COMPLY54

THE ENFORCEMENT GAP

Why Africa's AI Regulations Need Runtime Enforcement

The Case for Runtime Compliance Infrastructure in African AI Agent Deployments

Oluwajuwon Omotayo

Founder, Comply54 · comply54.io

July 2026

Table of Contents

OVERVIEW

Executive Summary

Africa is not behind on AI regulation. It is behind on enforcement infrastructure. Across the continent, the laws already exist: data protection statutes, central bank directives, national identity acts, anti-money-laundering rules, and emerging AI-specific bills. What is missing, in nearly every production deployment today, is a mechanism by which an autonomous AI agent actually knows any of this before it acts.

This creates a structural mismatch. Regulators in Nigeria, Kenya, South Africa, and across the African Union are moving quickly: fining violators, drafting AI bills, and publishing national AI policies. Meanwhile, the agent frameworks that fintechs, health-tech platforms, and insurers deploy today (LangChain, LangGraph, CrewAI, AutoGen) carry no built-in awareness of jurisdiction, regulatory obligation, or legal consequence. An agent can approve a loan, move funds cross-border, or access a patient record with the same indifference to law as it would in a jurisdiction with no law at all.

This paper argues that the gap between African AI regulation and African AI practice is not a policy problem to be solved with more legislation. It is an infrastructure problem, and it is not hypothetical: in June 2026, the author's own African regulatory contribution to a major open-source AI governance toolkit was merged the same day it was submitted, only for a new Nigerian identity law to change those very rules eight days later. The correction filed the next day entered the same ordinary review queue as any other contribution, with no mechanism to prioritize it ahead of a statute already in force. We call the missing layer runtime enforcement: compliance logic that intercepts an AI agent's actions before execution and evaluates them against jurisdiction-specific policy in real time, offline, and independent of which agent framework is in use.

The paper is structured to first establish that the regulatory foundation is real and enforced (Sections 1 and 2), then to show concretely how today's agent architectures fail to respect it (Section 3), why existing global governance tooling does not close the gap (Section 4), what a runtime enforcement layer looks like in practice (Section 5), and what this means for builders, fintechs, and regulators going forward (Sections 6 and 7). Comply54, an open-source runtime governance layer covering statutory obligations across twelve African jurisdictions, is offered throughout as a working reference implementation, not as the only possible answer, but as proof that the infrastructure is buildable now.

$220M	8 days	12
largest privacy penalty issued by a Global South regulator to date	from a Nigerian statute changing to the global toolkit built to enforce it falling out of date	African jurisdictions with statutory obligations encoded as runtime policy today

SECTION 01

Section 1: Africa's AI Moment and Its Governance Paradox

1.1 Adoption is already at production scale

AI adoption across Africa is accelerating on a timeline that has surprised even close observers of the market, and it is concentrated precisely in the sectors where regulation bites hardest: finance, health, and identity. This is not adoption at the level of pilots and proofs of concept. It is adoption at the level of live systems making consequential decisions about real people, at volume, today.

In Nigeria, the Central Bank's 2025 Fintech Report found that 37.5 percent of surveyed fintechs already use AI for credit scoring and risk assessment, a further 37.5 percent apply it to customer onboarding and KYC, and 87.5 percent use it for fraud detection. In Kenya, the pattern is further advanced: the Central Bank of Kenya's March 2025 survey on AI in the banking sector found that half of all surveyed lenders had already adopted AI, and that among those, 65 percent apply it specifically to credit risk assessment, its single most common use. Adoption runs highest among digital credit providers, 80 percent of which use AI to score borrower risk, followed by microfinance banks at 75 percent; a further 83 percent of institutions not yet using AI said they intend to adopt it for credit scoring. These are decisions, made by software, that determine who receives credit and on what terms, taken against a backdrop of record loan defaults that gives lenders every incentive to delegate more of the judgment to models.[1][2]

The scale is made concrete by individual operators. MTN Group describes digital lenders moving from processing a few hundred loans a day to approving on the order of a hundred thousand in the same period, with large language models increasingly deployed to analyse the underlying data. M-KOPA, serving millions of underbanked customers across several African markets, reports processing hundreds of payments per minute and extending hundreds of thousands of additional credit lines through AI-driven risk assessment.[3]

[1]Central Bank of Nigeria, 2025 Fintech Report; see also TechAfrica News and Ecofin Agency summaries of the Report's AI-adoption findings.
[2]Central Bank of Kenya, Survey on the Use of Artificial Intelligence in the Banking Sector (March 2025), as reported by Business Daily Africa, “Credit risk assessment tops AI use by Kenyan lenders,” 6 July 2025.
[3]African.business, “Banks and fintechs drive surge in AI-approved loans,” April 2025 (MTN Group and M-KOPA figures).

The lending models themselves increasingly draw on deeply personal data (mobile-money histories, airtime top-up patterns, SMS content, location, and social-media activity), which are exactly the categories of sensitive personal data that Section 2's statutes were written to protect. Every one of these systems is, in regulatory terms, an automated decision engine processing sensitive personal data at scale, which is to say precisely the thing the continent's data-protection and financial-conduct laws now govern.

1.2 The regulatory response has kept pace

What is less widely appreciated outside the continent is that the regulatory response has kept pace, and in some respects moved ahead of global expectations. The African Union's Continental AI Strategy, endorsed by the AU Executive Council in Accra in July 2024, entered Phase 1 implementation across its member states in 2025, a phase focused on governance frameworks, national AI strategies, and resource mobilisation. Counts of exactly how many of the AU's 55 member states have adopted a national AI strategy vary by source and by the date of counting, with 2026 tallies ranging from roughly eight to sixteen countries with adopted strategies and several more in draft; what is consistent across every count is the direction and pace of travel, which has been accelerating steadily since 2023. Nigeria, for its part, climbed 31 places in the Oxford Government AI Readiness Index, from 103rd in 2023 to 72nd out of 195 countries in 2025, the fourth-highest ranking in Sub-Saharan Africa.[4][5]

Nor is the movement confined to strategy documents. As Section 2 sets out in detail, Nigeria's GAID took binding effect in September 2025, its NIMC Act was signed in June 2026, Kenya's Data Protection Act has been enforced since 2019 with a dedicated AI Bill now before the Senate, and South Africa's Information Regulator enforces POPIA's automated-decision provisions today. This is not a continent waiting for regulation to catch up with technology. It is a continent where regulation and technology are advancing on parallel tracks.

1.3 The paradox: parallel tracks that never meet

The trouble is that parallel tracks, by definition, do not intersect. The laws are real and the adoption is real, but there is no mechanism connecting the two at the point where it would matter: no layer that translates a statute into something an autonomous system can actually check itself against before it acts. A Kenyan digital lender's AI can ingest 250 behavioural signals to price a loan in seconds, but it holds no representation of the borrower's right, under the pending AI Bill, to a human review of that decision. A Nigerian fintech's fraud model can screen a hundred thousand transactions an hour, but nothing in it knows what a Tier 1 KYC ceiling is. The sophistication is entirely on the capability side and entirely absent on the compliance side.

