Jurnal Studi Multidisiplin Ilmu (JASMI)
ISSN: 3026-7412, Vol 4, No 2, 2026, 43-57 https://doi.org/10.35912/jasmi.v4.n2.p43-57.2026

Global AI Cooperation and Responsible Use: Balancing Security, Social Dynamics, and Moderated Adoption

Adams Adah Otakwu[1*]
Centre for Civil-Military Cooperation (CIMICO), Abuja, Nigeria[1]
cimicocentre@gmail.com[1*]

Article History
Received on 15 March 2026
1st Revision on 28 March 2026
2nd Revision on 19 April 2026
Accepted on 21 May 2026

Abstract

Purpose: This study examines how the world can pursue global cooperation on Artificial Intelligence (AI) while keeping its deployment responsible, secure and socially sustainable, focusing on the security risks, social consequences and governance gaps that accompany AI's rapid diffusion, and arguing for moderated adoption rather than unrestricted acceleration or blanket resistance.

Methodology:A qualitative documentary design was used, drawing on recent reports from the United Nations, UNESCO, African Union, Organisation for Economic Co-operation and Development (OECD), United Nations Conference on Trade and Development (UNCTAD), International Labour Organization (ILO), International Monetary Fund (IMF), International Energy Agency (IEA), and World Economic Forum, supported by peer-reviewed literature published mainly between 2021 and 2026 and analysed thematically.

Results: AI already raises productivity, strengthens cybersecurity and expands access to services, yet it also widens inequality, enables synthetic disinformation and exposes roughly a quarter of the global workforce to task level disruption. Governance remains fragmented, and African countries face exclusion from rule making alongside limited capacity to capture AI's economic benefits.

Conclusions: A risk based, human centred and internationally coordinated approach to AI adoption can reconcile innovation with safety.

Limitations: The study relies on secondary institutional and scholarly sources rather than primary interviews or original quantitative modelling.

Contributions: It offers a five pillar framework linking AI security, social protection, governance, digital inclusion and responsible innovation, with attention to Africa's development priorities.

Keywords: *Artificial Intelligence, Global Governance, International Security, Moderated Adoption, Social Dynamics*

How to Cite: Otakwu, A. A. (2026). Global AI Cooperation and Responsible Use: Balancing Security, Social Dynamics, and Moderated Adoption. *Jurnal Studi Multidisiplin Ilmu*, 4(2), 43-57.

1. Introduction

1.1 Background

Artificial Intelligence (AI) has moved from a niche computing specialism to a general purpose technology that touches nearly every field of public and economic life, from healthcare and agriculture to finance, education and national defence (United, & Development, 2025; United, 2026). Unlike earlier waves of automation confined to particular industries, contemporary AI systems can read, write, classify, predict and generate content across many domains at once, and the arrival of generative AI has pushed these capabilities into ordinary workplaces and households within a few short years (United & Development, 2025). That speed of diffusion is precisely what makes the technology difficult to govern: rules, institutions and professional norms simply have not had time to catch up.

The scale of the shift is visible in market figures. United and Development (2025) estimates that the global AI market could grow from roughly 189 billion dollars in 2023 to 4.8 trillion dollars within a decade, an expansion of about twenty five times. Yet this growth is far from evenly spread. The United States and China together account for close to a third of AI publications and around sixty percent of patents, and computing infrastructure remains concentrated in a small group of advanced economies (Cerutti, Kido, Li, Melina, & Wingender, 2025). Countries with strong digital infrastructure, research capacity and skilled labour are positioned to capture most of the value AI creates, while countries that lack these foundations risk experiencing the technology mainly as a source of disruption rather than opportunity.

AI's reach into security affairs is equally significant. The same systems that improve cybersecurity, intelligence analysis and disaster response can be repurposed to automate cyber attacks, manufacture convincing synthetic media and probe digital defences at a scale no human team could match (Organisation, & Development, 2024; World, 2026). Generative tools have lowered the cost of producing deceptive text, audio and video so sharply that citizens and institutions increasingly struggle to tell authentic material from manipulated content, with consequences for elections, public health messaging and civic trust (Romero, 2024).

Labour markets are being reshaped in parallel. The International Labour Organization estimates that roughly one worker in four holds a job with meaningful exposure to generative AI, though most affected occupations are likely to be transformed rather than eliminated outright (Gmyrek, Berg, Kamiński, Konopczyński, Ładna, Nafradi, Rosłaniec, & Troszyński, 2025). Large language models behave like a genuine general purpose technology whose influence spans wage levels and sectors rather than a narrow automation tool confined to routine tasks (Eloundou, Manning, Mishkin, & Rock, 2024). Field evidence tells a similar story: assistance from a generative AI tool has been shown to raise output quality and cut completion time on professional writing tasks while narrowing the gap between weaker and stronger performers, and comparable productivity gains have been documented among real world customer support agents (Noy, & Zhang, 2023; Brynjolfsson, Li, & Raymond, 2025).

Africa sits at the centre of this debate because it carries both the promise and the peril of the AI transition. The African Union's Continental Artificial Intelligence Strategy frames AI as a strategic asset for Agenda 2063, capable of transforming health, agriculture, finance and education, provided that development proceeds on ethical, equitable and Africa centred terms (African, 2024). Realising that promise depends on closing persistent gaps in infrastructure, skills and research capacity, since AI adoption without those foundations tends to widen rather than narrow existing disparities (Mwaura, 2024).

Global cooperation has therefore become a practical necessity rather than a diplomatic aspiration. UNESCO's Recommendation on the Ethics of Artificial Intelligence, endorsed by all 194 member states, sets out shared commitments to human rights, transparency, fairness and human oversight, while the Organisation for Economic Co-operation and Development (OECD) AI Principles emphasise inclusive growth, robustness and accountability across the AI lifecycle(United, 2026).

These frameworks matter, yet coordination among states, firms and researchers still lags well behind the pace of technical change, leaving a governance gap that this study sets out to examine (Roberts, Hine, Taddeo, & Floridi, 2024).

