PARLIAMENT
OF THE REPUBLIC OF SOUTH AFRICA



# A CONSTITUTIONAL, DEVELOPMENTAL & AUDITABLE ARTIFICIAL INTELLIGENCE LEGISLATIVE FRAMEWORK FOR SOUTH AFRICA

## PORTFOLIO COMMITTEE ON COMMUNICATIONS AND DIGITAL TECHNOLOGIES POSITION PAPER

JUNE 2026

PARLIAMENTARY RESEARCH UNIT (PRU)
WWW.PARLIAMENT.GOV.ZA

# A CONSTITUTIONAL, DEVELOPMENTAL & AUDITABLE ARTIFICIAL INTELLIGENCE LEGISLATIVE FRAMEWORK FOR SOUTH AFRICA

Table of Contents


1. EXECUTIVE SUMMARY 3
2. INTRODUCTION 3
3. THE GLOBAL STATE OF ARTIFICIAL INTELLIGENCE GOVERNANCE 4
4. THE COMPARATIVE ANALYSIS OF EXISTING AI GOVERNANCE FRAMEWORKS 5
The European Union (EU) AI Act 5
China 6
The Brazil AI Bill 6
Canada: AI Governance Maturity 7
Germany 7
Kenya: National AI Strategy (2025-2030) 7
Namibia 8
Comparative Finding: The Reasoning Governance Gap Across Countries 8
5. AI SOVEREIGNTY AND ECOSYSTEM DEVELOPMENT 9
6. THE REASONING GOVERNANCE GAP: WHY CURRENT AI LEGISLATION GOVERNS THE SYSTEM BUT NOT THE REASONING PROCESS 9
6.1 Gaps in the existing legislation 9
6.2 Gap analysis 10
6.3 South Africa's Withdrawn Draft Policy 10
7. THE PROPOSED LEGISLATIVE ARCHITECTURE: A SOUTH AFRICAN ARTIFICIAL INTELLIGENCE ACT 11
8. THE INSTITUTIONAL FRAMEWORK FOR AI GOVERNANCE 13
9. KEY CONSIDERATIONS FOR THE COMMITTEE 14
10. CONCLUDING REMARKS 15
11. REFERENCES 16

## 1. EXECUTIVE SUMMARY

Artificial Intelligence (AI) is one of the most consequential technologies of the 21st century, transforming governance, labour markets, healthcare, education, and democratic processes globally. In the South African context, AI presents a dual imperative: leveraging its potential for developmental advancement while addressing significant risks, including increased inequality, mass surveillance, algorithmic discrimination, misinformation, and threats to democracy.

According to a 2026 study by Comparitech, only 33 of 178 countries have enacted AI legislation, and 27 of these are European Union member states governed by the EU AI Act (Regulation 2024/1689). Denmark, France, and Greece scored highest, each receiving 13 out of 14 by implementing additional national protections alongside the EU AI Act. Beyond Europe, South Korea has become the second country to enact comprehensive AI legislation, consolidating 19 separate bills into a single Framework Act.

A critical review of major AI legislative frameworks, including the EU AI Act, South Korea's AI Basic Act, China's Generative AI Measures, Brazil's Draft Bill, and emerging models in Canada, Germany, Italy, Kenya, and Namibia, identifies a universal gap in governing reasoning. Existing laws primarily regulate the external environment of AI, such as risk management, transparency, and documentation, but do not address the internal reasoning processes that generate outputs. As a result, current frameworks cannot ensure source traceability, reproducibility, or logical grounding, all of which are essential to constitutional accountability under Sections 33 (just administrative action) and 32 (access to information) of South Africa's Constitution.

Building on the Kuala Lumpur Declaration on Responsible AI, adopted by over 200 parliamentarians from 65 countries and advocating traceability, human oversight, democratic accountability, and independent audit, this paper proposes a five-layer governance model. The fifth layer, focused on process governance, represents a potential South African contribution to global AI governance.

The paper recommends enacting a dedicated South African AI Act and establishing the National AI Commission to track high-impact AI systems, establish AI certification frameworks, create regulatory sandboxes, and enhance parliamentary oversight. Through this approach, South Africa can position itself not only as a consumer of international AI governance frameworks but also as a contributor to the next generation of global AI regulation.

## 2. INTRODUCTION

Artificial Intelligence (AI) has evolved from a specialised technological domain into a foundational technology that influences nearly every sector of society. Machine learning, large language models, generative AI, computer vision, and autonomous systems have accelerated global AI adoption and fundamentally transformed how governments, businesses, and citizens interact with digital technologies.