[4]African Union, Continental Artificial Intelligence Strategy, endorsed by the Executive Council at its 45th Ordinary Session, Accra, 18–19 July 2024.
[5]Oxford Insights, Government AI Readiness Index 2025. Nigeria ranked 72nd of 195 (from 103rd in 2023).

This is the governance paradox at the centre of this paper. It is not a policy problem, because writing more law will not close a gap that already has law on both sides of it. It is not a capability problem, because the AI is demonstrably capable. It is an infrastructure problem: the missing piece is the connective layer that carries the law into the moment of machine action. The remainder of this paper is about what that layer is, why nothing currently deployed provides it, and what it looks like to build.

SECTION 02

Section 2: What Africa Actually Has: The Regulatory Reality

A common assumption among global AI governance commentators is that Africa represents a regulatory vacuum, fertile ground for AI deployment precisely because no one is watching. That assumption does not survive contact with the actual statutory landscape, or with recent enforcement history. This section works through the evidence jurisdiction by jurisdiction, because the specifics matter: the strength of the argument that follows depends on these being real, current, and verifiable instruments rather than a generalised impression of “emerging regulation.”

2.1 Nigeria

Nigeria has the most developed statutory stack on the continent, and it is worth setting out precisely, because it is also the jurisdiction where Comply54's policy packs are most mature.

The Nigeria Data Protection Act (NDPA) 2023 and the GAID 2025

The NDPA 2023 gave Nigeria's data protection regime, for the first time, real legislative force, replacing the older Nigeria Data Protection Regulation (NDPR) 2019, which had operated without a comparable statutory backbone. The Nigeria Data Protection Commission (NDPC) subsequently issued the General Application and Implementation Directive (GAID) on 20 March 2025, which took effect on 19 September 2025 after a six-month transition period. The GAID is the operative instrument that translates the NDPA's obligations into practice, and it speaks directly to AI: it requires that any deployment of an emerging technology, including artificial intelligence, IoT, or blockchain, be preceded by a Data Protection Impact Assessment (DPIA). That DPIA must be signed by an accredited Data Protection Officer, follow the GAID's Schedule 4 template, cover privacy-by-design measures and disparate-impact evaluation, and be filed within four months of deployment for processing of sensitive personal data, or within six months for processing that predates the GAID. The DPIA outcome becomes part of the organisation's annual Compliance Audit Return. This is the statutory hook behind this paper's central claim about Nigerian fintechs: an organisation running an AI credit-scoring or KYC model without a documented, DPO-certified

DPIA on file is not in a grey area. It is out of compliance with a directive that has been in force since September 2025.[6]

Enforcement: real, but uneven

Nigerian regulators have shown they are willing to act, though the record is more layered than a single headline figure suggests. In July 2024, following a 38-month joint investigation by the Federal Competition and Consumer Protection Commission (FCCPC) and the NDPC into Meta and WhatsApp's data practices, the FCCPC issued a $220 million administrative penalty for discriminatory and exploitative treatment of Nigerian consumers and unauthorised data sharing, a decision the Competition and Consumer Protection Tribunal upheld in full on 25 April 2025 after Meta's appeal failed. It remains, by most measures, the largest privacy-related penalty issued by a Global South regulator to date, though it is worth being precise that the finding was made jointly and the penalty itself was issued under the FCCPC's consumer-protection authority (the FCCPA 2018 and the pre-NDPA NDPR 2019), not as an NDPA action in isolation.[7]

The NDPC's own enforcement under the NDPA specifically has had a more mixed run. In February 2025, the NDPC issued a separate $32.8 million remedial fine directly against Meta under the NDPA, over unconsented behavioural advertising, processing of non-users' data, and unauthorised cross-border transfers affecting an estimated 61 million Nigerians. That fine, however, was set aside in November 2025 through a consent judgment agreed between Meta and the NDPC at the Federal High Court in Abuja, a resolution the Data Privacy Lawyers Association of Nigeria has since formally challenged, arguing the Commission lacks statutory authority to waive a remedial fine once issued. Whichever way that dispute resolves, it is a useful illustration in its own right: even where a regulator successfully applies the NDPA to an AI-adjacent processing failure, the path from finding to enforced remedy is not guaranteed to hold.[8]

The cleaner and less contested precedent is the NDPC's action against Multichoice Nigeria. Following an investigation opened in the second quarter of 2024, the NDPC imposed a ₦766,242,500 (approximately ₦766.2 million) administrative penalty on 6 June 2025 for unlawful cross-border transfer of subscriber and non-subscriber personal data and for data processing the Commission described as intrusive, unfair, and disproportionate under the NDPA. Unlike the Meta cases, this fine has not been challenged or set aside, and it stands as the clearest example to date of the NDPC directly enforcing the NDPA against a

[6]Nigeria Data Protection Commission, General Application and Implementation Directive (GAID) 2025, issued 20 March 2025, effective 19 September 2025; see also Banwo & Ighodalo and Lexology analyses of the GAID's emerging-technology and DPIA provisions.

[7]Federal Competition and Consumer Protection Commission (FCCPC) v. Meta Platforms / WhatsApp; $220m penalty upheld by the Competition and Consumer Protection Tribunal, 25 April 2025. Reported by TechPoint Africa and Silicon.co.uk.

[8]Nigeria Data Protection Commission $32.8m remedial fine against Meta (February 2025) and subsequent Federal High Court consent judgment (November 2025); reported by TheNigeriaLawyer and SaharaReporters. The consent judgment has been challenged by the Data Privacy Lawyers Association of Nigeria.

private processor over cross-border data handling, exactly the kind of processing an AI agent moving data between systems and jurisdictions would need to be checked against.[9]

National identity: the NIMC Act 2026

President Bola Tinubu signed the National Identity Management Commission (NIMC) Act 2026 into law on 26 June 2026, repealing the NIMC Act 2007 after nearly two decades without an update. The new Act designates NIMC as Nigeria's Root Certification Authority for its National Public Key Infrastructure and Digital Public Infrastructure, positions the National Identification Number as the country's foundational identity credential under a "one person, one identity" principle, and, importantly for this paper's purposes, brings NIN-linked data explicitly under the NDPA's framework for the first time. Third-party access to National Identity Database records is now consent-first by default, with narrowly defined exceptions for court orders, criminal investigation, and public interest. The Act also sharply raises the cost of getting this wrong: unauthorized access to the database carries a minimum penalty of five years' imprisonment or a ₦10 million fine for individuals, rising to a minimum ₦20 million fine for corporate entities, with personal criminal liability reaching responsible executives. An AI agent that bulk-exports or shares NIN-linked records outside these narrow exceptions is now operating in direct violation of a specific, recently strengthened criminal statute, not a general privacy principle.[10]

Layered on top of these is the Central Bank of Nigeria and the Nigerian Financial Intelligence Unit's existing anti-money-laundering and Know-Your-Customer tiering regime, and the CBN's 2025 Open Banking Framework, which classifies credit-scoring data as "high and sensitive risk" and requires standardised APIs and strict consent management, alongside the FCCPC's 2025 digital lending rules, which require registration for qualified digital lenders and prohibit undisclosed, opaque, "black box" interest-rate setting.[11]