1.2 Statement of the Problem

The central difficulty is not a shortage of technological capability but a mismatch between the speed of AI deployment and the readiness of the institutions meant to govern it. AI can raise productivity, extend services and strengthen security, yet its rapid and uneven expansion also generates real harm through misinformation, cyber exposure, privacy erosion, labour disruption and widening inequality (Organisation, & Development, 2024). A first strand of the problem is regulatory fragmentation. AI systems are trained in one jurisdiction, distributed through another and used by people scattered across dozens of legal regimes, yet safety standards, data protection rules and liability arrangements remain largely national. United and Development (2025) reports that by 2024 only seventy five countries were participating in even one of seven major AI governance initiatives, leaving one hundred and eighteen countries, mostly developing economies, outside every one of them. Those excluded countries are left absorbing the consequences of technologies they had no part in shaping (Oloyede, 2025).

A second strand concerns security. The dual use character of AI means that a system built to defend a network can be adapted to attack one, and defensive gains achieved in one domain can be matched or outpaced by offensive uses in another (Papadopoulos, Katsikas, & Pitropakis, 2025). This same duality runs through military applications, where autonomous and semi-autonomous systems raise unresolved questions about human control, accountability and the meaningful exercise of judgement in the use of force (Bode, Huelss, Nadibaidze, Qiao-Franco, & Watts, 2024). A third strand is social disruption. Occupational exposure at the scale documented by the ILO represents a substantial adjustment for workers, and without deliberate investment in reskilling, social dialogue and income support that adjustment risks translating into unmanaged insecurity rather than shared prosperity (Gmyrek, Berg, Kamiński, Konopczyński, Ładna, Nafradi, Rosłaniec, & Troszyński, 2025).

A fourth strand is inequality between and within countries. Because AI capability clusters around economies with strong infrastructure, capital and skilled labour, the IMF's modelling suggests advanced economies could see growth effects markedly larger than those available to low income countries under plausible adoption scenarios (Cerutti, Kido, Li, Melina, & Wingender, 2025). Left unaddressed, AI could therefore harden rather than soften the global economic hierarchy. A final and easily overlooked strand is the tendency to equate speed of adoption with quality of adoption. A system can complete a task quickly while still producing biased, opaque or harmful outcomes if it was poorly designed, trained or supervised. High stakes applications in health, policing, employment, finance and security therefore need stronger safeguards than low stakes uses, and the practical challenge is building governance capable of making that distinction in real time. For Africa, this challenge is sharpened by the need to close infrastructure and skills gaps at the same moment that poorly governed systems could introduce new social, economic and security vulnerabilities (African, 2024).

1.3 Objectives of the Study

This study sets out to examine the principal opportunities and security risks associated with artificial intelligence, to assess how AI adoption is reshaping employment, inequality, information integrity and human welfare, to explain why fragmented governance undermines effective oversight and why international cooperation is required, and to propose measures for responsible and moderated AI adoption that protect security and social stability without foreclosing legitimate innovation.

1.4 Research Questions

The study is guided by four questions. What are the major opportunities and security risks associated with artificial intelligence? How is AI adoption affecting employment, inequality, information integrity and human welfare? Why is international cooperation necessary for effective AI governance

and security? What measures can promote responsible and moderated AI adoption while preserving the benefits of technological innovation?

2. Literature Review

2.1 Conceptual Clarification

2.1.1 Artificial Intelligence

Artificial intelligence refers to computational systems capable of performing tasks commonly associated with human cognition, including learning, prediction, classification, language processing, problem-solving, and content generation (Organisation & Development, 2024). Contemporary AI encompasses a broad range of technologies, including machine learning systems, generative AI models, large language models, predictive algorithms, and increasingly autonomous applications capable of making or supporting complex decisions. The rapid expansion of these technologies has extended AI applications into education, healthcare, finance, public administration, security, communication, and employment, thereby increasing its influence on both institutional processes and everyday social interaction. In this study, AI is understood as a socio-technical system rather than as an isolated technological tool because its social consequences depend not only on technical capabilities but also on the quality of training data, the purposes for which systems are designed, the environments in which they are deployed, and the governance mechanisms surrounding their use (United & Cultural, 2021). Consequently, the same AI system may generate substantially different social outcomes depending on who controls it, how it is implemented, and what safeguards are established to prevent misuse or unintended harm.

2.1.2 Global AI Cooperation

Global AI cooperation refers to the arrangements through which states, international organisations, technology companies, research institutions, professional communities, and civil society actors collaborate to establish shared principles, technical standards, governance frameworks, and capacity-building mechanisms for the development and use of artificial intelligence (United & Development, 2025). Such cooperation has become increasingly important because the development, deployment, and consequences of AI frequently extend beyond national jurisdictions. AI systems may rely on globally sourced data, operate across digital platforms accessible in multiple countries, and generate consequences for international labour markets, information environments, cybersecurity, trade, and political communication. As a result, fragmented national regulation may be insufficient to address risks that emerge across borders. Global cooperation can therefore support greater consistency in regulatory principles, facilitate information sharing and technical assistance, strengthen international standards, and reduce disparities between countries with advanced technological capacity and those with limited institutional or digital resources. Effective cooperation is particularly important in ensuring that emerging AI governance arrangements reflect not only the interests of technologically advanced states and corporations but also the developmental concerns and policy priorities of lower- and middle-income countries.

2.1.3 Responsible AI Use

Responsible AI use refers to the development, deployment, and management of artificial intelligence in ways that respect human rights, safety, privacy, transparency, fairness, accountability, and human agency throughout the AI lifecycle (United & Development, 2025). This concept requires developers, governments, organisations, and other institutions to identify potential risks before AI systems are introduced, assess their possible effects on individuals and communities, and establish appropriate safeguards proportional to the seriousness of those risks. Responsible use also requires transparency regarding how important decisions are generated, mechanisms for meaningful human oversight, procedures for contesting harmful or inaccurate outcomes, and institutional accountability when AI produces adverse consequences. These principles are particularly important where AI systems influence employment, public services, financial access, education, healthcare, law enforcement, or political communication because automated decisions in such areas may directly affect fundamental rights and social opportunities. Responsible AI should therefore not be interpreted simply as compliance with technical standards but as an ongoing governance process through which innovation is aligned with human wellbeing and institutional responsibility.