The World Economic Forum (WEF), Organisation for Economic Co-operation and Development (OECD), United Nations Educational, Scientific and Cultural Organisation (UNESCO), African Union (AU), and Inter-Parliamentary Union (IPU) have identified AI as a transformative technology with the potential to drive economic development, enhance public service delivery, advance scientific research, and improve societal wellbeing. However, these organisations caution that insufficient governance could result in significant harms, including algorithmic discrimination, erosion of privacy, the dissemination of misinformation, the concentration of economic power, the manipulation of democratic processes, and increased inequality[1].

[1] UNESCO, (2021); IPU, (2024); OECD, (2019)

In the South African context, AI presents both significant opportunities and risks. AI can accelerate economic growth, improve public service delivery, strengthen healthcare and education systems, enhance agricultural productivity, support industrialisation, and facilitate digital transformation. The Presidential Commission on the Fourth Industrial Revolution (PC4IR) has recognised AI's substantial potential to address developmental challenges while fostering innovation, competitiveness, and inclusive economic growth. Conversely, persistent challenges in South Africa, including inequality, unemployment, poverty, digital exclusion, and uneven infrastructure, create distinct vulnerabilities that require targeted governance and regulatory frameworks.

The Kuala Lumpur Declaration on Responsible AI, adopted at an international parliamentary conference convened by the IPU, the Parliament of Malaysia, the Commonwealth Parliamentary Association (CPA), and the United Nations Development Programme (UNDP), frames AI as a governance challenge that requires democratic oversight, ethical leadership, and international cooperation. The Declaration cautions that AI systems may reshape economies, labour markets, governance structures, education, healthcare, and democratic institutions, resulting in both positive and negative outcomes.

This position paper pursues three primary objectives. First, it assesses the current international landscape of AI governance. Second, it critically evaluates South Africa's policy trajectory, including the withdrawn Draft National Artificial Intelligence Policy and emerging parliamentary perspectives on AI governance. Third, it proposes a legislative framework that balances innovation, rights protection, democratic accountability, and developmental objectives, while introducing innovative approaches to AI governance.

## 3. THE GLOBAL STATE OF ARTIFICIAL INTELLIGENCE GOVERNANCE

AI governance represents a significant policy challenge for governments worldwide. Although the transformative potential of AI is widely acknowledged, regulatory responses remain fragmented and inconsistent.

A 2026 study conducted by Comparitech evaluated 178 countries across 14 metrics, including the adoption of risk-based approaches, the existence of regulatory bodies, copyright disclosure requirements, transparency obligations, and penalties for non-compliance. The findings indicate that only 33 countries are subject to AI legislation, with 27 of these being European Union member states governed by the EU AI Act [2].

**Table 1: Global AI Legislative Coverage[3]**

| Region | Existing AI Law | Primary Model | Focus |
|---|---|---|---|
| European Union | 27 Member States (EU AI Act) | Risk-based (4 tiers) | Fundamental rights, market harmonization |
| East Asia | South Korea (AI Basic Act) | High Impact and Innovation | National competitiveness, trust |
| North America | None (Sectoral only) | Executive orders, voluntary | Innovation, limited rights |
| Africa | None | Non-binding | Developmental, ethical principles |

[2] Comparitech, (2026)
[3] OECD.AI Policy Observatory (2026)

| | | | |
|---|---|---|---|
| | | strategies | |

Denmark, France, and Greece have the highest-scoring national frameworks, each achieving 13 out of 14. These countries implement additional national protections, such as enhanced transparency requirements, public-sector AI registers, and strengthened data protection, thereby supplementing the EU AI Act framework.

Outside Europe, South Korea has emerged as the second country in the world to enact AI legislation. The Framework Act on the Development of Artificial Intelligence and Establishment of Trust (passed 26 December 2024, effective 22 January 2026) consolidated 19 separate bills proposed in the National Assembly over four years of debate. This aligns with Comparitech's principal finding that countries with AI legislation show significantly higher public trust in AI deployment. Nevertheless, no current law requires algorithmic reproducibility or source traceability for generative outputs, highlighting a persistent governance gap.

The Kuala Lumpur Declaration supports this conclusion, observing that technological advancement is outpacing the capacity of governments, institutions, and societies to adapt effectively. The declaration advocates for international cooperation to promote safety, accountability, interoperability, and responsible innovation.

## 4. THE COMPARATIVE ANALYSIS OF EXISTING AI GOVERNANCE FRAMEWORKS

This analysis evaluates the principal characteristics of the EU AI Act, South Korea's AI Basic Act, China's Generative AI Measures, Brazil's Draft Bill, and emerging models in Canada, Germany, Italy, Kenya, and Namibia.

### The European Union (EU) AI Act[4]

The EU AI Act introduces the first legally binding risk taxonomy for artificial intelligence systems:

1. **Unacceptable risk**: Prohibited practices including social scoring by public authorities, biometric mass surveillance without a warrant, AI manipulating vulnerable groups, and real-time facial recognition in public spaces.
2. **High-risk**: Systems affecting healthcare, education, employment, critical infrastructure, law enforcement, and democratic processes are subject to conformity assessments and ongoing monitoring.
3. **Limited risk**: Transparency obligations for chatbots and synthetic content.
4. **Minimal risk**: Voluntary compliance.

