KPMG

AI Governance and Control

September 2025

kpmg.com/ng

Contents

Introduction

Artificial intelligence (AI) has the potential to transform society and improve lives through advancements in fields like commerce, healthcare, transportation, and cybersecurity. It can drive economic growth and support scientific breakthroughs that enhance our world. However, these benefits come with significant risks. AI technologies can negatively impact individuals, groups, and even the environment. Similar to other technologies, these risks can emerge in various ways and can be either long-term or short-term.

From biased algorithms that reinforce systemic discrimination, to opaque machine learning models making high-stakes decisions without human oversight, the risks posed by poorly governed AI systems are not theoretical, they are unfolding in real time. Regulators across jurisdictions are responding with new frameworks such as the EU AI Act and updates to ISO standards, signaling a shift from innovation-at-any-cost to responsible AI deployment.

But stream of these rising risks, the most forward-thinking organizations understand that governance is not just a regulatory checkbox; it is a strategic capability.

In this piece, we examine the structural elements of a mature AI governance program, highlight leading practices and frameworks, and offer actionable insights into how businesses can build effective controls that scale with their AI ambitions; without stifling innovation. **The goal is not to fear AI, but to govern it wisely**.

Have a great read!

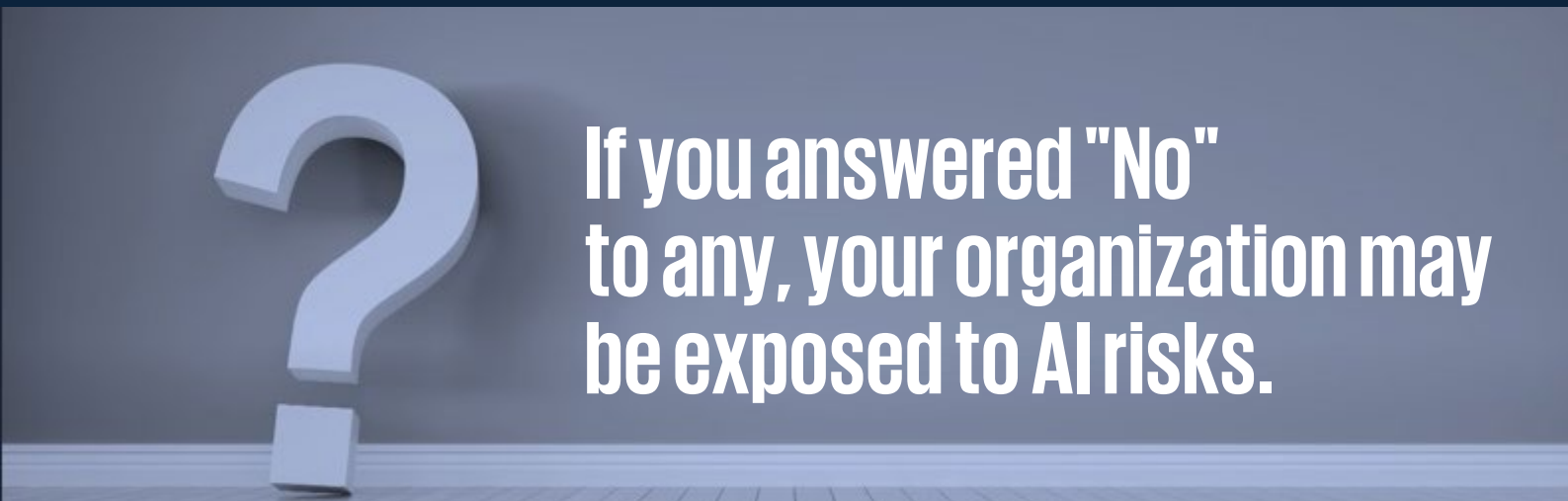
Lawrence Amadi
Partner & Head,
Technology Risk
KPMG in Africa.