2.1.4 Social Dynamics

Social dynamics refers to the processes through which individuals, groups, organisations, and institutions interact, adapt, and respond to social, economic, political, and technological change. In the context of artificial intelligence, these dynamics include changes in employment patterns, education, communication, access to information, social trust, privacy, institutional relationships, and interactions between citizens and public authorities (United & Development, 2025). AI can positively influence social dynamics by expanding access to knowledge, increasing productivity, supporting public-service delivery, improving communication, and facilitating more efficient decision-making. At the same time, poorly governed AI may contribute to employment displacement, digital exclusion, privacy violations, misinformation, social manipulation, or declining trust in public and private institutions (World, 2026). The social effects of AI are therefore not inherently beneficial or harmful but depend upon patterns of access, institutional safeguards, public literacy, and the distribution of technological benefits and risks. Understanding these dynamics is essential for assessing whether AI adoption strengthens social resilience and inclusion or instead intensifies existing inequalities and institutional vulnerabilities.

2.1.5 Moderated AI Adoption

Moderated AI adoption refers to the deliberate and proportionate integration of artificial intelligence into economic, governmental, security, and social systems at a pace that corresponds with institutional readiness, social capacity, risk assessment, and regulatory capability (United & Cultural, 2021). Moderation in this context does not imply opposition to technological innovation or an attempt to prevent AI development. Rather, it recognises that different AI applications generate different levels of social and institutional risk and therefore require differentiated governance responses. Low-risk applications may be introduced with relatively limited regulatory intervention, while systems capable of significantly affecting employment, privacy, public safety, political participation, or fundamental rights require stronger oversight and safeguards. Moderated adoption therefore seeks to balance technological advancement with institutional preparedness by ensuring that societies possess adequate legal frameworks, technical expertise, public awareness, monitoring mechanisms, and accountability structures before highly consequential AI systems become deeply embedded in social and economic life. This approach is particularly relevant for countries where rapid technological adoption may outpace regulatory capacity and thereby increase exposure to social disruption or technological dependency.

2.2 Theoretical Framework

2.2.1 Human Security Theory

Human Security Theory broadens the concept of security beyond state territory to include threats to people's economic, food, health, environmental, personal, community, and political wellbeing. Applied to AI, the theory suggests that technology becomes a security concern when it threatens livelihoods, privacy, access to reliable information, social cohesion, or fundamental rights. Recent scholarship similarly shows that AI creates both opportunities and social-security risks, requiring its effects on human wellbeing to be considered alongside technological benefits (Malmio, 2024). In security-related applications, the increasing autonomy of AI systems also raises concerns regarding human oversight, accountability, and compliance with humanitarian principles (Putra et al., 2025). Synthetic disinformation can therefore erode public trust, while employment disruption and other AI-related transformations may generate economic and social insecurity among affected populations (Gmyrek et al., 2025).

2.2.2 Technology Governance Theory

Technology Governance Theory emphasises that technological innovation is shaped by laws, institutions, market structures, professional norms, and political decisions. Effective AI governance therefore requires transparency, accountability, risk management, ethical standards, and meaningful human oversight. Recent studies demonstrate that responsible AI governance depends on translating broad ethical principles into structural, procedural, and relational practices within organisations (Papagiannidis et al., 2025). Public institutions likewise require governance structures that integrate ethical principles, standards, training, and oversight into AI implementation (Almeida & Santos

Júnior, 2025). Evidence from Nigeria further shows that infrastructure limitations, digital literacy, privacy concerns, human-capital readiness, and supportive regulatory frameworks significantly influence AI adoption in public governance (Nwosu et al., 2024). Governance should therefore function not merely as a restriction on innovation but as an institutional mechanism for enabling technological development while limiting disproportionate social harm.

2.2.3 Digital Divide and Development Perspective

The Digital Divide and Development Perspective highlights how unequal access to infrastructure, data, digital skills, and institutional capacity shapes the ability of countries and social groups to benefit from AI. Research shows that previous digital experiences, socioeconomic differences, and digital confidence can influence people's engagement with and attitudes toward AI technologies (Bentley et al., 2024), while access to information and communication technologies and digital skills also contribute significantly to AI literacy (Celik, 2023). Evidence from Indonesia demonstrates that infrastructure and government policy remain important factors in addressing regional digital inequality (Yulianita et al., 2024). Similarly, recent research on AI and Big Data adoption shows that technological readiness, managerial support, and organisational capacity affect the ability of smaller enterprises to benefit from digital transformation (Pratiwi et al., 2026). AI democratisation may reduce some barriers by expanding access to affordable and scalable technologies, although institutional and capability differences continue to shape outcomes in emerging economies (Chinwendu et al., 2026). Governance that manages technological risks without expanding access and capacity may therefore reinforce rather than reduce existing inequalities.

2.2.4 Application of the Theories

Taken together, these perspectives provide an integrated framework for analysing AI, security, governance, and development. Human Security Theory focuses on the protection of individuals, Technology Governance Theory explains the role of institutions in regulating AI, and the Digital Divide Perspective highlights inequalities in access and capacity. This combined framework is used to assess whether contemporary AI development and global cooperation support responsible innovation, social stability, and inclusive growth.

3. Research Methodology

The study employed a qualitative documentary research design, which is particularly appropriate for examining a rapidly evolving policy field where relevant evidence is primarily contained in institutional reports, regulatory frameworks, policy documents, and recent scholarly literature rather than in data suited to survey or experimental analysis (United, & Development, 2025). This approach enabled the study to synthesise diverse forms of evidence while maintaining attention to the institutional and policy contexts in which AI governance is developing.