A significant gap in the EU AI Act concerns the regulation of reasoning processes. Although the Act provides a detailed framework for process documentation, including technical documentation, record-keeping, and conformity assessments, it does not address the verifiability of inferential logic. Compliance with technical standards, such as ISO/IEC 42001:2023, does not ensure that AI outputs are grounded in verifiable fact. The ISO/IEC 42001 standard governs organisational processes for AI systems, including establishing, implementing, maintaining, and improving an AI management system. However, it does not regulate the generative or inferential mechanisms themselves[5].

[4] European Union, (2024)
[5] ISO, (2023)

### South Korea's AI Basic Act

The Act on the Development of Artificial Intelligence and Establishment of Trust integrated 19 separate bills proposed in the National Assembly over four years of debate. Key provisions include [6]:

- National AI Committee to coordinate strategy across government departments.
- Designation of high-impact AI for systems significantly affecting human life, safety, or fundamental rights, including healthcare, public administration, employment, education, law enforcement, transportation, and financial decision-making.
- Explicit inclusion of generative AI as a distinct category requiring specific governance consideration.
- Obligations for transparency, safety, trustworthiness, and international cooperation

The Act is designed to build trust and improve national competitiveness, but it does not require that AI systems explain how they reach specific decisions after the fact. For example, an AI system used in public administration may make decisions that affect citizens' rights. While the Act requires documentation of how these systems are managed, it does not require a clear, independent explanation of how a particular decision was made[7].

### China

China became the first country to mandate security assessments and algorithm registration before market launch. The Cyberspace Administration of China (CAC) requires training data to be legal, compliant, and free from discrimination, and requires generated content to uphold socialist core values. Providers are prohibited from generating content that subverts state power, incites terrorism, or disrupts social stability [8].

China's approach demonstrates the feasibility of ex ante security assessments, a concept that could be adapted for high-risk AI in South Africa. However, the Chinese model lacks independent parliamentary oversight. As Sheehan notes, regulation is state-led and top-down, with no mechanisms for legislative review, public contestation, or judicial challenge to algorithm registration decisions. Despite security assessments, a governance gap persists, as the internal inferential processes of generative AI are not subject to independent audit or citizen contestation [9].

### The Brazil AI Bill

The Bill explicitly defines AI as a computational system that learns, perceives, interprets, and interacts with its environment using techniques including supervised learning, Bayesian inference, and logic-based systems. This Bill requires agents involved in AI development to identify and mitigate biases that conflict with national legislation, including anti-discrimination provisions. [10]

Figueira et al. (2024) identify three flaws[11]:

---

[6] Republic of South Korea, (2025)
[7] IAPP, (2025) & Stimson Centre, (2025)
[8] China, (2023)
[9] Sheehan, (2023)
[10] Brazil , (2020)
[11] Figueira, (2024)

- The Brazilian General Data Protection Law (LGPD) guarantees data subjects the right to review automated decisions and receive clear information regarding criteria and procedures. This allows AI providers to withhold information on the grounds of commercial and industrial secrecy, creating a direct legal contradiction.
- The Bill refers to future guidelines rather than establishing direct and fault-based liability, meaning no entity is held accountable for discriminatory or erroneous AI outputs.
- The Bill relies on voluntary codes of conduct rather than independent audits or regulatory enforcement.

### Canada: AI Governance Maturity[12]

A 2024 review of 84 AI governance initiatives in Canada revealed that while the country was the first nation to implement a national AI strategy, it faces persistent challenges, which include public mistrust of AI technologies, inadequate public engagement on major AI governance initiatives, accountability concerns in public funding, and difficulty retaining Canadian AI innovation and talent.

Daly mapped AI use in the Government of Canada and found seven use cases:

- Enhancing accessibility of public-facing resources,
- Creating models of natural and human activity,
- Assessing performance,
- Managing enforcement resources,
- Advising on eligibility,
- Triaging applications, and
- Assisting with or making eligibility decisions.

The Canadian experience further demonstrates that public trust is essential for AI adoption, a lesson reflected in the Kuala Lumpur Declaration.

### Germany

The German Bundestag has developed an AI Cloverleaf Model that categorises parliamentary AI applications into three overlapping areas: (1) legislative process (drafting, committee preparation, debate organisation); (2) support services (research, library, translation, technical infrastructure); and (3) internal administration (HR, procurement, IT, compliance). The model prioritises use cases that rely on centralised guidance, risk management, and human oversight, enabling scalability beyond individual pilot projects while ensuring compliance with the EU AI Act [13].