Interactive Checkpoint:

Is Your Organization AI-Ready?

Answer these quick questions to assess your governance maturity

1. Do you have clear policies defining who owns AI risk?
2. Can you explain how your AI models make decisions?
3. Do you regularly audit AI systems for bias or drift?
4. Are employees trained in ethical AI principles?



If you answered "No" to any, your organization may be exposed to AI risks.

Core Technology Risks in AI Systems

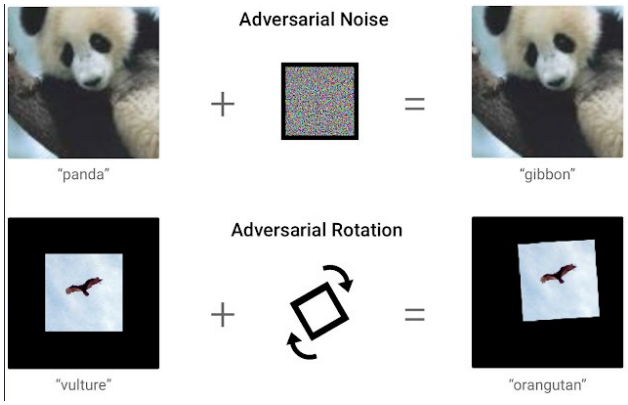
As AI systems become deeply embedded in critical decision-making, the risks inherent in their design and operation demand close attention. Two central technology risks in AI systems are **Robustness and Safety failures**, where models produce harmful or unreliable outputs under unexpected conditions, and **Bias, Fairness, and Representation issues**, where uneven data or design choices reinforce inequities. These risks not only affect technical performance but also have significant social and ethical implications, making them critical challenges in AI governance.

Robustness and Safety Failures

Adversarial Examples

Adversarial examples are inputs that have been deliberately and subtly altered in ways nearly invisible to humans, yet they can cause an AI system to produce highly inaccurate outputs.

For instance, by adding barely noticeable pixel noise to an image of a cat, an image classifier might confidently misidentify it as a dog. To the human eye, it still appears to be a perfectly ordinary cat, but the model is completely misled.

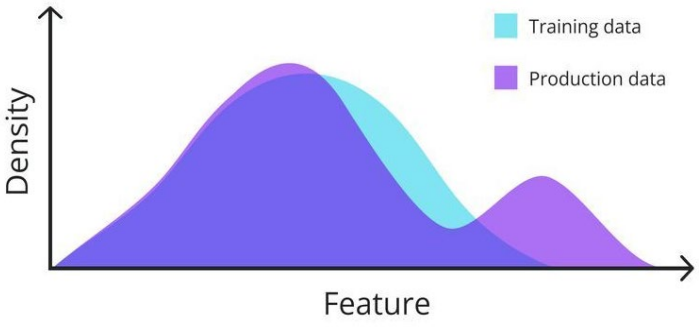


Google Research.

A study showed that changing just one pixel in an image misled deep learning models, causing up to 67.97% of CIFAR-10 images to be misclassified with 74.03% average confidence.

Distribution Shift

Another key safety risk in AI is distribution shift, when a system that is trained on one type of data encounters different conditions in the real world. For example, a voice assistant trained on clean studio recordings may falter in noisy environments, or a self-driving car trained mainly in sunny weather may struggle in rain or snow.



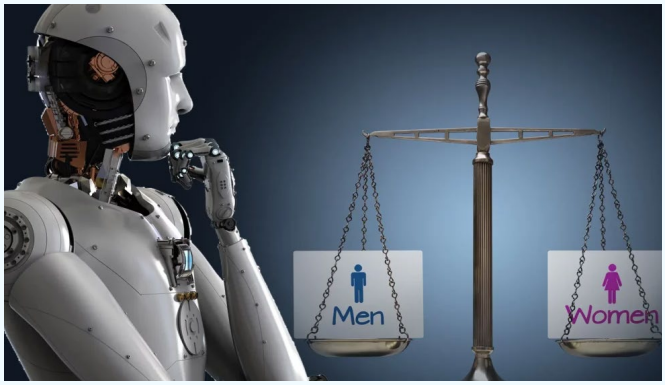
Distribution Shift - (2023)