2.2 Kenya

Kenya's Data Protection Act (KDPA) 2019 already gives citizens meaningful rights over automated processing, including the right to object to processing based on automated decision-making. Building on that base, the Artificial Intelligence Bill, 2026, sponsored by nominated Senator Karen Nyamu, was introduced in February 2026 and received its first reading in the Senate on 2 April 2026, where it was committed to the Standing Committee on ICT for public input.[12][13]

[9]Nigeria Data Protection Commission penalty of ₦766,242,500 against MultiChoice Nigeria, 6 June 2025, for unlawful cross-border data transfer under the NDPA. Reported by TechAfrica News and Business Post Nigeria.
[10]National Identity Management Commission (NIMC) Act 2026, signed 26 June 2026; reported by State House Abuja and Premium Times. Repeals the NIMC Act 2007.
[11]Central Bank of Nigeria, Open Banking Framework (2025); FCCPC digital lending / consumer-credit regulations (2025).
[12]Kenya Data Protection Act, 2019 (No. 24 of 2019).
[13]Artificial Intelligence Bill, 2026 (Kenya), sponsored by Senator Karen Nyamu; first reading in the Senate 2 April 2026. Analysis by Bowmans and Business Daily Africa. As of mid-2026 the Bill remains pending and is not yet enacted law.

As drafted, the Bill would create an Office of the Artificial Intelligence Commissioner with powers to register high-risk AI systems, inspect deployments, and impose administrative fines. Its substantive protections are significant for this paper's argument: it grants citizens a right to human review of automated decisions in areas such as employment screening, loan approvals, welfare support, and insurance assessment, a right to a plain-language explanation of how the decision was reached, and a right to challenge the outcome. Companies deploying AI systems would be required to disclose the system's purpose and limitations and to explain the steps taken to identify and address algorithmic bias, with fines of up to KES 1 million for non-disclosure under some reporting and up to KES 5 million and custodial penalties for more serious violations under others. Critically for the argument in Section 3, the Bill requires “human-in-the-loop” oversight designed into any AI system making decisions with significant effects on individuals.

It is important to be precise about the Bill's current status: as of mid-2026 it remains a Senate Bill that has not yet been tabled before the full Senate for a vote, and because it touches on matters affecting county governments, it will also need to go before the National Assembly before it can reach presidential assent. Legal commentators, including Bowman's and contributors to Business Daily, have raised substantive concerns that the Bill may be premature relative to Kenya's National AI Strategy 2025 to 2030 and the parallel national AI policy process the Senate itself directed the ICT Ministry to develop in March 2026, and that its institutional design, three new regulatory bodies layered onto an existing Data Protection Commissioner with overlapping authority, risks duplicating rather than closing gaps. The Bill is therefore best read as a strong signal of regulatory direction and an increasingly firm floor of expected obligations, rather than as settled, enforceable law today.

2.3 South Africa

South Africa's Protection of Personal Information Act (POPIA) is the operative statute today, and it already reaches automated decision-making directly: Section 71 gives individuals a right not to be subject to decisions based solely on automated processing that significantly affects them, enforced by the Information Regulator.[14]

On 10 April 2026, South Africa's Department of Communications and Digital Technologies published a Draft National AI Policy in the Government Gazette for a 60-day public comment period, having cleared Cabinet approval in March 2026. The draft proposed a sector-specific, risk-based governance model layered onto existing law rather than a standalone AI Act, alongside new bodies including a National AI Commission, an AI Ethics Board, an AI Regulatory Authority, and an AI Ombudsperson. It is necessary to report accurately what happened next, since it bears directly on this paper's argument. On 26 April 2026, just over two weeks after publication, the government withdrew the draft policy after journalists and academic reviewers found that roughly ten percent of the 67 academic references in its bibliography

[14]Protection of Personal Information Act 4 of 2013 (South Africa), s.71 (automated decision-making), enforced by the Information Regulator.

were fabricated, citing articles in journals that confirmed the pieces had never been published. The pattern was consistent with unverified generative-AI output: a policy document about governing AI appears to have been drafted, in part, using an AI tool whose citations were never checked. As of this writing, the draft has not been reissued, no timeline has been confirmed for a replacement, and South Africa's operative AI-relevant law remains POPIA alone, applied through the Information Regulator and, for automated decisions specifically, Section 71.[15]

This episode is worth stating plainly rather than passing over, because it is a live, concrete demonstration of exactly the failure mode this paper is about: a system, in this case a policy drafting process, that produced confident, plausible-looking output with no mechanism checking it against ground truth before it was acted on. The stakes for a lending or health-data agent making the same kind of error in production are considerably higher than an embarrassed retraction.

2.4 Ghana, Rwanda, Egypt, Mauritius, and Ethiopia

Beyond Nigeria, Kenya, and South Africa, a further tier of African states has adopted national AI strategies, though, consistent with the continent-wide pattern, few have yet followed strategy with binding legislation. Mauritius adopted the continent's first national AI strategy in 2018. Egypt followed with a National AI Strategy in 2019, updated by an Egyptian Charter for Responsible AI in 2023, both administered through a national AI Centre of Excellence. Ghana's National AI Strategy, first adopted in 2022, received renewed Cabinet approval in February 2026. Rwanda's National AI Policy, adopted in 2023 with support from Germany's GIZ and the World Economic Forum, frames AI development explicitly around building local skills alongside global partnerships. Ethiopia adopted a National Artificial Intelligence Policy in 2024, administered by the Ethiopian Artificial Intelligence Institute, which is also tasked with drafting future AI-specific legislation. In each case, the same pattern recurs, and Section 2.5 draws it together: a strategy document sets direction, but no operative mechanism yet exists to make an AI system deployed in that country check its own actions against the obligations the strategy describes.[16]

2.5 The continental layer: the AU Continental AI Strategy

The African Union Executive Council endorsed the Continental AI Strategy at its 45th Ordinary Session in Accra on 18 to 19 July 2024. The Strategy runs from 2025 to 2030 across two implementation phases: Phase 1 (2025 to 2026), the phase currently underway, is focused on establishing governance structures, developing national AI strategies, mobilising resources, and building institutional capacity, including AI advisory boards and centres of excellence; Phase 2, beginning in 2028 following a 2027 review, will focus on executing the Strategy's core projects. The Strategy explicitly calls on member states without existing AI governance frameworks to develop them, and calls for harmonisation of data protection principles

[15]South Africa Draft National AI Policy, published in the Government Gazette 10 April 2026 and withdrawn 26 April 2026 after fabricated academic citations were identified; reported by CNBC Africa and Polity.org.za. See also DLA Piper commentary.
[16]National AI strategies: Mauritius (2018), Egypt (2019; Charter for Responsible AI 2023), Ghana (2022; Cabinet re-approval February 2026), Rwanda (2023), Ethiopia (2024).

across the continent's fragmented regulatory landscape. Independent research tracking the Strategy's first eighteen months of implementation has found progress concentrated geographically, with the large majority of AI-related funding flowing to a handful of countries, a reminder that continental direction and uniform national capacity are two different things.[17]

Taken together, the pattern across every jurisdiction examined in this section is consistent, and it is the central claim of the paper: Africa does not have an AI regulation gap. It has an accountability gap: a gap between what the law already requires (in force, enforceable, and in Nigeria's case actively enforced) and what the AI systems currently deployed across the continent are capable of checking themselves against before they act. And accountability gaps, unlike regulation gaps, are not closed by writing new statutes. They are closed by building the infrastructure that makes the statutes already on the books enforceable at the point of execution.