Sources were selected purposively based on three main criteria: institutional authority, relevance to the study's four research questions, and recency. Greater weight was given to publications produced by institutions directly involved in the policy processes under examination, with particular emphasis on materials published between 2021 and 2026 due to the rapid development of generative AI and related governance debates (Gmyrek, Berg, Kamiński, Konopczyński, Ładna, Nafradi, Rosłaniec, & Troszyński, 2025; United and Development, 2025). Principal sources included publications from the United Nations, UNESCO, the African Union, OECD, UNCTAD, International Labour Organization, International Monetary Fund, International Energy Agency, and World Economic Forum. These materials were complemented by peer-reviewed studies on AI governance, security, labour economics, and African digital development, identified through keyword searches and citation tracking.

The collected evidence was organised thematically according to four analytical areas derived from the research questions: AI opportunities and security risks, social and economic transformation, global governance and cooperation, and responsible and moderated AI adoption. Within each theme, the study identified key actors, mechanisms, risks, and outcomes, and then compared evidence across multiple sources to determine areas of convergence, divergence, and remaining gaps in the literature. The analysis adopted a cautious interpretive approach to causal relationships. Where available

evidence demonstrated only association, exposure, or emerging risk, these findings were reported accordingly rather than treated as established causal effects. The analytical framework therefore followed a sequential logic linking AI development with economic and social transformation, emerging security and social risks, governance and international cooperation, responsible safeguards, and moderated adoption, ultimately assessing their implications for more inclusive and secure AI development.

4. Results and Discussions

4.1 Global AI Opportunities and Security Implications

4.1.1 AI and Economic Productivity

AI carries real potential to lift productivity by automating routine cognitive work, sharpening forecasts and accelerating research (Cerutti, Kido, Li, Melina, & Wingender, 2025). United and Development (2025) documents applications spanning content creation, product design, coding and production systems, while IMF modelling suggests that AI driven productivity gains could meaningfully lift global output under favourable adoption conditions (Cerutti et al., 2025). These gains are not automatic, however. They depend on access to infrastructure, data, skills and capable institutions, so AI alone cannot substitute for the structural investment that development still requires (Cerutti et al., 2025).

4.1.2 AI and National Security

AI has become integral to intelligence analysis, cybersecurity, surveillance and the protection of critical infrastructure, in part because it can process volumes of information and detect patterns beyond the reach of conventional systems (Organisation, & Development, 2024). The same capability is available to hostile actors, who can use AI to automate attacks, generate deceptive content or exploit digital vulnerabilities at scale (World, 2026). Comparative research on major military powers shows that governance of this risk remains fragmented, with Chinese, Russian and United States practices diverging sharply on what counts as meaningful human control over autonomous systems (Bode, Huelss, Nadibaidze, Qiao-Franco, & Watts, 2023). That divergence is shaped partly by how national leaderships narrate AI as a marker of strategic status, a pattern documented in Russian and Chinese discourse alike (Bode, Huelss, Nadibaidze, Qiao-Franco, & Watts, 2024; Bode et al., 2023. While public narratives about autonomous weapons continue to influence how policymakers frame acceptable levels of risk (Watts & Bode, 2023). A broader research programme on algorithmic warfare concludes that accountability gaps persist across the weapons lifecycle regardless of which state is doing the developing (Bode et al., 2024).

4.1.3 AI and Information Security

Generative AI has sharply lowered the cost of producing convincing synthetic text, audio, images and video, expanding both the volume and the sophistication of potentially misleading content (Organisation, & Development, 2024). Legal scholarship increasingly treats generative deepfakes as a distinct human rights concern rather than an ordinary content moderation problem, given their capacity to damage reputations, distort elections and enable coercion (Romero, 2024). Bibliometric analysis confirms that research on generative AI's role in misinformation has expanded sharply since large language models became widely available, with clear links to several of the Sustainable Development Goals (Raman, Nair, Nedungadi, Sahu, Kowalski, Ramanathan, & Achuthan, 2024). The practical governance challenge is to curb harmful manipulation while preserving legitimate speech and creative use of AI generated content.

4.1.4 AI and Cybersecurity

AI strengthens cyber defence through anomaly detection, automated monitoring and faster incident response, yet it can equally sharpen phishing, automate intrusion attempts and identify vulnerabilities more efficiently than a human attacker could (Organisation & Development, 2024). Cybersecurity scholarship identifies this dual use tension as one of the field's most persistent structural challenges, since AI strengthens both defence and offence at the same time (Papadopoulos et al., 2025). Because

cyber infrastructure and digital services cross borders almost by definition, meaningful resilience requires international information sharing, common technical standards and coordinated incident response rather than purely national defences (United & Development, 2025).

4.1.5 AI, Autonomous Systems and Military Security

AI is increasingly embedded in intelligence, surveillance, logistics and decision support systems, and the governance stakes rise sharply where such systems could influence decisions involving the use of force (United & Cultural, 2021). UNESCO's framework insists that AI must never displace ultimate human responsibility, which implies that high risk military applications require stronger human oversight, clear chains of accountability and appropriate legal safeguards. The evidence reviewed in section 4.1.2 above suggests those safeguards remain unevenly developed across the world's major military powers.

4.2 AI and Social Dynamics

4.2.1 Employment and Labour Transformation

AI is reorganising work by automating some tasks and augmenting others. The ILO's finding that roughly a quarter of the global workforce holds a job with meaningful generative AI exposure indicates that this transformation is already well underway rather than a future prospect (Gmyrek et al., 2025). Because most occupations combine tasks that AI can perform with tasks that still require human judgement, the ILO expects transformation to be far more common than wholesale job elimination (Gmyrek et al., 2025). Task level evidence reinforces this reading: large language models affect a substantial share of work across occupations once complementary software is included, positioning them as a genuine general purpose technology rather than a narrow automation tool (Eloundou et al., 2024). Controlled experimental evidence adds further support, showing that generative AI assistance can raise output quality and cut completion time on professional writing tasks while narrowing performance gaps between workers (Noy & Zhang, 2023). A pattern echoed in field evidence from customer support agents using an AI assistant, where productivity gains coincided with reduced staff turnover (Brynjolfsson et al., 2025). The policy implication is that governments and employers should treat reskilling, lifelong learning and social dialogue as central to responsible adoption rather than optional extras.