In high-stakes areas like healthcare, the consequences can be severe. A diagnostic model trained on images from one hospital might fail when used in another with slightly different equipment, potentially missing a tumor or misidentifying healthy tissue. Such failures go beyond technical errors, they pose real safety risks and can undermine public trust in AI systems

Bias, Fairness, and Representation Issues

AI systems learn patterns from the data they are trained on, which means they also absorb any biases present in that data. As a result, these models can reinforce and amplify existing societal inequalities. For example, word embedding models trained on Google News data associated “man” with professions like engineer or programmer, while linking “woman” to roles such as nurse or homemaker. These seemingly subtle biases can manifest in real-world applications, shaping automated decisions and content in ways that unfairly disadvantage certain groups.

Gender Bias in LLMs



Despite the increasing development on artificial intelligence framework, Generative AI’s outputs still reflect a considerable amount of gender and sexuality-based bias.

The leading Large Language models draw reasoning on vast datasets scraped from the internet that make it inherit the same biases that are already present in news articles, media, and online culture. As a result, when prompted in seemingly neutral ways, the outputs tend to link men with technical, business, and leadership roles such as engineer, CEO, or scientist, while women are more often associated with caregiving or domestic positions like nurse, teacher, or homemaker.

This skewed representation is not limited to text generation alone. In visual outputs, women are frequently depicted in stereotypical ways, while men dominate professional or authoritative settings.

Socio-cultural Bias

Another critical area is linguistic and cultural bias. LLMs are trained on predominantly English text which make them exhibit a latent bias favoring Western cultural values.
This creates a risk that AI-driven tools will reinforce a form of digital colonialism, where Western perspectives dominate and local voices are marginalized.

Research has shown that without cultural prompting, large language models consistently align most closely with values found in Western societies, particularly those in the Anglosphere and Protestant Europe. In contrast, all models showed the greatest distance from cultural values in African-Islamic regions, with countries such as Jordan, Libya, Tunisia, Ghana, and Moldova representing the least alignment. Overall, the findings highlight that models, by default, reflect a Western cultural orientation rather than a globally neutral stance.

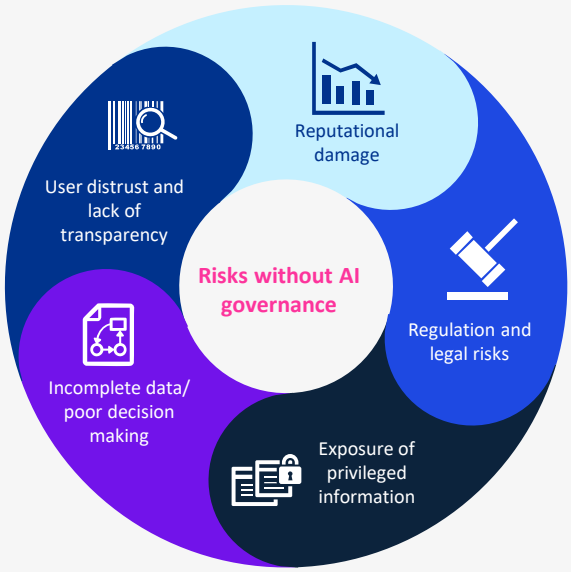
Effective AI governance is essential to ensure that systems are not only technically reliable but also fair and accountable, preventing robustness failures and biased outcomes from causing real harm in society.

AI Governance

Effective AI governance requires robust oversight mechanisms that not only mitigate risks such as bias, privacy violations, security breaches, and misuse, but also foster responsible innovation and build public trust.

Governance in this sense is not just about compliance, it is about ensuring that AI technologies are developed and applied in ways that are safe, fair, and aligned with human values.