> ***Africa does not have an AI regulation gap. It has an accountability gap.***

Sources for the claims in this section are given in the footnotes. Regulatory status reflects the position as of July 2026; Kenya's AI Bill and South Africa's national AI policy were both in active development at the time of writing.

SECTION 03

Section 3: How AI Agents Currently Operate in African Regulated Sectors

The dominant agent orchestration frameworks in production today (LangChain, LangGraph, CrewAI, and AutoGen) are genuinely capable systems. They can reason over multi-step plans, call external APIs, move funds, and read and write to sensitive data stores. What none of them do, by default or by design, is carry any awareness of jurisdiction, applicable regulation, or legal consequence. A LangGraph agent making a lending decision in Lagos operates on exactly the same substrate as one operating in a jurisdiction with no financial regulation whatsoever.

3.1 Why the blindness is architectural, not accidental

It is worth being precise about why this is so, because the answer determines what kind of fix is possible. An agent framework's core abstraction is the tool-calling loop: the model reasons about a goal, selects a tool from the set it has been given, produces the tool's arguments, and the framework executes the call and feeds the result back for the next reasoning step. Within that loop, every tool the agent has been granted is, by construction, equally permissible. The framework's job is orchestration, not judgment; it faithfully executes whatever call the model produces, with whatever arguments the model produces.

[17] African Union Continental AI Strategy (2025–2030), Phase 1 (2025–2026) and Phase 2 (from 2028, following a 2027 review). Independent implementation tracking indicates AI-related funding concentrated in a small number of member states.

There is no native concept of a call that is technically valid but legally prohibited, no notion that transfer_funds with amount=5,000,000 differs in kind from the same call with amount=5,000, and no representation anywhere in the stack of who the customer is under a regulatory classification scheme like the CBN's KYC tiers.

This is not a criticism of the frameworks; jurisdiction-awareness was never their design goal, and building it into a general-purpose orchestration library would be the wrong layer for it. But it means the blindness cannot be prompted away. Instructing an agent in its system prompt to “comply with CBN regulations” produces a system that complies only when the model happens to recall the rule, interpret it correctly, and choose to apply it. Those are three probabilistic steps, each of which fails some percentage of the time, on a task where the acceptable failure rate is zero. Compliance that depends on a language model's recall of a statute is not compliance. It is a well-intentioned guess, repeated thousands of times a day.

> ***Compliance that depends on a language model's recall of a statute is not compliance. It is a well-intentioned guess, repeated thousands of times a day.***

What follows are four failure scenarios showing what that architecture does when it meets real African statutory obligations. They are adapted from the Agent Disaster Lab, a set of public demonstration scenarios maintained as part of the Comply54 open-source project, and each is anchored to a specific, verifiable provision of Nigerian law. None requires a malicious actor, a jailbreak, or a broken model. In every case the agent does exactly what it was built to do.

3.2 Scenario one: the Tier 1 transfer

A digital lender deploys a payments agent that can initiate transfers on customer instruction. A customer holding a Tier 1 account, the lowest verification level in the CBN's three-tiered KYC framework, opened with only a name, phone number, and photograph, requests a transfer of ₦5,000,000. The agent validates that the account exists, that the balance covers the amount, and that the instruction is well-formed, and it executes.

Under the CBN's tiered KYC framework, a Tier 1 account is capped at a maximum single deposit of ₦50,000 and a maximum cumulative balance of ₦300,000, precisely because the identity behind it has barely been verified. The transfer the agent just executed is one hundred times the single-transaction ceiling for that account class. Nothing in the agent's execution path represented the concept of a KYC tier at all: the account lookup returned an account, the balance check returned a number, and the tool executed. The tiering system exists to make low-verification accounts useless for laundering exactly this kind of sum; an agent that cannot see tiers dismantles that control silently, at scale, on every request. The institution's exposure is not hypothetical: it is a per-transaction breach of a CBN prudential control,

discoverable in any routine examination, multiplied by every similar transfer the agent has ever processed.[18]

3.3 Scenario two: the health record disclosure

A hospital group deploys an assistant agent with read access to its patient record system, intended to help administrative staff answer routine queries. An HR officer at a partner organisation asks the agent, through an integration, whether a named job candidate has any relevant medical history. The agent retrieves the record, finds an HIV-positive status, and includes it in a helpful summary.

Two separate Nigerian statutes were just violated in one tool call. Section 26(1) of the National Health Act 2014 makes all information concerning a patient's health status, treatment, or stay in a health establishment confidential, and Section 26(2) permits disclosure only with the patient's written consent, under a court order or legal requirement, or where non-disclosure poses a serious threat to public health; Section 29 separately obliges the health establishment to maintain access controls preventing exactly this kind of unauthorised retrieval. And because the condition disclosed was HIV status and the recipient was an employer, the HIV and AIDS (Anti-Discrimination) Act 2014 is also engaged: it specifically prohibits disclosure of an employee's or candidate's HIV status without consent. The agent had no representation of any of this. To the retrieval tool, an HIV status is a string in a field, indistinguishable from a blood type. The failure is not that the model was careless; it is that consent, purpose limitation, and the special status of certain data categories exist nowhere in the layer where the action happened.[19]

3.4 Scenario three: the bulk identity export

A fintech's data engineering team deploys an agent to automate analytics pipeline tasks. Asked to prepare a dataset for a churn analysis, the agent selects the most efficient path: it exports a customer table, including 12,000 Bank Verification Numbers, to a cloud storage bucket in a us-east-1 region, no consent records checked, no data transfer agreement in place.

This is a near-exact replay of the fact pattern behind the largest uncontested NDPA enforcement action to date. As Section 2 detailed, the NDPC fined Multichoice Nigeria ₦766.2 million in June 2025 for unlawful cross-border transfer of personal data without appropriate safeguards; Part VIII of the NDPA permits transfers outside Nigeria only where the destination ensures an adequate level of protection or where safeguards such as binding corporate rules or standard contractual clauses are in place, and the GAID's Schedule 5 sets out the permitted mechanisms in detail. A human data engineer might know this; a data team's review process might catch it. The agent optimised for the task it was given, and the task

[18]CBN tiered Know-Your-Customer framework (Circular FPR/DIR/GEN/CIR/07/003 and related guidance): Tier 1 accounts are subject to a maximum single deposit of ₦50,000 and a maximum cumulative balance of ₦300,000. Encoded in the comply54 nigeria/cbn policy pack.

[19]National Health Act 2014 (Nigeria), ss. 26 and 29 (confidentiality of and access controls over patient information); HIV and AIDS (Anti-Discrimination) Act 2014, prohibiting disclosure of HIV status without consent.

said nothing about Part VIII. The organisation now has 12,000 citizens' foundational financial identifiers sitting in a foreign jurisdiction with no legal basis, a fact pattern a regulator has already shown it will fine at scale, created in the time it took one tool call to complete.[20]

3.5 Scenario four: the screening that never happened

A cross-border payments platform deploys an agent to route and settle transactions. A transfer instruction arrives for a counterparty that appears on a published sanctions list. The agent checks liquidity, selects the cheapest corridor, and settles. No screening step exists anywhere in its tool chain, because nobody built one, and the framework certainly did not supply one.