4.2.2 AI and Inequality

AI can generate substantial productivity gains, but who captures those gains depends heavily on access to capital, infrastructure and skills (Cerutti et al., 2025). IMF analysis suggests that advanced economies could benefit disproportionately because they already possess stronger AI readiness and broader access to digital infrastructure (Cerutti et al., 2025). Without deliberate inclusion policies, this dynamic risks becoming another channel through which existing global inequalities are reproduced rather than reduced, a concern illustrated by AI-powered digital identity systems whose benefits have not always reached the vulnerable populations they were designed to serve (Siad & Sagar, 2025). Complementary macroeconomic modelling similarly finds that labour income and wealth inequality could widen within countries unless regulatory frameworks and labour reallocation support are upgraded alongside AI adoption (Cazzaniga et al., 2024).

4.2.3 AI and Social Interaction

AI increasingly mediates communication, information retrieval, education and professional work (United & Development, 2025). These uses can widen access to knowledge and services, yet they can also reshape how people interact and make decisions. Overreliance on automated outputs may weaken independent judgement and deepen dependence on systems whose limitations users do not fully grasp (United & Cultural, 2021). The emotional dimension of this dependence is visible in the growing use of conversational AI for informal mental health support, where users report genuine comfort even though such tools were never designed or clinically validated as therapeutic interventions (Siddals et al., 2024). AI literacy should therefore include critical evaluation of system limitations alongside purely technical skills.

4.2.4 AI and Education

AI offers real educational value through personalised learning, translation, teaching support and expanded access to materials, benefits that matter most where teachers and resources are unevenly distributed (African, 2024). United and Cultural (2021) nonetheless stresses ethical, privacy and quality concerns that accompany these applications. Institutions should therefore treat AI as a complement to, rather than a replacement for, teachers and human learning, and should adopt clear, regularly updated academic integrity policies rather than reacting piecemeal to each new generative capability as it appears (Fowler, 2023).

4.2.5 AI, Privacy and Human Rights

AI systems typically depend on large volumes of data, which raises serious questions about consent, surveillance and discrimination. Automated profiling can affect people in ways that are difficult to understand or contest where transparency and accountability mechanisms are weak. Research on the relationship between data protection law and AI surveillance infrastructure cautions that formal legal compliance does not by itself prevent pervasive monitoring from becoming entrenched (Yew, Qin, & Venkatasubramanian, 2024), which is part of why UNESCO's framework grounds AI governance explicitly in human rights rather than compliance checklists alone (Bakiner, 2023; United, & Cultural, 2021). These safeguards matter most where AI is used by governments, law enforcement or other authorities exercising significant power over citizens.

4.3 The Global AI Governance Gap

AI development is outpacing the capacity of laws, regulators and courts to adapt, and because AI systems and data routinely cross borders, purely national governance cannot address many of the risks involved. The OECD AI Principles, built around inclusive growth, human centred values, transparency, robustness, safety and accountability, remain one of the most widely referenced international frameworks, requiring systematic risk management across the AI lifecycle and appropriate mechanisms for human oversight (Organisation & Development, 2024). UNESCO's Recommendation on the Ethics of Artificial Intelligence, adopted in 2021 and applicable to all 194 member states, places human rights and dignity at the centre of AI governance and adds commitments on sustainability and adaptive governance (United & Cultural, 2021).

Despite these frameworks, governance remains fragmented in practice. United and Development (2025) finds that one hundred and eighteen countries took part in none of seven major AI governance initiatives by 2024, even as the most advanced capabilities remain concentrated among a small number of states and firms. Existing regulatory authorities could be used more collaboratively to close parts of this gap without waiting for entirely new institutions, particularly across Africa (Oloyede, 2025). The underlying challenge can be reduced to three linked questions: who develops AI, who benefits from it, and who bears its risks. A durable governance system has to answer all three rather than focusing narrowly on restricting harmful uses. The stakes of leaving these questions unresolved are visible in the political domain, where campaigns increasingly use AI for targeting, content generation and voter outreach faster than electoral integrity regulation can keep pace (Foos, 2024).

4.4 Africa and the Imperative of Responsible AI Adoption

Africa has genuine opportunities to apply AI to agriculture, health, education, finance and public administration (African, 2024). The African Union's Continental Artificial Intelligence Strategy frames AI as a strategic asset for Agenda 2063 and calls for an Africa centred, development focused approach grounded in ethics, equity and regional cooperation. Realising that vision depends on stronger infrastructure, skills, data governance and research capacity, and on avoiding a purely digital divide reading of Africa's challenge, since the continent's digital transformation is itself proceeding unevenly across countries and social groups (Mwaura, 2024). Precision agriculture illustrates both the promise and the limits of this agenda: AI tools have meaningfully improved resource use and yields for small scale African farmers, but their societal benefit remains constrained by uneven access to training, connectivity and finance (Aroba & Rudolph, 2024).

Africa's task is therefore not simply to adopt AI but to build the capacity needed to evaluate, adapt and where possible develop it locally, since dependence on externally built systems can leave countries with little control over data, algorithms or standards. Low risk applications that address clear development needs can move quickly, while higher risk uses should undergo the testing, oversight and impact assessment appropriate to their potential consequences (Organisation, & Development, 2024; United, & Cultural, 2021).

4.5 The Energy and Environmental Dimension of AI

AI depends on computing infrastructure, servers and data centres that consume substantial electricity (International, 2025). At the same time, AI can improve energy forecasting, grid management and system efficiency, so the technology sits on both sides of the energy ledger at once. This creates a genuine policy paradox, such as AI can help make energy systems more efficient even as the infrastructure that supports AI adds new pressure on electricity demand. For countries with limited generating capacity, large scale AI infrastructure can compete directly with other essential uses of scarce power, so responsible adoption needs to weigh energy efficiency, renewable integration and transparent infrastructure planning alongside the technology's economic promise (International, 2025).

