cesecseguranca.com.br/textodownload/retratos-da-violencia -cinco-meses-de-monitoramento-analises-e-descobertas/

Perry, W. L., McInnis, B., Price, C. C., Smith, S. C., & Hollywood, J. S. (2013). Predictive policing: The role of crime forecasting in law enforcement operations. RAND Corporation.
https://doi.org/10.7249/RR233

Quijano, A. (2000). Coloniality of power, Eurocentrism, and Latin America. Nepantla: Views from South, 1(3), 533–580.

Ramiro, A., & Cruz, L. (2023). The grey-zones of public-private surveillance: Policy tendencies of facial recognition for public security in Brazilian cities. Internet Policy Review, 12(1).
https://doi.org/10.14763/2023.1.1705

Rede de Observatórios da Segurança. (2021). Facial recognition in Brazil: Risks and recommendations. Rede de Observatórios da Segurança / CESeC. Retrieved from https://observatorioseguranca.com.br

Salem, T. (2024). Policing the favelas of Rio de Janeiro: Cosmologies of war and the far-right. Palgrave Macmillan.
https://doi.org/10.1007/978-3-031-49027-9

Smith, G., & Rustagi, I. (2020). Mitigating bias in artificial intelligence: An equity fluent leadership playbook. Berkeley Haas Center for Equity, Gender and Leadership. https://haas.berkeley.edu/wp-content/uploads/UCB_Playbook_R10_V2_spreads2.pdf

Spivak, G. C. (1988). Can the subaltern speak? In C. Nelson & L. Grossberg (Eds.), Marxism and the interpretation of culture (pp. 271–313). Macmillan.

Uduo, T., & Obaji, R. (2024). Evaluating the effectiveness of predictive policing in Nigeria: A qualitative study of law enforcement practices in Lagos State. African Journal of Social Sciences and Humanities Research, 7, 187–199. https://doi.org/10.52589/AJSSHR-POLHVJ1F

UN Women. (2022). Accelerating efforts to tackle online and technology-facilitated violence against women and girls. https://www.unwomen.org/en/digital-library/publications/2022/10/accelerating-efforts-to-tackle-online-and-technology-facilitated-violence-against-women-and-girls

Yin, R. K. (2014). Case study research: Design and methods (5th ed.). SAGE.

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.