Under the Money Laundering (Prevention and Prohibition) Act 2022, Nigerian financial institutions are required to conduct customer due diligence and to file suspicious transaction reports with the NFIU within a statutorily mandated window. Those obligations were written on the assumption that a human or a purpose-built compliance system stands between an instruction and its settlement. An autonomous routing agent inverts that assumption: it is precisely the component standing between instruction and settlement, and it carries none of the obligations. This scenario also illustrates a subtler point that matters for Section 5: a partial control can mask the gap. If the platform happens to enforce an unrelated per-transaction cap, large sanctioned transfers get blocked by coincidence and the institution believes screening is happening; a transfer below the cap to the same counterparty settles untouched. Controls that fire for the wrong reason are, in some ways, worse than no controls, because they generate false confidence precisely where scrutiny is most needed.[21]

3.6 The systemic picture

None of these four failures requires anything to go wrong in the ordinary engineering sense. The model performed. The tools executed. The task was completed. The failures live entirely in the gap between what the system was asked to do and what the law required it to refuse. And the scale of the exposure is not marginal: under the GAID's DPIA mandate, in force since 19 September 2025, any organisation deploying AI over personal data without a documented, DPO-certified impact assessment is already in a state of technical non-compliance, before any of the scenarios above ever occurs. The problem, in other words, is systemic rather than exceptional. It is the default state of the market, not the failure of a few bad actors, and it will remain the default until jurisdiction-awareness is built into the layer where the actions actually happen.

SECTION 04

[20]Nigeria Data Protection Act 2023, Part VIII (cross-border transfer of personal data); GAID 2025, Schedule 5 (permitted transfer mechanisms).

[21]Money Laundering (Prevention and Prohibition) Act 2022 (Nigeria); Nigerian Financial Intelligence Unit (NFIU) AML/CFT guidelines, including customer due diligence and suspicious-transaction-report obligations.

Section 4: Why Existing Solutions Don't Close the Gap

It is tempting to assume that global AI governance tooling already solves the problem described in Section 3, and simply needs to be adopted more widely across African deployments. This section examines the two main categories of existing tooling, general-purpose agent governance platforms and content-level guardrail frameworks, and shows why each, despite being genuinely valuable, leaves the accountability gap open. The examination is deliberately specific rather than dismissive: these are serious tools, and understanding precisely what they do well is the fastest way to see precisely what they do not do.

4.1 General-purpose agent governance: the Agent Governance Toolkit

The most significant entry in this category is Microsoft's Agent Governance Toolkit (AGT), released as an open-source project under the MIT licence on 2 April 2026. AGT deserves detailed treatment because it is, architecturally, the closest thing to the runtime enforcement model this paper argues for, and because it demonstrates that the world's largest software companies now agree that governance belongs in the execution layer. Its policy engine intercepts agent actions before execution with sub-millisecond evaluation, supports policies written in YAML, Rego, or Cedar, and integrates with agent frameworks through their native extension points: LangChain's callback handlers, CrewAI's task decorators, and Microsoft's own Agent Framework middleware, among others. Microsoft positions it as the first toolkit to address all ten risks in the OWASP Top 10 for Agentic Applications, the taxonomy published in December 2025 covering goal hijacking, tool misuse, identity abuse, memory poisoning, cascading failures, and rogue agents. On the architectural argument of this paper, in other words, AGT is an ally, not a counterexample: pre-execution interception is exactly right.[22][23]

The gap is in what the policies know. AGT's compliance package maps governance evidence to the regulatory frameworks its designers anticipated: the EU AI Act, HIPAA, and SOC 2. These are the frameworks of the markets the toolkit was built in and for. Nothing in the shipped toolkit knows what a CBN KYC tier is, what the GAID's DPIA mandate requires, what the NIMC Act 2026 prohibits, or what NDPA Part VIII demands before personal data crosses a border. The enforcement machinery is world-class; the African regulatory content it would need to enforce is simply not there.

The author can speak to this gap firsthand. In June 2026, the author opened an issue against AGT stating the problem plainly: the toolkit covered OWASP, NIST AI RMF, the EU AI Act, SOC 2, and HIPAA, but had no African regulatory framework coverage at all, while AI agents were already deployed in Nigerian fintech, Kenyan healthtech, and South African banking. The accompanying contribution, an African regulatory policy pack drawn from the author's agt-policies-nigeria project and covering the NDPA 2023,

[22]Microsoft, “Introducing the Agent Governance Toolkit: Open-source runtime security for AI agents,” Microsoft Open Source Blog, 2 April 2026. AGT is released under the MIT licence and maps compliance evidence to the EU AI Act, HIPAA, and SOC 2.
[23]OWASP Top 10 for Agentic Applications (2026), published December 2025.

CBN and NFIU obligations, BVN and NIN protection, Kenya's Data Protection Act 2019, and POPIA, was accepted and merged the same day it was submitted (PR #3077, 16 June 2026), followed two days later by an East African pack covering Uganda's DPPA, Tanzania's PDPA, and Ethiopia's data protection framework (PR #3110, merged 18 June 2026). Every element of this record is publicly verifiable on the AGT repository.[24]

The experience was instructive in both directions. It proved that AGT's architecture could carry African regulatory logic, and that the maintainers welcomed it. But it also produced a live, measurable demonstration of the structural problem. On 26 June 2026, eight days after the second pack merged, Nigeria signed the NIMC Act 2026 into law, materially changing the obligations around NIN data that the merged pack encoded. The author filed a follow-up contribution adding the new Act's obligations the very next day (PR #3196, 27 June 2026). Whether that specific correction took a day or a season to review, the fact that it entered an ordinary review queue at all, with no mechanism to prioritize it ahead of a statute already in force, is the accountability gap in miniature, timestamped in a public repository. Policy packs that are not continuously maintained against a moving statutory landscape decay into a compliance risk of their own, one that is arguably worse than no coverage, because it carries false authority. A global toolkit can host African compliance content, and to its credit AGT does; what it cannot do, by construction, is own the obligation to keep that content current. That obligation has to live somewhere whose entire purpose is African regulatory fidelity.

4.2 Content-level guardrails: validating outputs, not actions

The second category operates at a different layer entirely. Frameworks such as Guardrails AI, an open-source Python framework with a hub of over one hundred composable validators, and NVIDIA's NeMo Guardrails, which uses a domain-specific language to define conversational rails, intercept the inputs and outputs of language models and validate the content: detecting toxicity, personally identifiable information, hallucinated claims, jailbreak attempts, and structural malformation. Within that scope they are effective and widely used, and nothing in this paper argues against deploying them.[25]

But content validation answers a different question than compliance enforcement. A content guardrail inspects what the model said; a compliance layer must evaluate what the agent is about to do, against who is affected and what the law permits. Consider Section 3's Tier 1 transfer through a content lens: the agent's output, a well-formed transfer instruction, contains no toxicity, no PII leak, no hallucination, and no jailbreak. Every content validator passes, because the content is unimpeachable. The violation lives entirely in the relationship between the action's parameters and a regulatory fact the validator has no access to: the customer's KYC tier and the CBN cap attached to it. The same is true of the health record

[24] microsoft/agent-governance-toolkit, issue #3043 and PR #3077 (merged 16 June 2026), PR #3110 (merged 18 June 2026), and PR #3196 (filed 27 June 2026), contributed by GitHub user kingztech2019. The underlying policy source is kingztech2019/agt-policies-nigeria (306 passing OPA tests). Readers can check the current status of any of these directly on the AGT repository.