4.6 Why the World Needs Moderated AI Adoption

Moderated adoption rests on a simple premise: the fact that a system can perform a task does not automatically mean it should perform that task without human oversight (United & Cultural, 2021). Adoption should be risk based, so that low risk applications move quickly while high risk applications in health, policing, employment, finance and national security carry stronger assessment and safeguards. Evidence from public health practice underlines why this matters, since AI deployed in clinical and population health settings without equity safeguards can worsen rather than reduce existing health disparities (Dankwa-Mullan, 2024).

Adoption should also be human centred, since United and Cultural (2021) insists that AI must not displace ultimate human responsibility, particularly where automated decisions carry significant consequences for individuals or communities. It should be evidence based, following the OECD's call for systematic risk management across the AI lifecycle, including transparency, traceability and accountability mechanisms capable of flagging harmful system behaviour before it spreads. It should be socially prepared, given the scale of employment exposure documented by the ILO, which requires real investment in skills, education and social dialogue rather than an assumption that markets will absorb the adjustment on their own (Gmyrek, Berg, Kamiński, Konopczyński, Ładna, Nafradi, Rosłaniec, & Troszyński, 2025). Finally, adoption should be internationally coordinated, because AI development and use cross borders and no state can manage the resulting risks alone (United & Development, 2025).

4.7 A Global Framework for Responsible and Moderated AI Adoption

Table 1 sets out a practical framework built around five interconnected pillars: AI security, social protection, AI governance, digital inclusion and responsible innovation. For each pillar, the table distinguishes an immediate response that governments and institutions can undertake now from a medium to long term response that requires sustained institution building. Read across the pillars, the table shows that no single ministry or agency can deliver responsible AI on its own, since security institutions, education systems, regulators, infrastructure planners and innovation policymakers each hold a piece of the solution.

Table 1. A Five pillar framework for responsible and moderated AI adoption

Pillar	Immediate Response	Medium to Long Term Response
AI Security	Risk assessment and human oversight for high risk systems	International safety standards and cross border AI security cooperation

Pillar	Immediate Response	Medium to Long Term Response
Social Protection	Worker transition assistance and targeted AI literacy programmes	Institutionalised reskilling, lifelong learning and adaptive social protection
AI Governance	National AI guidelines, audits and transparency requirements	Interoperable international AI governance frameworks
Digital Inclusion	Expanded connectivity, AI literacy and affordable access	Investment in computing infrastructure, data systems and local AI capacity
Responsible Innovation	Controlled testing and regulatory sandboxes	Indigenous AI research, ethical innovation and sustainable AI infrastructure

Delivering this framework requires coordinated action across security institutions, education systems, labour ministries, finance ministries, justice systems, technology firms, researchers and civil society (United & Cultural, 2021). The underlying objective is a whole of society approach in which government, industry, researchers, workers and citizens all participate in setting the conditions for safe and socially beneficial AI adoption (United & Development, 2025).

4.8 Discussion

Taken together, the evidence reviewed here shows that AI is neither inherently beneficial nor inherently harmful. Its consequences depend heavily on how it is designed, deployed and governed (United & Cultural, 2021) . Human security theory helps explain why AI governance must reach beyond technical performance: because AI touches employment, income, privacy, information access and political participation, technological risk can translate quickly into broader human security harm, so AI's implications need to be judged by human security criteria and not only by national security ones (United, 1994).

Technology governance theory reinforces the case for institutions capable of restraining high risk applications while still permitting beneficial innovation, and the OECD and UNESCO frameworks demonstrate that this balance is achievable without rejecting technological progress outright (United & Cultural, 2021). The digital divide perspective, meanwhile, highlights an international distribution problem that governance alone cannot solve: AI research, patents, computing infrastructure and investment remain concentrated in a limited set of countries, and IMF evidence suggests advanced economies could see substantially larger AI driven growth than low income countries under plausible scenarios (Cerutti et al., 2025).

The employment evidence reinforces the case for social preparedness. The ILO's finding that roughly a quarter of workers worldwide already hold jobs with meaningful generative AI exposure shows that occupational transformation is well underway (Gmyrek et al., 2025), and because transformation is expected to outpace outright elimination, policy should concentrate on reskilling, worker participation and social dialogue rather than treating AI purely as an unemployment threat (Gmyrek et al., 2025). The security evidence points in the same direction: AI strengthens cybersecurity and intelligence work even as it expands the potential for cyber abuse and synthetic disinformation, so security governance has to manage both the beneficial and the harmful sides of the same capability at once (World, 2026).

For Africa specifically, the central lesson is that AI must be approached as an opportunity and a governance challenge simultaneously. The African Union's Continental AI Strategy supplies a sound policy foundation for equitable, Africa centred AI development, but turning that foundation into results depends on sustained investment in infrastructure, skills, data and institutional capacity (African Union Commission, 2024). The study therefore rejects both unrestricted technological acceleration and reflexive resistance to AI. Moderated adoption offers a middle path in which low risk applications can advance quickly while high risk systems face proportionate testing, oversight, transparency and accountability.

4.9 Policy Implications

Four policy implications follow from this analysis. AI governance needs to become part of mainstream international security governance, since cyber threats, disinformation, autonomous systems and digital vulnerabilities all cross national boundaries (World, 2026). AI policy needs to become part of social policy, because employment, education, inequality and social protection will determine whether AI's benefits are broadly shared or narrowly captured (Cazzaniga et al., 2024; Gmyrek et al., 2025). Global AI cooperation needs to address technological inequality directly, since gaps in infrastructure, data and computing power translate into unequal economic outcomes if left unaddressed (Cazzaniga et al., 2024). Finally, AI adoption needs to remain proportionate to risk, with stronger testing, human oversight, transparency and accountability reserved for high risk applications, while lower risk uses are not burdened with unnecessary constraints.

