

Ethical AI for the Developing World: A Comparative Study of Global Frameworks and Their Applicability in Pakistan

Hafiz Muhammad Huzaifa(Corresponding Author)

Undergraduate Student, Ziauddin University, College of Computing and Software Engineering, Karachi, Pakistan. hafizmuhamamdhuzaifa13@gmail.com

Marium Ehtesham

Lecturer, Ziauddin University, College of Computing and Software Engineering, Karachi, Pakistan. marium.ehetesham@gmail.com

Zainab Rafi

Undergraduate Student, Ziauddin University, College of Computing and Software Engineering, Karachi, Pakistan. zainabraf164@gmail.com

Riffat Sehar Haroon

Lecturer, Ziauddin University, College of Computing and Software Engineering, Karachi, Pakistan. riffatsehar4561@gmail.com

Hamna Yasir

Undergraduate Student, Ziauddin University, College of Computing and Software Engineering, Karachi, Pakistan. hamnayasir2021@gmail.com

ABSTRACT

As of early 2026, the pace of expansion of Generative AI (Generative AI) has far exceeded the governing structures that have been developed between 2019 and 2021. To developing countries such as Pakistan, Western-centric laws, such as the EU AI Act, pose a kind of Compliance-Innovation Paradox, making it prohibitively expensive, yet it does not mitigate the real-world risks like Digital Colonialism and cultural hallucinations in Large Language Models (LLMs). A Qualitative Systematic Review (QSR) of 32 primary sources, assessed using a special

Triple-Lens Matrix, is used in this study to determine gaps in system implementation. The study concludes that the current international norms lack the essential characteristics to manage the non-deterministic character of Generative AI and its socio-technical influence on the Global South. In filling these gaps, this paper will present the Three-Pillar Adaptive Governance Model of Pakistan that will combine regulatory innovation by including SMEs in safe-to-fail sandboxes, technical sovereignty by curating indigenous Urdu-centric datasets, and ethical regulation based on the Trusteeship (I'timāni) Paradigm. The model enables us to use Adl (Justice) and Maslaha (Public Interest) as the functional auditing standards to give us a blueprint on how to move out of the Algorithmic Dependence to Epistemic Sovereignty. The report concludes that value-based indigenized framework is necessary to ensure that Pakistan can spearhead a digital future that is both technologically proficient, economically inclusive, and culturally oriented.

Keywords: AI Governance, Generative AI, Pakistan AI Policy, Digital Colonialism, Epistemic Sovereignty, Islamic AI Ethics, Regulatory Sandboxes, Urdu LLMs, Algorithmic Auditing.

1. Introduction

The fast development of Artificial Intelligence (AI) has shifted the concept to a realistic possibility, which is a structural pressure that is essentially changing profoundly healthcare, finance, education, and creative arts. In the past ten years, AI governance has been unable to keep up with this technological pace. The earliest frameworks were created in 2019-2021 e.g., the Ethics Guidelines of Trustworthy AI produced by the EU and the Recommendations by the UNESCO mostly revolved around normative principles of transparency, fairness, and accountability [1], [2]. Nonetheless, with the advent of Generative AI (GenAI) and Large Language Models (LLMs) the technological landscape has radically changed. Only conservative forms of governance have failed to solve the problem of the quick rise of strong models such as GPT-4. Generative AI generates new information, unlike earlier generations of AI, which analyzed existing data, which poses previously unknown systemic risks like hallucinations, deepfakes, and copyright infringement [4]. Recent 2026 studies show that these models have reached more autonomous levels, and that there is a need to migrate these models to adaptive models that can handle reasoning-aware safety and moral responsibility of

autonomous agents [30]. The general disadvantage of the current global discourse is that it is geopolitically minded; the largest part of AI ethics literature is created by the Global North. Individualistic and socio-economic realities at the community level are generally not compatible with privacy-oriented centric approaches. of the Global South. Data poverty, absence of infrastructure and data poverty are common issues in third world countries like Pakistan. this misalignment is added to by language marginalization. This kind of environment fosters what has been referred to as. as Digital Colonialism, a hypothetical 2026 model that describes how international technology giants introduce digital standards, marginalizing local bodies of knowledge and language. such as Urdu [31]. Moreover, implementing strict laws such as the EU AI Act can introduce the prohibitive compliance costs on Small and Medium Enterprises (SMEs), which may deaden the local innovation ecosystems [32].

Research Gap (Motivation)

Although the development of global safety is general, there is a serious gap of research comparing GenAI governance with the socio-cultural and infrastructural realities of Pakistan. This advanced-level incompatibility is addressed in this study with an Indigenized Governance model suggested. It aims to supplant imported procedural norms with a paradigm of a Trusteeship (I'timāni) to guarantee that AI is understood as a moral Amanah (Trust) based on communal well-being (Maslaha).