[25] Guardrails AI (open-source Python validation framework) and NVIDIA NeMo Guardrails (conversational rail DSL).

disclosure, where a PII detector might flag that health data is present but cannot know whether this recipient, for this purpose, with this consent state, is lawful, which is the entire question. Content guardrails and compliance enforcement are complements operating at different layers, and mistaking one for the other is how an organisation ends up confidently non-compliant.

4.3 The temporal failure of the audit model

The third existing answer is not a tool but a practice: the post-deployment compliance audit. Audits and compliance review cycles are built around a cadence, monthly or quarterly, that made sense when the systems being reviewed were humans making discrete decisions. An autonomous agent making thousands of micro-decisions per hour operates on a completely different clock. A compliance team that reviews AI lending decisions quarterly cannot intercept a transaction that violates CBN policy the moment it is initiated; by the time the audit happens, the transaction has already settled, the data has already crossed the border, the disclosure has already been made. This is the core argument for what security engineering calls “shifting left”: governance cannot be layered on top of a system after deployment. It has to be embedded in the execution layer itself, evaluated before the action happens, not after. The audit does not become useless in this model; it becomes verifiable, because a runtime layer that records every decision gives the auditor evidence instead of reconstruction.

Taken together, the survey yields a precise specification of what is missing. The world has execution-layer enforcement machinery without African regulatory content (the AGT category), content validation without action-level compliance semantics (the guardrails category), and compliance practice without runtime speed (the audit model). What none of the three provides, and what Section 5 describes, is the intersection: pre-execution enforcement, carrying maintained African statutory logic, evaluated at the moment of action.

SECTION 05

Section 5: Runtime Enforcement: What the Solution Looks Like

Runtime enforcement is a specific technical claim: that compliance evaluation should happen at the moment an AI agent attempts an action, a tool call, an API request, a data access, and before that action executes, rather than after the fact. This is the pre-execution interception model, and it stands in direct contrast to the post-deployment audit model that dominates compliance practice today.

5.1 The interception model

In practice, this means placing a policy evaluation step between an agent's decision and the tool or API call it is about to execute. The agent proposes an action, for example a funds transfer, a NIN lookup, or a data export; the policy layer evaluates that action's parameters against the applicable jurisdiction- and

sector-specific rules before the underlying tool call is allowed to run. A working instance of this pattern already exists: Comply54, an open-source runtime governance layer, ships this exact interception point as a single node that can be added to an existing agent graph. Dropped into a LangGraph pipeline, for instance, the compliance check becomes one additional node with conditional routing back to the agent or forward to the tool, requiring no rewrite of the agent's existing logic and no separate policy-engine binary to install.[26]

Each evaluated action resolves to one of a small number of outcomes: allow, where the action proceeds unmodified; block, where the action is halted before execution; or escalate, where the action is paused pending human review. This matters specifically for regulated industries, where a binary allow-or-deny model is too blunt. Many real compliance obligations, including the human-review right Kenya's pending AI Bill would create, require a human in the loop rather than an outright refusal. What distinguishes a genuine compliance decision from a vague warning is specificity: a useful implementation does not simply say a transaction was blocked, it returns the exact regulation, the specific section, and the rule that fired. This is precisely what would have caught Section 3's first scenario. Faced with the same ₦5,000,000 transfer on a Tier 1 account, a policy-aware layer blocks the action and returns the reason: the transfer exceeds the ₦50,000 single-transaction ceiling that the CBN's tiered Know-Your-Customer framework attaches to a Tier 1 account, cited to the specific rule, rather than a generic compliance error.

5.2 Proof of enforcement, not just enforcement

A requirement that becomes obvious once regulators start asking questions is that blocking a violation is only half the job; the other half is proving, after the fact, that the check actually ran. This is why a compliance layer should attach a tamper-evident record to every decision. Comply54 does this by emitting, on request, a compliance certificate: a structured record of the action evaluated, the decision reached, and the specific rule that fired, carrying a SHA-256 integrity hash so that any later alteration of the record is detectable.[27]

For an organisation answering to the NDPC or the CBN after an incident, a certificate of this kind converts “we believe our systems were compliant” into a specific artifact tied to the exact transaction under review, produced at the moment of the decision rather than reconstructed afterwards. Stronger cryptographic guarantees are possible in principle, for example signing each record with a private key the deploying organisation holds itself, so that a regulator can verify authorship offline without trusting the operator, and this is a natural direction for the tamper-evidence model to develop; the essential point for this paper is that the evidentiary artifact is generated as a by-product of enforcement, not assembled later from logs.

[26] comply54, open-source project (github.com/comply54/comply54), Apache 2.0 licence; distributed via PyPI (comply54) and npm (@comply54/core). Policy evaluation is performed in-process via the regopy Rego interpreter, with no OPA binary required.
[27] comply54 emits a compliance certificate carrying a SHA-256 integrity hash for each evaluated decision (see the certificate() interface in the project documentation).

5.3 Framework agnosticism

A further design requirement is framework agnosticism. Any governance layer that only works with one agent orchestration framework is not infrastructure, it is a plugin, tied to the fortunes of whichever framework it was built for. A governance layer intended to serve as genuine infrastructure needs adapters across the frameworks actually in production use today, LangChain, LangGraph, CrewAI, and AutoGen chief among them, so that the compliance guarantee travels with the deployment rather than the tooling choice. This is achievable without excessive engineering overhead: a governance layer built as an installable package, distributed via PyPI for Python agents or npm for TypeScript and Node agents, with thin adapters per framework, can typically be added to an existing agent in under an hour.

5.4 Offline enforcement

A final requirement is specific to the African context: offline enforcement. Data sovereignty concerns across the NDPA, POPIA, and comparable statutes mean that a compliance decision should never require an external API call, particularly one that might route sensitive financial or health data outside the organisation's own environment merely to check whether an action is permitted. A compliance layer that phones home to a third-party service to answer “is this allowed?” reproduces the very data-sovereignty risk the underlying law was written to prevent. This argues for policy evaluation that runs entirely inside the deploying organisation's own infrastructure, with no dependency on an external service and no separate policy-engine binary to install and maintain.