5. Conclusions

5.1 Conclusion

Artificial intelligence is one of the most consequential technological shifts of the current international system because its reach extends across production, health, education, agriculture, finance, public administration and security all at once. The evidence examined in this study shows that AI generates real productivity and development benefits, but those benefits arrive bundled with risks touching misinformation, cybersecurity, privacy, inequality, employment and social cohesion. AI governance can no longer be treated as a purely domestic matter of technology policy. AI systems, data flows, infrastructure and markets cross borders freely, and their consequences reach populations far beyond the jurisdictions where the underlying systems were built. Global cooperation is therefore necessary to build common standards, share knowledge, strengthen capacity and manage risks that no single state can handle alone.

The distribution of AI capability underlines why equity has to sit at the centre of that cooperation. Research and computing power remain heavily concentrated, and modelling evidence suggests advanced economies could obtain far larger economic benefits than low income countries under current trends. A governance framework that manages risk without also improving access to infrastructure, skills and technology could end up widening the global AI divide it was meant to close. The social stakes are equally clear. Roughly a quarter of the global workforce already holds a job with meaningful exposure to generative AI, and while transformation is expected to be far more common than outright job loss, that transformation still demands real investment in skills, education and social dialogue. Security considerations reinforce the same conclusion, since AI strengthens both defensive and offensive capability at the same time, so innovation has to be matched with proportionate safeguards rather than left to outpace them. For Africa, the challenge is developmental and strategic in equal measure. The African Union has identified AI as central to Agenda 2063 and has called for ethical, equitable and cooperative AI development, but delivering on that vision will require sustained investment in infrastructure, data, skills, research and governance capacity so that African countries can participate meaningfully in the AI economy rather than remain largely dependent consumers of systems built elsewhere.

The study ultimately argues against two extremes. Unrestricted acceleration risks letting technological harm outrun institutional and social capacity, while excessive restriction risks denying societies benefits they could legitimately capture. Moderated adoption offers the more workable path, matching the speed and scope of AI deployment to the level of risk involved and to the genuine readiness of institutions and society to manage it. That means risk classification, human oversight, impact assessment, transparency, accountability, worker preparation and international cooperation working together, so that AI governance combines innovation with responsibility and national capability with international partnership, and so that technological progress is measured by its contribution to human welfare and inclusive development and not only by the sophistication of the machines involved.

5.2 Research Limitations

This study relies on a qualitative documentary design and does not include primary interviews with regulators, technology developers, workers or affected communities, so it reflects the formal

institutional and scholarly record rather than lived experience of AI adoption on the ground. Reliance on institutional self reporting from bodies such as the United Nations, the OECD and the African Union, while offset by cross referencing across independent sources, may still understate practical implementation gaps that do not appear in official communications. The analysis is confined to documentary and secondary literature and does not include original quantitative modelling of labour market outcomes, AI investment flows or governance effectiveness. Because generative AI and its surrounding governance are changing quickly, some figures and policy positions cited here may already have moved on by the time this study is read.

5.3 Suggestions and Directions for Future Study

Future research could usefully track how national and regional AI governance frameworks are actually implemented rather than merely adopted, since the gap between formal commitment and practical enforcement remains poorly understood. Primary research with African policymakers, technology developers and workers would help test whether the Africa centred approach set out in continental strategy documents is translating into measurable outcomes in infrastructure, skills and firm level adoption. Longitudinal labour market studies could clarify how occupational transformation associated with generative AI evolves over several years rather than in the short experimental windows examined by most current evidence. Comparative work across regions could also examine how differing governance philosophies, spanning risk based regulation, principles based guidance and market led approaches, perform against shared benchmarks for safety, equity and innovation.

Acknowledgement

The author is grateful to colleagues at the Centre for Civil-Military Cooperation (CIMICO), Abuja, for comments on earlier drafts of this study, and to the anonymous reviewers whose observations helped sharpen the analysis presented here. Any remaining errors are the author's own.

References

African Union Commission. (2024). Continental artificial intelligence strategy. *African Union*. Retrieved from https://au.int/en/documents/20240809/continental-artificial-intelligence-strategy

Almeida, P. G. R. de, & Santos Júnior, C. D. dos. (2025). Artificial intelligence governance: Understanding how public organizations implement it. *Government Information Quarterly*, *42*(1), 102003. https://doi.org/10.1016/j.giq.2024.102003

Aroba, O. J., & Rudolph, M. (2024). Systematic literature review on the application of precision agriculture using artificial intelligence by small-scale farmers in Africa and its societal impact. *Journal of Infrastructure, Policy and Development*, *8*(13), 1-19. https://doi.org/10.24294/jipd8872

Bakiner, O. (2023). The promises and challenges of addressing artificial intelligence with human rights. *Big Data & Society*, *10*(1). https://doi.org/10.1177/20539517231205476

Bentley, S. V., Naughtin, C. K., McGrath, M. J., Irons, J. L., & Cooper, P. S. (2024). The digital divide in action: How experiences of digital technology shape future relationships with artificial intelligence. *AI and Ethics, 4, 901–915*. https://doi.org/10.1007/s43681-024-00452-3

Bode, I., Huelss, H., Nadibaidze, A., Qiao-Franco, G., & Watts, T. F. A. (2023). Prospects for the global governance of autonomous weapons: Comparing Chinese, Russian, and US practices. *Ethics and Information Technology*, *25*(1), 5. https://doi.org/10.1007/s10676-023-09678-x

Bode, I., Huelss, H., Nadibaidze, A., Qiao-Franco, G., & Watts, T. F. A. (2024). Algorithmic warfare: Taking stock of a research programme. *Global Society, 38, 1–23*. https://doi.org/10.1080/13600826.2023.2263473

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*, *140*(2), 889-942. https://doi.org/10.1093/qje/qjae044

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A. J., Pizzinelli, C., Rockall, E. J., & Tavares, M. M. (2024). *Gen-AI: Artificial intelligence and the future of work (IMF Staff Discussion Note No*.