Achieving this balance calls for collaboration across multiple stakeholders, including developers, business leaders, policymakers, regulators, ethicists, and end-users, so that AI systems reflect societal priorities while remaining adaptable to emerging challenges...



Pillars of AI Governance

These core pillars of AI governance form the foundation for responsible AI use and are operationalized through AI control frameworks that translate these principles into practical measures and safeguards.

Accountability & Oversight
Clear accountability, oversight bodies, and mechanisms for redress to ensure responsible decision-making

Transparency & Explainability
Thorough documentation, explainability, and open communication of risks and system boundaries to all stakeholders

Fairness & Bias Mitigation
Addressing biases, conducting impact assessments, and promoting inclusive design that reflects diverse populations.

Safety & Robustness
Stress-testing, continuous monitoring with incident response, and proven performance under real-world, high-stakes conditions.

Compliance & Alignment with Standards
Align with laws and standards, adopt industry frameworks, and evolve practices as regulations change

Stakeholder Engagement & Public Trust
Stakeholder engagement, alignment with societal values, and ethical, proactive deployment.

Privacy & Security
Robust privacy safeguards, legal compliance, and strong protections against misuse and security threats.

AI Control Framework

An AI Control Framework serves as the operational backbone of AI governance, turning high-level principles into concrete policies, processes, and safeguards.

It defines the controls, checks, and procedures needed to ensure that AI systems function as intended, reliably, safely, and fairly, while also remaining transparent, accountable, and aligned with organizational and societal expectations.

Its key objectives are

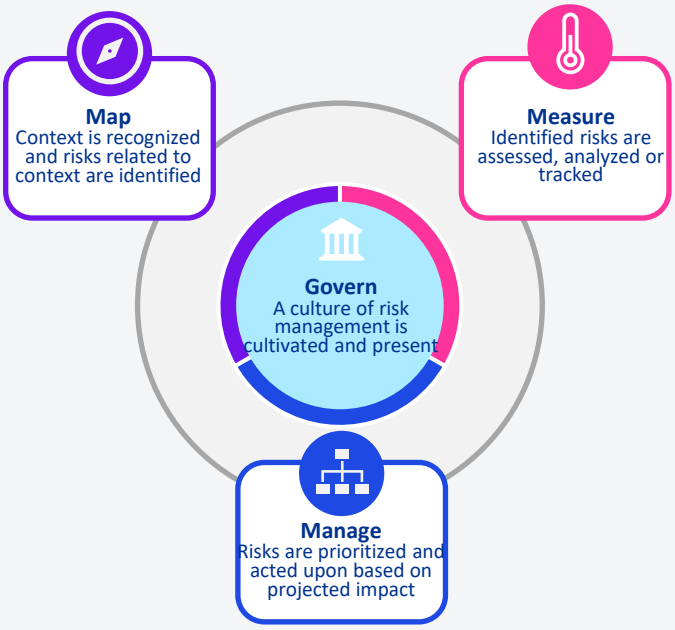
1. Operationalize governance into day-to-day AI processes
2. Detect and mitigate risks during data preparation, model development, and deployment
3. Provide mechanisms for monitoring, auditing, and continuous improvement
4. Demonstrate compliance and accountability through traceable controls

The table shows examples of AI control frameworks, which together turn governance principles into practical safeguards..

Framework	Purpose & Strengths
NIST AI RMF	Governs AI risks across development stages, widely used in US and enterprises
ISO/IEC 42001	Establishes formal AI governance systems with risk assessment and continuous improvement
Unified Control Framework (UCF)	Offers efficient, unified controls adaptable across jurisdictions
ModelOps	Governs AI model lifecycle in production with traceability and governance controls
Trustmarkinitiative.ai	Focuses on conversational AI standards, transparency, and compliance training
Framework Convention on AI (Council of Europe)	Legally binds participating countries to human rights-oriented AI governance principles

Adoptable IA Control Frameworks

One of the adoptable AI control frameworks is the **NIST AI Risk Management Framework (AI RMF),** which provides voluntary guidance to help organizations identify, assess, and mitigate AI risks while fostering trustworthy and responsible AI systems.