A use-case prioritisation framework enables parliaments to adopt AI incrementally, beginning with low-risk domains such as transcription and document search, and then progressing to higher-risk applications such as legislative drafting and budget analysis. The Cloverleaf model incorporates checks and balances at each stage, including ethics committees, oversight mechanisms, and risk management.

### Kenya: National AI Strategy (2025-2030)

Kenya's National AI Strategy (adopted in March 2025 and running through 2030) positions the country as Africa's leading AI hub for model innovation, advancing sustainable development, economic growth, and social inclusion. The strategy incorporates data sovereignty, a cybersecurity

[12] Daly, (2023)
[13] IPU, (2022)

framework, and ethical oversight as foundational principles. The Kenya Robotics and AI Society Bill (2023), currently before Parliament, proposes establishing a statutory body to promote responsible AI development, including expanding internet coverage, and raising public awareness[14] .

Kenyan Members of Parliament have publicly expressed concerns regarding AI-related risks, including privacy breaches, deepfakes, algorithmic discrimination, autonomous weapons, job displacement, economic inequality, social manipulation, and financial market manipulation. Several MPs have advocated for a dedicated legal framework beyond the current Bill, emphasising the need for the Ministry of Information, Communication, and the Digital Economy to lead the development of ethical guidelines and enforcement mechanisms [15].

### Namibia

Namibia launched its National AI Strategy in 2025, alongside three draft laws:

The Draft Artificial Intelligence Bill, the Data Protection Bill, and the Cybercrime Bill. An AI Readiness Assessment, launched in August 2025 in partnership with UNESCO, identified infrastructure gaps, skill shortages, and policy coordination challenges as impediments to the coherent deployment of AI. The Namibian Parliament has also pursued a multi-year e-Parliament Strategy (2022-2027), revamping its website, introducing live-streaming of sessions, and developing an online bill-tracking system[16].

### Comparative Finding: The Reasoning Governance Gap Across Countries

**Table 2: Comparative Assessment of Reasoning Governance**

| Country | Primary Governance Focus | Regulates Reasoning Traceability? | Mandates Independent Audit of Inference? |
|---|---|---|---|
| EU AI Act | Risk management, documentation, transparency | No | No |
| South Korea AI Basic Act | High-impact AI, trust, national strategy | No | No |
| China Generative AI Measures | Security assessments, algorithm registration | No (state audit only) | No (no independent mechanism) |
| Brazil Bill 21/2020 | Definitions, bias mitigation, self-regulation | No | No |
| Canada | Innovation, procurement, centralised governance | No | No |
| Germany | Use-case prioritisation, human oversight | No | No |
| Kenya | National strategy, Robotics | No | No |

[14] Kenya, (2025)
[15] Beigon, (2025)
[16] Xinhua, (2025)

| | | | |
|---|---|---|---|
| | & AI Society Bill | | |
| Namibia | Legislative groundwork, readiness assessment | No | No |

A common limitation across the frameworks reviewed is that they emphasise governing inputs, risks, and outputs more than governing the inferential processes through which AI systems reach specific conclusions.

The Kuala Lumpur Declaration on Responsible AI points out this gap. It says AI can change democracies in both positive and negative ways. It stresses that AI decisions affecting people should be open to challenge, review, and correction. It also calls for clear tracking, compelling evidence, and independent checks of how AI decisions are made ideas that match the Committee's proposed focus on process oversight and making AI reasoning more transparent[17]

## 5. AI SOVEREIGNTY AND ECOSYSTEM DEVELOPMENT[18]

Although this paper addresses governance and accountability, the governance gap cannot be examined apart from broader issues concerning a nation's AI capabilities. The Naspers and Prosus Africa's AI Imperative report states that effective AI governance requires a supporting ecosystem, including digital infrastructure, compute capacity, quality datasets, skilled personnel, sustainable financing, and reliable data governance arrangements. These enabling factors determine a country's ability to regulate AI and actively participate in its development and deployment.

## 6. THE REASONING GOVERNANCE GAP: WHY CURRENT AI LEGISLATION GOVERNS THE SYSTEM BUT NOT THE REASONING PROCESS

The reasoning governance gap is the absence of regulatory mechanisms to independently verify, reproduce, trace to source, or audit the inferential process through which an AI system arrives at a conclusion, recommendation, or decision.

### 6.1 Gaps in the existing legislation

The current legislation focuses on three domains:

- **Inputs:** Data quality, provenance, privacy compliance (POPIA).
- **Environment:** Risk management systems (ISO/IEC 42001), human oversight, documentation.
- **Outputs:** Transparency labels, error rates, post-market monitoring.

None of the frameworks regulates inference integrity – the logical step from data to decision. For generative AI, this manifests as hallucinations (factually incorrect yet confident outputs). For discriminative AI, it manifests as spurious correlations.