5.5 A working reference implementation

Comply54 is offered throughout this paper as a concrete, working instance of this model, not as the only possible implementation, but as evidence that the infrastructure described above is buildable today rather than a future research problem. As of this writing, the open-source project ships policy packs across twelve African jurisdictions, organised into two layers: individual regulatory packs (covering, among others, Nigeria's NDPA, CBN tiered-KYC, BVN/NIN, NFIU AML, National Health Act, and NAICOM rules; Kenya's KDPA; South Africa's POPIA; and data-protection statutes for Ghana, Rwanda, Egypt, Ethiopia, Mauritius, Tanzania, and Uganda), plus a set of universal agent-safety packs mapped to the OWASP Agentic AI risks. These are composed into ready-to-use sector classes, such as Nigeria fintech, healthcare, and insurance bundles, a Kenya fintech bundle, and a Pan-African fintech bundle spanning ten markets, so that a single import wires up every framework relevant to a given deployment. Coverage depth varies by jurisdiction, with the Nigerian packs the most detailed. Policy logic is evaluated in-process through the regopy Rego interpreter, with no OPA binary or subprocess required, which is what allows it to run in serverless environments and fully offline. It is released under the Apache 2.0 licence, distributed via PyPI and npm, and its source is publicly auditable on GitHub, which matters for a compliance tool specifically: an organisation adopting it can verify for itself that the policy logic matches the statutory text it claims to enforce, rather than taking a vendor's word for it. Full technical

documentation, including the sector-specific policy packs and framework integration guides referenced above, is available at comply54.io.

SECTION 06

Section 6: Implications for Builders, Fintechs, and Regulators

The preceding sections established a factual claim: African AI regulation is real and enforced, deployed agents cannot see it, existing tooling does not close the gap, and runtime enforcement can. This section translates that claim into consequences for the three groups who can act on it.

6.1 For AI developers: compliance-first architecture

Compliance-first architecture means treating jurisdiction-awareness as a first-class input to agent design, in the same category as authentication or rate limiting, not as a checklist item addressed after the system is already in staging. In practice this changes three habits. First, the policy check belongs in the design diagram from day one, as a named component between the agent's decision step and its tool execution, so that every new tool added to the agent inherits governance automatically rather than requiring a separate review. Second, regulatory context becomes part of the action's data model: an agent that will move money needs the customer's KYC tier in scope at the moment of the transfer call, and an agent that will touch personal data needs the consent state and destination jurisdiction in scope at the moment of the export call, because a policy engine can only evaluate what it can see. Third, the compliance decision itself becomes a first-class output, logged, citable, and attached to the action it governed, so that the system accumulates evidence as a by-product of operating.

The economic argument for doing this early rather than late is straightforward. Retrofitting governance after deployment requires reverse-engineering every tool call an agent has ever been permitted to make and reasoning, after the fact, about which of them should have been blocked, a forensic exercise conducted under regulatory pressure. Building it in from the start is one additional node in an agent graph and a policy pack import. The asymmetry compounds with scale: every month of ungoverned operation adds to the population of decisions that may later need to be explained.

One habit deserves specific warning, because Section 3 showed why it fails: compliance instructions in the system prompt are not a control. They are a request to a probabilistic system, and the appropriate role for them is as a complement to deterministic enforcement, never a substitute. A developer who can answer the question “what, other than the model's judgment, prevents this agent from executing a prohibited action?” with a specific component has a governed agent. A developer who cannot has a liability with good intentions.

6.2 For fintechs: the obligation stack is already in force

For Nigerian fintechs specifically, the exposure described in this paper is not prospective. It is an accumulation of obligations that are individually in force today, and it is worth setting them out as the stack a deployed AI agent actually sits inside.

The GAID's DPIA mandate has applied since 19 September 2025: any deployment of AI over personal data requires a Data Protection Impact Assessment signed by an accredited Data Protection Officer, following the GAID's Schedule 4 template, submitted within four months for software processing sensitive personal data, and folded into the organisation's annual Compliance Audit Return. Organisations classified as Data Controllers or Processors of Major Importance (DCPMI), a category that includes fintechs and payment providers by design, carry registration obligations and, at the higher tiers, annual audit filings. The FCCPC's 2025 digital lending regulations separately require qualified digital lenders to register and explicitly target opaque algorithmic lending, requiring transparency in how rates are set. The CBN's Open Banking framework classifies credit-scoring data as high and sensitive risk, with consent-management requirements attached. And the NIMC Act 2026 attaches criminal penalties to unauthorised handling of NIN-linked data: a minimum ₦10 million fine or five years' imprisonment for individuals, a minimum ₦20 million fine for corporate entities, and personal liability reaching responsible executives, which for a fintech means the officers who approved the agent deployment.

The enforcement climate around this stack is active, not theoretical. Beyond the individual actions detailed in Section 2, the NDPC announced in August 2025 a sector-wide investigation into more than 1,300 organisations across banking, insurance, pensions, and gaming, demanding audit reports, evidence of technical safeguards, DPO appointments, and DCPMI registration within 21 days.[28]

An organisation whose AI deployment cannot produce, on demand, a documented DPIA and a record of what its agents have been permitted and refused to do is not in a defensible position when that letter arrives. Runtime enforcement does not merely reduce the probability of a violation; it converts the response to a regulator's inquiry from reconstruction into retrieval.

The practical sequencing for a fintech already running agents in production is therefore: inventory every action an agent can currently take against the obligation stack above; complete and file the DPIAs the GAID already requires; and place a policy evaluation layer in front of the highest-risk actions first (transfers, data exports, and identity lookups), where a single ungoverned call can constitute a criminal offence under the NIMC Act or a fined breach under the NDPA.

6.3 For regulators and policymakers: from post-incident audit to deployment standard

[28]Nigeria Data Protection Commission, 2025 sector-wide compliance investigation reportedly covering more than 1,300 organisations across banking, insurance, pensions, and gaming, with a 21-day response requirement.

The recommendation this paper puts to regulators is specific: make runtime, pre-execution compliance evaluation an expected, documented characteristic of AI systems deployed in regulated sectors, rather than relying primarily on post-incident audit and investigation. Three concrete openings exist to do this.

First, within Nigeria, the GAID already points in this direction. Its emerging-technology provisions require organisations deploying AI to test in low-risk environments, evaluate the likelihood of discriminatory outcomes, carry out periodic re-testing, and establish structures for continuous monitoring of the technology. A runtime enforcement layer is the natural, and arguably the only scalable, implementation of a continuous monitoring structure for an autonomous agent: guidance from the NDPC clarifying that continuous monitoring of agentic AI systems is expected to operate at the point of action would convert an existing obligation into an unambiguous deployment standard without requiring new legislation.

Second, at the continental level, the AU's implementation calendar creates a specific window. Phase 1 of the Continental AI Strategy concludes in 2026, a formal review is scheduled for 2027, and Phase 2 begins in 2028. The 2027 review should assess not only how many member states have adopted strategies, the metric Phase 1 was built around, but whether any enforcement infrastructure exists to make those strategies operative; and Phase 2's deployment guidance should specify pre-execution compliance evaluation as a documented standard for AI systems in regulated sectors, against which conformity can be assessed. A standard of this kind is deliberately technology-neutral: it names a property the system must have, not a vendor or product that must supply it.

Third, regulators should require proof of enforcement, not just enforcement. An assertion that a compliance layer exists is unauditable; a tamper-evident record of each evaluation, verifiable offline by the regulator without trusting the operator, is not. Specifying evidence of this kind in supervisory expectations would give examiners a concrete artifact to request, and would reward exactly the organisations that have invested in doing this properly.

South Africa's withdrawn draft policy is itself a cautionary argument for the whole recommendation: a governance document about AI, produced without verification infrastructure, was undone by unverified AI output. Verification at the point of production would have caught the fabricated citations before publication, just as verification at the point of execution catches a non-compliant agent action before settlement. The lesson generalises, and the continent that internalises it first will be the one whose AI deployments regulators elsewhere learn to trust.