Celik, I. (2023). Exploring the determinants of artificial intelligence (AI) literacy: Digital divide, computational thinking, cognitive absorption. *Telematics and Informatics, 83, 102026*. https://doi.org/10.1016/j.tele.2023.102026

Cerutti, E. M., Garcia Pascual, A. I., Kido, Y., Li, L., Melina, G., Mendes Tavares, M., & Wingender, P. (2025). *The global impact of AI: Mind the gap (IMF Working Paper No*.

Chinwendu, W. S., Arachie, A. E., Okeke, O. I., & Chiedozie, F. C. (2026). Does artificial intelligence democratisation enhance small business operations in emerging economies? Annals of Management and Organization Research, 8(1), 125–142. https://doi.org/10.35912/amor.v8.n1.p125-142.2026

Dankwa-Mullan, I. (2024). Health equity and ethical considerations in using artificial intelligence in public health and medicine. *Preventing Chronic Disease, 21, 240245*. https://doi.org/10.5888/pcd21.240245

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, *384*(6702), 1306-1308. https://doi.org/10.1126/science.adj0998

Foos, F. (2024). The use of AI by election campaigns. *LSE Public Policy Review*, *3*(3), 1-7. https://doi.org/10.31389/lseppr.112

Fowler, D. S. (2023). AI in higher education: Academic integrity, harmony of insights, and recommendations. *Journal of Ethics in Higher Education, 3, 127–143*. https://doi.org/10.26034/fr.jehe.2023.4657

Gmyrek, P., Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Nafradi, B., Rosłaniec, K., & Troszyński, M. (2025). Generative AI and jobs: A 2025 update. *International Labour Organization*.

International Energy Agency. (2025). *Energy and AI*.

Malmio, I. (2024). Artificial intelligence and the social dimension of sustainable development: Through a security perspective. *Discover Sustainability, 5, 466*. https://doi.org/10.1007/s43621-024-00677-6

Mwaura, J. (2024). Navigating the digital horizon: Future trends and predictions in Africa. *Dialogues on Digital Society, 1–5*. https://doi.org/10.1177/29768640241252714

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187-192. https://doi.org/10.1126/science.adh2586

Nwosu, C. C., Obalum, D. C., & Ananti, M. O. (2024). Artificial intelligence in public service and governance in Nigeria. *Journal of Governance and Accountability Studies*, *4*(2), 109-120. https://doi.org/10.35912/jgas.v4i2.2425

Oloyede, R. (2025). Collaborative regulation: Leveraging existing authorities for effective AI governance in Africa. *International Review of Law, Computers & Technology, 1–20*. https://doi.org/10.1080/13600869.2025.2506916

Organisation for Economic Co-operation and Development. (2024). *OECD AI principles*.

Papadopoulos, P., Katsikas, S., & Pitropakis, N. (2025). Editorial: Cybersecurity and artificial intelligence: Advances, challenges, opportunities, threats. *Frontiers in Big Data, 7, 1537878*. https://doi.org/10.3389/fdata.2024.1537878

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, *34*(2), 101885. https://doi.org/10.1016/j.jsis.2024.101885

Pratiwi, L. A., Subandi, S., Padriyadi, Y., & Oktaria, E. T. (2026). Artificial intelligence and Big Data adoption in MSMEs: A systematic literature review. *Jurnal Studi Pemerintahan dan Akuntabilitas*, *6*(1), 29-44. https://doi.org/10.35912/jastaka.v6i1.6039

Putra, R. A., Buaton, T., & Prasetyo, A. H. (2025). Legal analysis of the use of artificial intelligence systems in military operations carried out by the TNI reviewed from the principles of humanitarian law. *Dynamics of Politics and Democracy*, *3*(1), 39-47. https://doi.org/10.35912/dpd.v3i1.2978

Raman, R., Nair, V. K., Nedungadi, P., Sahu, A. K., Kowalski, R., Ramanathan, S., & Achuthan, K. (2024). Fake news research trends, linkages to generative artificial intelligence and sustainable development goals. *Heliyon, 10, e24727*. https://doi.org/10.1016/j.heliyon.2024.e24727

Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2024). Global AI governance: Barriers and pathways forward. *International Affairs*, *100*(3), 1275-1286. https://doi.org/10.1093/ia/iiae073

Romero Moreno, F. (2024). Generative AI and deepfakes: A human rights approach to tackling harmful content. *International Review of Law, Computers & Technology*, *38*(3). https://doi.org/10.1080/13600869.2024.2324540

Siad, R., & Sagar, A. (2025). AI-powered digital identity systems and the new digital divide: The case of World. *Big Data & Society*, *12*(2). https://doi.org/10.1177/27523543251340713

Siddals, S., Torous, J., & Coxon, A. (2024). "It happened to be the perfect thing": Experiences of generative AI chatbots for mental health. *npj Mental Health Research*, *3*(1), 41. https://doi.org/10.1038/s44184-024-00097-4

United Nations. (2026). *Artificial intelligence*.

United Nations Conference on Trade and Development. (2025). Technology and innovation report 2025: Inclusive artificial intelligence for development. Paper presented at United Nations.

United Nations Development Programme. (1994). *Human development report 1994: New dimensions of human security*.

United Nations Educational, Scientific and Cultural Organization. (2021). *Recommendation on the ethics of artificial intelligence*.

Watts, T., & Bode, I. (2023). Machine guardians: The Terminator, AI narratives and US regulatory discourse on lethal autonomous weapons systems. *Cooperation and Conflict*. https://doi.org/10.1177/00108367231198155

World Economic Forum. (2026). *The global risks report 2026*.

Yew, R.-J., Qin, L., & Venkatasubramanian, S. (2024). You still see me: How data protection supports the architecture of AI surveillance. Paper presented at Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society.

Yulianita, A., Subardin, S., & Zulfikri, Z. (2024). Government size and digital inequality in Indonesia. *Journal of Governance and Accountability Studies*, *4*(1), 31-41. https://doi.org/10.35912/jgas.v4i1.1981