### 6.2 Gap analysis

[17] Kuala Lumpur Declaration on Responsible AI, (2025)
[18] Naspers/Prosus Report on AI (2026)

A 2025 study by the Algorithmic Justice League found that there is a high percentage of high-risk AI deployments (healthcare, hiring, credit scoring) for which developers could not explain why a specific output was generated19. Within the framework of South Africa's Constitution, the divergence between reasoning and governance presents a significant challenge. Section 33 of the Constitution requires administrative action to be lawful, reasonable, and procedurally fair, while the Promotion of Administrative Justice Act (PAJA) requires decisionmakers to provide adequate reasons for their administrative decisions. These legal obligations presuppose that decision-makers can identify, evaluate, and articulate the rationale underlying a decision. When AI systems generate recommendations, assessments, or decisions through opaque reasoning processes, public institutions may struggle to demonstrate compliance with constitutional and administrative justice requirements. The issue extends beyond technical accuracy to encompass constitutional accountability. If a public official cannot explain the basis for an AI system's conclusion, meeting constitutional standards of transparency, rationality, and justification becomes difficult. Consequently, governance mechanisms must ensure AI-assisted decision-making remains subject to meaningful human oversight and constitutional review. An inability to interrogate or independently verify AI-assisted decisions may raise questions regarding compliance with constitutional standards of reasonableness, rationality, and procedural fairness [20].

The Kuala Lumpur Declaration reinforces this by stating that Parliaments should ensure that AI systems do not undermine the right to administrative justice, and that automated decisions affecting legal rights are subject to meaningful human review and independent audit.

### 6.3 South Africa's Withdrawn Draft Policy[21]

The gazette of the Draft National Artificial Intelligence Policy in April 2026 marked South Africa's initial substantive effort to create a national AI governance framework. The policy explicitly recognised AI as a strategic general-purpose technology and warned that the absence of a coherent national policy could exacerbate inequality, widen digital divides, and reduce national competitiveness.

Shortly after its publication, the policy became the focus of considerable controversy. The Department of Communications and Digital Technologies (DCDT) later informed Parliament that it had withdrawn the policy after discovering that generative AI tools had been used during drafting and that certain references in the document were fictitious or unverifiable.

The withdrawal is significant for three primary reasons. First, it underscored the governance challenges that the policy aimed to address. Second, it demonstrated the necessity of verification, traceability, and accountability in public-sector decision-making processes involving AI systems. Third, it reinforced concerns about excessive reliance on AI-generated content without sufficient human validation.

The paper then gives a practical example of the increasing need for governance measures which can verify the outputs of AI and stresses the importance of tackling the governance gap concerning reasoning by means of legislation.

## 7. THE PROPOSED LEGISLATIVE ARCHITECTURE: A SOUTH AFRICAN ARTIFICIAL INTELLIGENCE ACT

[19] Algorithmic Justice League, (2025)
[20] Constitution of the Republic of South Africa, (1996)
[21] Department of Communications and Digital Technologies, (2026)

This paper proposes a 5-layer governance model for AI. The first four layers draw on leading global standards, including the EU AI Act, South Korea's AI Basic Act, and the ISO/IEC 42001 management standard. The fifth layer introduces South Africa's distinctive approach, addressing gaps in current international practices by tackling the ***reasoning-governance gap*** and meeting the Kuala Lumpur Declaration's requirements for traceability, evidence-based decision-making, and independent audits.

### Layer 1: Protection of Constitutional Guarantees

Layer 1 sets clear legal boundaries for unacceptable AI use. The paper recommends that the following practices be explicitly banned in future legislation [22]:

- Bias
- Unlawful biometric mass surveillance without judicial warrant.
- AI systems designed to manipulate vulnerable populations (children, persons with disabilities, economically disadvantaged groups).
- AI systems that intentionally facilitate discrimination on grounds prohibited by Section 9 of the Constitution.
- AI systems intended to undermine democratic processes, including electoral manipulation.
- Real-time facial recognition in public spaces except for narrowly defined serious crime prevention with judicial authorisation.

These prohibitions align with the EU AI Act and Parliament should revisit them regularly to keep pace with technological advances.

### Layer 2: Risk-Based Regulation

Layer 2 uses a risk-based approach, similar to the EU AI Act and South Korea's AI Basic Act.

- High-Risk Systems (Require certification): These include applications that impact healthcare, education admissions, job hiring and firing, financial credit decisions, criminal justice, public benefit eligibility, elections, and critical infrastructure.
- Limited-Risk Systems (Require disclosure): Examples are chatbots, synthetic content generators, and emotion recognition tools (when not used in high-risk settings). These must be clearly labelled and users must be notified.
- Minimal-Risk Systems (Voluntary compliance): Things like spam filters, video game AI, and simple automation fall in this category and are only subject to voluntary best practices.