1.1 Research Objectives

In order to fill these governance gaps, this study aims to:

- Assess technical and ethical constraints of current global AI frameworks (EU, UNESCO) to reduce non-deterministic risks of Generative AI, namely, hallucinations and cross-cultural validation failures.
- Examine how Western liberal AI ethics and the realities of developing economies are socio-technically disconnected, with an emphasis on digital colonialism, data poverty, and the compliance cost to local SMEs.
- Offer a Three-Pillar Strategic Model, which integrates regulatory innovation with the Trusteeship paradigm, based on Adl and Maslaha to develop a culturally adaptive governance model.

1.2 Research Questions

- **RQ1:** How do the current global AI governance regimes (e.g., EU AI Act, UNESCO) fail to reduce the non-deterministic risks of Generative AI, e.g., the cultural hallucinations, and cross-cultural validation errors?
- **RQ2:** How does the Digital Colonialism paradigm in 2026 intensify algorithmic dependence, and data poverty in the socio-technical infrastructure in Pakistan?
- **RQ3:** How can an example of localized governance based on the Islamic concept of Adl (Justice) and I'timani (Trusteeship) offer a practical substitute to Western procedural ethics to the Global South?

1.3 Original Contributions

1. The Trusteeship (I'timāni) Ethical Framework: This article presents a new model of ethics concerning AI that goes beyond procedural transparency to moral accountability. The study grounds AI as an Amanah (Divine Trust) giving the Algorithms Trust a metaphysical explanation based on the Islamic communal values [28, 29].
2. Strategic Pathways to Resource-Constrained SMEs: This study determines a set of Strategic Pathways to the SMEs in Pakistan, and suggests a model of a "Regulatory Sandbox" which permits innovation without the crippling expense of international compliance requirements [32].
3. Counter-Colonial Data Sovereignty Model: The research adds a framework to combat the so-called Epistemic Colonialism by promoting indigenous Urdu-AI datasets and localized LLM fine-tuning, as a roadmap to Pakistan gaining a sense of Data Sovereignty [31].

2. Literature Review:

Another aspect that is still in the state of a paradigm shift in the world of AI governance is that it changes the emphasis on abstract theorization of ethics to practical and implementable systems. This section provides an overview of how the system of governance evolved, how a systemic change was created by Generative AI (GenAI), and why a Global South perspective is so desperately necessary, particularly, with regard to Islamic ethics as a robust alternative to western proceduralism.

2.1 The Era of "Toothless" Principles (2019–2021)

The initial AI governance saw an international consensus on high-level normative principles. A study by Jobin et al. [1] arrived at five pillars of convergence which comprised of transparency, justice, non-maleficence, responsibility, and privacy. Nevertheless a huge chasm of responsibility existed between these ideals and their implementation [21]. Hagendorff [2] was a harsh critic, describing such voluntary codes as toothless and subject to ethics washing, in which companies use ethical rhetoric to do their PR but continue to engage in exploitative data practices. The years have proved that a soft law and voluntary compliance cannot restrict strong corporate interests [3, 12].

2.2 Generative AI Disruption & the Robustness Crisis (2022–2026)

The 2019-era guidelines were disclosed as unsound with the introduction of Large Language Models (LLMs). Weidinger et al. [4] added GenAI-specific harms, including hallucinations and systemic disinformation, to the risk taxonomy. The academic emphasis has changed to Reasoning-aware Safety and Strict Liability by 2026 [5, 30]. Modern surveys [30] point out that, in spite of technical progress, LLMs still experience issues of cross-cultural validation failure and cultural hallucinations, which require the evaluation frameworks that are able to keep pace with changing model capabilities. The governance question that is central has changed to be not what AI should not do, but who is responsible in the non-deterministic failures of autonomous agents [14, 20].

2.3 Beyond Western Autonomy: Digital Colonialism and Sovereignty

One of the main constraints of the current literature is the Western-centric bias of AI models. The systems such as the EU AI Act focus on the values of agency and autonomy- values that tend to oppose the societal socio-economic needs of the Global South [6, 9]. In other countries, such as Pakistan, Digital Colonialism and Data Poverty are the greatest risks, not only the violation of personal privacy [8, 31].

A theory that has developed in 2026 is Algorithmic Dependence, a type of epistemic colonialism, in which the Global South is dependent on Western-standardized models that place local knowledge and languages, like Urdu, into the periphery [31]. As a result, Data Sovereignty by using indigenous datasets is no longer a cultural choice but a strategic governance necessity [7, 31].

2.4 The Islamic "Trusteeship" (I'timāni) Paradigm

In order to properly reconcile Western individualism and communal good, the recent scholarship suggests Pluralist Ethical Benchmarking based on Islamic Jurisprudence [7, 8]. Going beyond the procedural ethics, the Trusteeship (I'timani) model bases AI accountability on the notion of Amanah (Divine Trust) [28].