SECTION 07

Section 7: The Way Forward

Two futures are available from here. In the first, African AI agents continue to be deployed without jurisdiction awareness, and regulatory risk compounds quietly as enforcement activity intensifies, each

new fine, each new AI bill, widening the gap between what agents are doing and what they are permitted to do, until the exposure surfaces as a crisis for some institution that assumed compliance was someone else's problem.

In the second, African engineers build the governance infrastructure that makes compliance a structural property of the system rather than an optional afterthought, and in doing so, produce something exportable: a model for encoding jurisdiction-specific regulatory obligation into enforceable, machine-readable policy that the rest of the Global South, facing an identical mismatch between fast AI adoption and fast-moving regulation, can adapt rather than reinvent.

This is deliberately framed as work that cannot be outsourced. Encoding Nigerian, Kenyan, or South African law into policy that an AI agent can correctly interpret requires people who understand both the statutory text and the systems that need to obey it, a combination that exists, overwhelmingly, among African engineers building for African markets. The infrastructure gap this paper describes is also, in that sense, an opportunity that belongs specifically here.

> ***Encoding African law into policy an AI agent can correctly interpret is work that cannot be outsourced.***

REFERENCE

Appendix

A. Regulatory Reference Table

Jurisdiction	Key AI-Relevant Law(s)	Core Obligation
Nigeria	NDPA 2023; GAID 2025 (effective 19 Sep 2025); NIMC Act 2026 (26 Jun 2026); CBN/NFIU rules	DPIAs for AI/emerging tech; consent-first NIN data access; AML/KYC screening
Kenya	KDPA 2019; Artificial Intelligence Bill 2026 (Senate first reading 2 Apr 2026, pending)	Right to human review and explanation of automated decisions
South Africa	POPIA (s.71 automated decisions); Draft National AI Policy published 10 Apr 2026, withdrawn 26 Apr 2026	Right not to be subject to solely automated decisions; AI-specific policy pending redraft
Ghana	National AI Strategy (2022; Cabinet re-approval Feb 2026)	Strategic framework; no binding AI legislation yet
Rwanda	National AI Policy (2023)	Strategic framework; no binding AI legislation yet

Egypt	National AI Strategy (2019); Charter for Responsible AI (2023)	Strategic and ethical framework; AI Centre of Excellence
Mauritius	National AI Strategy (2018, Africa's first)	Strategic framework; no binding AI legislation yet
Ethiopia	National AI Policy (2024)	Strategic framework administered by Ethiopian AI Institute
Pan-African	AU Continental AI Strategy (endorsed Jul 2024; Phase 1, 2025-2026)	Governance-framework and national-strategy coordination across AU member states

Note: this table reflects the regulatory status of each instrument as of July 2026. Kenya's AI Bill and South Africa's national AI policy were both in active development at time of writing and are the entries most likely to have since changed.

B. Technical Glossary

- Runtime enforcement: evaluating an AI agent's proposed action against policy before that action executes, rather than reviewing it after the fact.
- Policy-as-code: expressing regulatory or organisational rules in a machine-evaluable format (e.g. Rego) so they can be checked programmatically.
- Pre-execution interception: the point in an agent's tool-calling chain where a proposed action is paused for policy evaluation before being sent to the underlying API or system.
- Allow, Block, Escalate, Audit: the four possible outcomes of a policy evaluation: permit, halt, route to human review, or permit while logging for compliance record.
- Framework agnosticism: a governance layer's ability to operate consistently across multiple agent orchestration frameworks (LangChain, LangGraph, CrewAI, AutoGen, etc.) rather than being tied to one.
- Offline enforcement: performing policy evaluation locally, without requiring an external network call, to avoid exposing sensitive data during the compliance check itself.
- DPIA (Data Protection Impact Assessment): a documented assessment of privacy risk, required under the GAID 2025 for AI and other emerging-technology deployments.

C. Architecture Overview: The Runtime Enforcement Model

At a high level, the runtime enforcement model inserts a policy evaluation layer between an agent's decision-making step and the tool or API call that would carry out that decision. The agent proposes an action; the policy engine evaluates that action against the relevant jurisdiction-specific policy pack; the engine returns one of the four outcomes; and only on an Allow, a cleared Escalate, or an Audit does the original tool call proceed. A Block terminates the action before it ever reaches the tool. Every evaluation can emit a tamper-evident compliance certificate, and the entire layer, including the policy packs, runs inside the deploying organisation's own environment, so no sensitive data leaves for a compliance check. The diagram below shows the full flow.

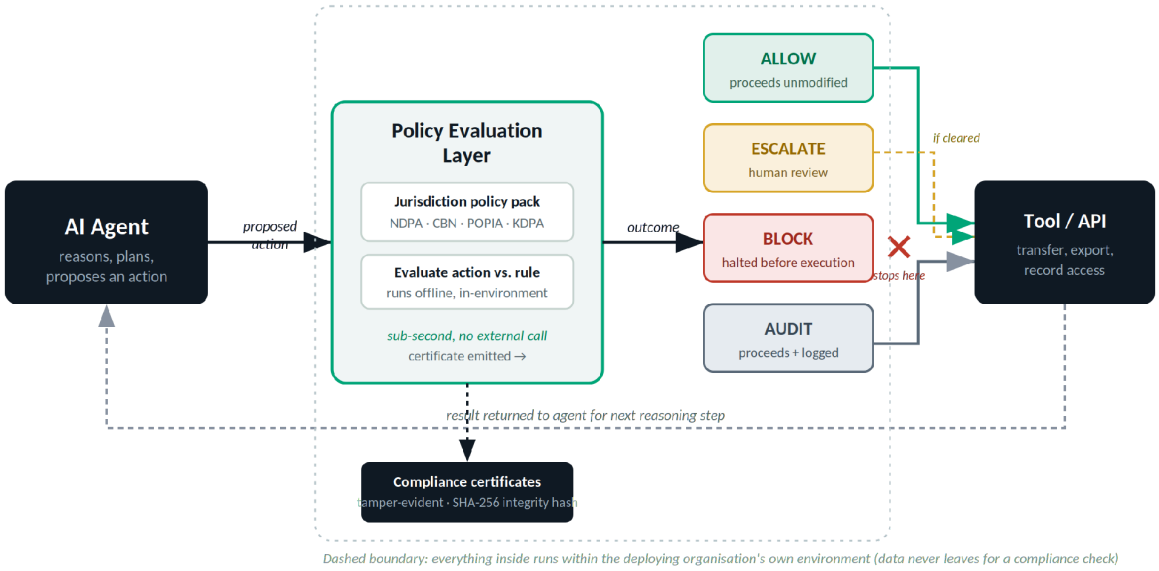

Figure 1. The runtime enforcement model. A proposed agent action is evaluated against a jurisdiction-specific policy pack before execution and resolved to one of four outcomes; only Allow, cleared Escalate, and Audit reach the tool, while Block halts the action. Every decision can emit a tamper-evident compliance certificate, and the whole layer runs inside the organisation's own environment.

End of document