This way, the rules match the level of risk and avoid over-regulating low-impact or innovative uses of AI.

### Layer 3: Organisational Accountability and Compliance

Layer 3 outlines what organisations that develop, use, or manage AI systems need to do under the ISO/IEC 42001:2023 standard.

Organisations deploying high-risk AI must implement [23]:

[22] Republic of South Africa, (2020)

- AI risk management systems with ongoing monitoring.
- AI governance frameworks with designated responsible AI Officers.
- Human oversight mechanisms at critical decision points.
- Incident reporting procedures to the National AI Commission.
- Internal audit mechanisms for AI system outputs.
- Data governance policies complying with POPIA.

Compliance requirements should scale to the organisation's size and risk level, so smaller companies and startups are not burdened by unnecessary red tape.

### Layer 4: Transparency and Human Oversight [24]

Layer 4 boosts transparency requirements, building on ideas from the EU AI Act and the Kuala Lumpur Declaration.

Human oversight is essential in high-risk settings. AI should support, but not replace, human judgement, especially when decisions could affect people's rights, opportunities, or wellbeing. Fully automated decisions that have negative legal impacts are not allowed.

This layer directly aligns with the South African e-government strategy by providing the foundational shared digital infrastructure that breaks down departmental silos, enabling seamless, integrated, and citizen-centric public service delivery, as mandated by the Digital Transformation Roadmap and Operation Vulindlela.

### Layer 5: Process Governance and Reasoning Auditability

Layer 5 is the Committee's main legislative innovation and highlights South Africa's potential impact on global AI governance.

The paper argues that existing frameworks regulate inputs, governance structures, transparency obligations, and outputs, but provide limited oversight of the inferential processes through which AI systems generate conclusions, recommendations, or predictions. The steps that connect inputs to outputs [25]. South Africa should set clear legal rules for high-risk AI systems, wherever it is practical and makes sense:

- **Source traceability**: AI systems must show where their information comes from, using clear and reliable sources, especially in sensitive areas like healthcare, finance, and government.
- **Input traceability**: Important inputs that affect decisions must be recorded, time-stamped, and available for review.
- **Reproducibility**: Similar inputs should lead to similar results, or the system must explain why the results are different.
- **Evidentiary grounding**: For important decisions, the AI must provide a confidence level and show the evidence behind its conclusion.
- **Independent auditability:** Independent auditors must be able to check how decisions were made, without needing access to the system's confidential internal details

[23] ISO, (2023)
[24] UNESCO, (2021)
[25] World Bank, (2022)

Layer 5 puts Sections 32 (access to information) and 33 (just administrative action) of the Constitution into practice in today's algorithm-driven world. If someone is affected by an AI decision, they have the right to know what information the decision was based on and to have an independent auditor check it.

## 8. THE INSTITUTIONAL FRAMEWORK FOR AI GOVERNANCE

### The Institutional Architecture: A Coordinated, Adaptive Regulatory Model

Establishing an appropriate institutional framework for AI governance is one of the most consequential decisions confronting the South African legislature. While the proposed five-layer legislative model remains the cornerstone of this framework, the institutional mechanisms for its enforcement must be grounded in a pragmatic assessment of the South African state's institutional capacity, fiscal constraints, and existing regulatory infrastructure. This paper therefore proposes a fundamental recalibration of the institutional architecture, moving away from establishing a new, standalone National AI Commission in favour of a coordinated, adaptive regulatory regime that empowers and aligns existing institutions.

### The Institutional Landscape: A Comparative Analysis

An examination of international approaches reveals considerable variation in institutional design. The European Union's AI Act establishes a complex governance structure comprising a European AI Board, national competent authorities, and a market surveillance system. This multi-layered approach, while comprehensive, imposes significant administrative burdens on member states. South Korea's Framework Act on the Development of Artificial Intelligence and Establishment of Trust (2024) establishes a National AI Committee to coordinate strategy across government departments, supported by a dedicated secretariat. This model demonstrates a commitment to whole-of-government coordination but entails substantial institutional investment [26]. China's approach, by contrast, vests authority in the Cyberspace Administration of China (CAC), reflecting a centralised, state-led model that prioritises security and state control over independent oversight.[27]

These comparative cases illustrate a fundamental tension: between comprehensive institutional design and administrative feasibility. For South Africa, this tension is particularly acute, given fiscal constraints and the institutional challenges that characterise the developmental state. The Portfolio Committee on Communications and Digital Technologies has therefore adopted a pragmatic position, articulated by its Chairperson at the Naspers/Prosus AI Governance and Policy Symposium, that "the existing institutional architecture is sufficient and just needs to be better coordinated" to achieve effective AI governance.