Govern

The Govern function is the foundation for managing AI risks. It embeds AI risk management into the organization's culture and policies, ensuring it isn't an afterthought. This involves leaders taking responsibility, assigning clear roles, and training staff. The goal is to align AI use with the company's mission and legal requirements while promoting transparency and a "safety-first" mindset. It also requires engaging with external stakeholders to understand the broader impact of AI systems.

Map

The Map function is about understanding the context of an AI system. Before development, organizations must clarify the system's purpose, its operating environment, and its potential benefits and harms. This includes documenting intended uses, potential misuses, and limitations, as well as identifying risks introduced by third-party components. By carefully mapping the scope and impact, organizations can make informed decisions about whether to proceed and lay the groundwork for later risk

Measure

The Measure function focuses on testing and evaluating AI systems. Once risks are mapped, organizations must validate that the AI performs as intended and meets criteria for trustworthiness, such as fairness and security. Measurement goes beyond simple accuracy; it includes assessing robustness, monitoring for new risks, and ensuring the system is explainable. It's an ongoing process that requires continuous monitoring throughout development and deployment, with an emphasis on independent evaluations from outside experts or

Manage

The Manage function is where organizations take action based on what they've learned from mapping and measuring. It involves prioritizing risks and directing resources toward mitigating the most significant harms. This includes deploying strategies to reduce risk, implementing controls, and, if necessary, deactivating a system if its risks are too high. It also requires having a plan to respond to inevitable failures quickly and transparently. Management is a continuous process, as risks and contexts evolve over time, requiring organizations to stay vigilant and update their risk posture.

AI for Customer Support in Nigerian Banks (Customer Experience & Operational Efficiency)

Scenario

Several Nigerian banks have deployed AI chatbots and virtual assistants to enhance customer support services. These AI systems can handle a wide range of customer queries, including account balances, transaction histories, and loan applications, improving customer experience and reducing wait times for service. The use of AI in this area has led to improved operational efficiency and customer satisfaction.

Risk Types

- If the AI system misinterprets customer queries, it could lead to incorrect responses or customer frustration.
- AI chatbots often require access to sensitive personal data, which must be protected to prevent data breaches.
- AI-driven interactions must comply with data privacy and financial regulations, particularly in terms of user consent and data handling.

Governance Controls

- **Data Privacy Controls:** Ensure that AI systems handling sensitive customer data comply with local data protection laws, such as Nigeria's Data Protection Regulation.
- **Customer Feedback Loops:** Implement feedback mechanisms to continuously improve the AI's ability to handle customer queries accurately.
- **Audit and Monitoring**: Regularly audit the AI systems to ensure they are performing as expected and are compliant with relevant regulations.

Mobile Money Fraud Detection in Ghana (Operational Risk & Privacy Risk)

Scenario:

Mobile money services are widely used in Ghana, allowing for financial transactions via cell phones. Some service providers have implemented AI-powered systems to detect fraudulent transactions. However, these systems have been found to generate a high number of false positives, blocking legitimate transactions and inconveniencing users. Additionally, concerns were raised about the misuse of customer data for fraud detection, leading to privacy violations.

Risk Types:

- **Operational Risk**: The AI system fails to accurately differentiate between legitimate and fraudulent transactions, leading to operational disruptions and customer dissatisfaction.
- **Privacy & Data Security Risk:** AI models process large volumes of sensitive customer data, and mishandling or unauthorized access could lead to data breaches or privacy violations.
- **Regulatory Risk:** Failure to comply with data protection laws (e.g., Ghana's Data Protection Act) could expose service providers to legal action and regulatory penalties.