### Proposed Institutional Model: AI Regulator's Forum

This paper therefore proposes a collaborative, sector-specific approach that leverages the existing regulatory infrastructure. The identified regulators and their AI-relevant competencies include:

#### *ICASA (Independent Communications Authority of South Africa):*

ICASA's mandate encompasses broadcasting, telecommunications, and online content regulation. The Authority is already grappling with AI-related challenges, including regulating AI in broadcasting and telecommunications, AI's influence on data privacy and misinformation,

[26] IAPP (2025); Stimson Centre (2025)
[27] Sheehan (2023)

and procurement transparency in AI-enabled systems. ICASA's existing competency in communications regulation provides a foundation for AI governance in these sectors.

***The Competition Commission:***

The Commission's mandate includes preventing anti-competitive conduct and promoting market competition. Its examination of the market concentration effects of AI is vital for preventing dominance by large technology platforms and ensuring barriers to entry for local AI innovators. The Commission's expertise in market analysis and competition economics is directly applicable to AI governance, particularly given AI markets' tendency toward concentration.

***The National Consumer Commission (NCC):***

The NCC's mandate includes consumer protection, including investigating unfair, unreasonable, or unjust commercial practices. The Commission is investigating how e-commerce platforms use algorithms to drive consumer engagement, expressing concern about data mining and the use of automated systems to influence consumer behaviour. Its consumer protection mandate aligns directly with protecting citizens from algorithmic harm.

***The Information Regulator:***

The Regulator's mandate under the Protection of Personal Information Act (POPIA) includes protecting personal information and overseeing automated decision-making. The Regulator has expressed concern about AI's impact on privacy rights and has asserted its competence to assess that impact. Its expertise in data protection and automated decision-making is directly relevant to AI governance.

This model aligns with international trends in AI institutional design. The United Kingdom, for example, has rejected a single AI regulator in favour of a "principles-based approach" where existing regulators adopt AI-specific responsibilities, balanced against their existing duties. This model, advocated by Lord Harper and others, is seen as more efficient and adaptable, allowing for innovation while mitigating risks [28]. The South African proposal mirrors this thinking, positioning the country to achieve effective AI governance while leveraging existing institutional capacity.

## 9. KEY CONSIDERATIONS FOR THE COMMITTEE

This institutional recalibration does not diminish Parliament's role in AI governance. The Portfolio Committee's responsibilities remain:

*Legislative:* Enacting the South African AI Act with its five-layer governance model, including constitutional protections, risk-based regulation, organisational accountability, transparency requirements, and process governance.

*Oversight:* Exercising oversight over the coordinated regulatory framework, including review of annual reports, budgetary oversight, and public hearings on AI governance issues.

[28] House of Lords (2024)

*Accountability:* Ensuring that responsibility for AI governance cannot be diffused into a "technological grey zone" but remains clearly allocated to accountable institutions.

A Joint Committee on AI could provide dedicated parliamentary oversight of this coordinated framework, with a mandate to review the regulators' forum's work, conduct inquiries into AI governance issues, and report to Parliament on the framework's effectiveness.

### Implementation Considerations

The transition to this coordinated model requires several implementation measures:

*First, legislative amendment:* The South African AI Act should include provisions establishing the AI Regulators Forum as the primary coordination mechanism for AI governance, with an expanded mandate and enhanced resources.

*Second, resource allocation:* Existing regulators require additional resources to develop AI expertise, including technical capacity, data science skills, and international engagement capabilities. These resources should be allocated through the budgetary process, with parliamentary oversight.

*Third, operational protocols:* The ICT Regulators Forum should develop operational protocols to coordinate AI governance, including information sharing, joint investigations, and consistent enforcement approaches.

*Fourth, international engagement:* The coordinated framework should include mechanisms for international engagement, ensuring that South Africa can participate in global AI governance discussions and learn from international experience.

## 10. CONCLUDING REMARKS

The rapid deployment of increasingly autonomous and generative AI systems is challenging traditional assumptions about accountability, transparency, and decision-making within democratic societies.

The analysis indicates that, although international initiatives have advanced risk management, transparency, and accountability in the governance of artificial intelligence, significant deficiencies persist. Among these is the persistent inability to consistently trace, reproduce, or independently verify the outputs and decision-making processes of AI systems. This governance gap, which is the absence of regulatory mechanisms to ensure the explainability and verifiability of AI-generated decisions, remains evident across all major frameworks reviewed.

South Africa commences its legislative engagement with artificial intelligence at a juncture when considerable international experience is available for critical reflection. This affords the country an opportunity to draw upon global successes and shortcomings, thereby enabling the development of a governance framework that aligns with its constitutional values and developmental priorities.

The Committee position paper maintains that South Africa should pursue a governance model for AI that is human-centred, innovation-enabling, and firmly rooted in constitutional and developmental imperatives. Rather than replicating existing legislative models, South Africa is well positioned to advance global discourse by introducing its own principles of process governance and reasoning accountability. The proposed Five-Layer Governance

Model, particularly through its emphasis on process and reasoning auditability, articulates a distinctively South African approach: no automated decision affecting an individual's rights should remain beyond independent scrutiny.

## 11. REFERENCES


African Union. 2024. *Continental Artificial Intelligence Strategy: Harnessing AI for Africa's Development and Prosperity*. Addis Ababa: African Union Commission.

Beigon, E. 2025. 'MP Kitany warns of AI risks in Kenya and calls for a national regulatory framework'. *Kenya Broadcasting Corporation*. Available at: https://www.kbc.co.ke (Accessed: 25 June 2026).

Brazil. Chamber of Deputies. 2020. *Bill No. 21/2020: Legal Framework for Artificial Intelligence in Brazil*. Brasília: Chamber of Deputies.

Comparitech. 2026. *AI Legislation: 178 Countries Ranked by How They Control and Safeguard AI Use*. Available at: https://www.comparitech.com (Accessed: 1 July 2026).

Daly, P. 2023. 'Mapping the use of artificial intelligence in the Government of Canada', *Revue Gouvernance / Governance Review*, 20(1), pp. 74-95.

Department of Communications and Digital Technologies (DCDT). 2026. *Draft National Artificial Intelligence Policy Framework*. Pretoria: Department of Communications and Digital Technologies.

European Union. 2024. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*, L 2024/1689.

Figueira, H.M.A., Garrote, M.G. and Zanatta, R.A.F. 2024. 'Regulating artificial intelligence in Brazil: The contributions of critical social theory to rethinking principles', *Data, Law and Decolonisation*, pp. 73-83.

House of Lords Communications and Digital Committee. 2024. *Large Language Models and Generative AI*. London: UK Parliament. Available at: https://committees.parliament.uk

International Association of Privacy Professionals (IAPP). 2025. *Analysing South Korea's Framework Act on the Development of Artificial Intelligence and Establishment of Trust*. Portsmouth, NH: IAPP.

Inter-Parliamentary Union (IPU). 2024. *Guidelines for AI in Parliaments*. Geneva: Inter-Parliamentary Union.

ISO/IEC. 2023. *ISO/IEC 42001:2023 Information Technology - Artificial Intelligence - Management System*. Geneva: International Organization for Standardization and International Electrotechnical Commission. ISBN: 978-92-671-1415-6.

Kenya Ministry of Information, Communications and the Digital Economy. 2025. *Kenya National Artificial Intelligence Strategy 2025-2030*. Nairobi: Government of Kenya.

Kuala Lumpur Declaration on Responsible Artificial Intelligence. 2025. *Kuala Lumpur Declaration on Responsible Artificial Intelligence*. Kuala Lumpur: Inter-Parliamentary Union, Commonwealth Parliamentary Association and United Nations Development Programme.

Naspers and Prosus. 2026. *Africa's AI Imperative: Insights Report*. Cape Town and Amsterdam: Naspers and Prosus.

OECD. 2019. *OECD Principles on Artificial Intelligence*. Paris: OECD Publishing. Available at: https://oecd.ai/en/ai-principles (Accessed: 4 July 2026).

OECD AI Policy Observatory. 2026. *National AI Policies and Strategies*. Available at: https://oecd.ai (Accessed: 14 July 2026).

Onyango, W. 2024. 'How people could shape the future of AI in Kenya: Using public participation to develop AI governance and literacy'. *Ada Lovelace Institute*. Available at: https://www.adalovelaceinstitute.org (Accessed: 7 June 2026).

Presidential Commission on the Fourth Industrial Revolution (PC4IR). 2020. *Report of the Presidential Commission on the Fourth Industrial Revolution*. Pretoria

Republic of Korea. 2025. *Framework Act on the Development of Artificial Intelligence and Establishment of Trust*. Seoul: Government of the Republic of Korea.

Republic of South Africa. 1996. *Constitution of the Republic of South Africa, 1996*. Pretoria
Republic of South Africa. 2000. *Promotion of Administrative Justice Act 3 of 2000*. Pretoria
Sheehan, M. 2023. *China's AI Regulations and How They Get Made*. Washington, DC: Carnegie Endowment for International Peace.

UNESCO. 2021. *Recommendation on the Ethics of Artificial Intelligence*. Paris: UNESCO.

Veale, M. and Borgesius, F.Z. 2025. 'Demystifying the EU AI Act', *Computer Law & Security Review*, 45, Article 105112.

Xinhua News Agency. 2025. 'Namibia launches AI readiness report'. *Xinhua News Agency*, August 2025.