Governance Controls:

- **False Positive Monitoring:** Implement continuous performance monitoring of AI systems to identify and address high rates of false positives, reducing operational disruptions.
- **Data Privacy & Security**: Establish robust data governance policies to ensure that customer data is handled securely, in compliance with local privacy laws.

Nigeria's Credit Scoring System (Bias and Financial Inclusion Risk)

Scenario:

In Nigeria, several financial institutions have begun implementing AI-based credit scoring systems to assess the creditworthiness of individuals and businesses. However, some of these AI models have been criticized for excluding large segments of the population, particularly individuals who lack a formal credit history or those in rural areas without access to digital records. The reliance on data such as utility bills or social media activity disproportionately affects poorer individuals, resulting in systemic exclusion.

Risk Types:

- **Bias & Discrimination**: The AI model unintentionally discriminates against individuals from lower-income or rural backgrounds, who do not have access to digital footprints or formal financial records.
- **Financial Risk**: The use of biased models leads to unequal access to credit.
- Reputational InclusionRisk: Banks and financial institutions may face backlash from consumers and regulators, damaging their reputation.

Governance Controls:

- **Bias Detection & Mitigation:** Regular audits of AI models to detect and address bias, ensuring that all demographic groups are treated equitably.
- **Regulatory Compliance:** Align AI-based credit scoring systems with regulatory frameworks for financial inclusion and consumer protection.

How can KPMG Tech Risk help?

As organizations accelerate their adoption of artificial intelligence, the need for strong governance, robust controls, and regulatory foresight becomes paramount. KPMG's Tech Risk Unit offers a multidisciplinary, forward-looking approach to AI governance; combining deep expertise in emerging technologies, regulatory intelligence, cybersecurity, ethics, and enterprise risk management. We help clients not only meet compliance requirements, but design AI systems that are trustworthy, transparent, and strategically aligned.

Our services span the full AI lifecycle, addressing both pre-deployment and post-deployment risks with a balance of rigor and practicality:

AI Risk Assessments: We conduct comprehensive assessments of AI initiatives, models, and systems to identify vulnerabilities across data sourcing, model training, algorithm design, deployment, and operational use. Our methodology considers ethical, legal, and operational risks such as bias, explainability, data misuse, unintended outcomes, and compliance with emerging global standards.

Control Design & Testing: KPMG works with organizations to design and validate internal controls tailored to AI systems. These controls promote fairness, transparency, accountability, and cybersecurity. We ensure that governance protocols are not only in place but also tested for effectiveness creating a defensible control environment.

Regulatory Compliance Alignment: The AI regulatory landscape is dynamic and fragmented. From the EU AI Act and OECD AI Principles, to the NIST AI RMF, our specialists help organizations interpret and operationalize relevant legal obligations and voluntary standards. We provide compliance assessments and implementation roadmaps.

Independent Third-Party AI Audits: Where external AI solutions or vendors are involved, we offer independent audit and assurance services to assess the integrity and trustworthiness of third-party models before adoption. This includes validating vendor claims, identifying hidden biases or weaknesses, and ensuring the models align with your organization's risk appetite, ethical principles, and performance expectations.

Continuous Monitoring & Model Lifecycle Oversight:. AI systems can degrade over time due to drift in data, shifting real-world conditions, or subtle bias creep. We help clients design and implement robust monitoring frameworks that enable early detection of performance issues, compliance failures, or ethical concerns ensuring that AI remains aligned with intended outcomes throughout its operational life.

Contact us

Lawrence Amadi
Partner and Head
Technology Risk
KPMG West Africa
T: +234-803-975-4017
E: lawrence.amadi@ng.kpmg.com

Chukwuemeke Igabari
Associate Director
Technology Risk
KPMG West Africa
T: +234 702 500 1098
E: chukwuemeke.igabari@ng.kpmg.com

home.kpmg/ng
home.kpmg/socialmedia