In contrast to Western paradigms, in which opt-out rights are core, a model based in Maslaha stipulates that deployment of AI must provide a net-positive benefit to the community (Ummah) as a legal requirement [29]. In this paradigm, Adl (Absolute Justice) is a remedial solution to the power imbalance between data extractors (corporations) and data providers (users) with equitable distribution of wealth and benefits [28, 29].

2.5 The Compliance Gap: Corporate Power vs. SME Survival

An increasing conflict exists between technical Red Teaming [15] by Big Tech and legislation by the state [10]. Corporate guidelines are technically advanced but have a tendency to be inherently conflicted [13]. In addition, adoption of Western Hard Law as cut-copy-paste (such as the EU AI Act) places prohibitive compliance costs on SMEs in developing economies with resource constraints [32]. Research into these economies in 2026 suggests that these regions need Strategic Pathways and innovations supported by the government instead of strict, punitive legal frameworks that hamper local development [32].

Table 1: Comparative Analysis of Global AI Governance Frameworks

Framework	Ethical Basis	Enforcement Mechanism	Relevance to Pakistan	2026 Identified Gap
Western (EU AI Act)	Individual Rights & Autonomy [3]	Hard Law (Fines/Bans)	Limited; emphasizes Western legal norms.	Overlooks communal welfare and SME costs [32].
Corporate (Big Tech)	Technical Safety (Red Teaming) [15]	Self-Regulation (Soft Law)	High; used by local developers.	Lacks public accountability and transparency

				[13].
Global South (General)	Anti-Colonialism & Data Equity [7]	Emerging Policy/Sovereignty	Crucial for data and epistemic sovereignty.	Lacks clear technical and auditing standards [31].
Islamic Ethics (Proposed)	Adl (Justice) & Amanah (Trust) [28]	Moral & Legal Benchmarking	Highly relevant for local adoption and trust.	Requires codification into technical "Adl Filters" [29].

The following table offers a comparative synthesis of the main frameworks of governance that are discussed in this paper. It emphasizes the shift towards Western centric procedural ethics to the new Islamic Trusteeship model, including the recent 2026 views of Digital Colonialism and SME Strategic Pathways.

Analytical Justification: Islamic Trusteeship Why?

The Islamic ethics is a practical rule of governance because of three technical reasons:

- Distributive Justice (Adl): Adl, in contrast to the mathematical measures of fairness, deals with the structural power asymmetry in the data economy [29].
- Public Interest (Maslaha): This introduces a conclusive "Success Metric" of AI auditing. When an automated system leads to social agitation or unemployment, it does not pass the Maslaha test and has to be limited [28].
- Trusteeship (Amanah): This turns AI not into a tool of extraction but a sacred trust, which requires human-in-the-loop conditions and moral responsibility of developers [28, 29].

Research Gap and Problem Statement.

Gap Identification: Although literature has established the toothless character of voluntary rules [2] and the technical hazards of Generative AI [4, 5], there is a significant geographic and conceptual gap. Western liberal individualism prevails in current frameworks and centers around individual privacy and autonomy [1]. It lacks scholarship that would help to translate communal ethics, namely, Islamic Adl, Maslaha, and Amanah, into a large-scale technical

system of governance in the Global South. Moreover, no empirical research is conducted that would allow how these localized ethics can be practically applied to curb Digital Colonialism and Algorithmic Dependence in developing countries such as Pakistan [31, 32].

Problem Statement: The uncontrolled spread of Generative AI in the Global South is being controlled with the erroneous Western models that are not in harmony with socio-cultural or economic realities of Pakistan. This incompatibility results in two acute crises:

- **Weak Regulation:** Western models do not consider the bias of local representation when working on Urdu datasets and cannot restrict the extraction of local data that cannot benefit a particular community in any way [31].
- **Implementation Deficit:** Rigid Western compliance norms are a tax on local SMEs to innovate, turning AI governance into a corporate one. seek instead of a social provision [32].

Hence, there is an immediate need to create an Adaptive Governance Framework that combines pluralist ethical benchmarking and technical auditing to protect the common good (Maslaha) in the era of Generative AI.

3. Research Methodology

The study will be a Qualitative Systematic Review (QSR) accompanied by a Comparative Gap Analysis. This qualitative method is opposed to quantitative ones that are concerned with the statistical frequency, according to which subtle interpretation of normative language, legal precedence, and cultural values can be made in AI governance [11, 12]. The methodology examines how abstract principles can be translated into enforceable rules of Generative AI through the prism of the Global South and Pakistan.

3.1 Data Collection Strategy

Data collection of the research was conducted in four work stages and with the assistance of 32 primary sources which were required to get a complete longitudinal picture. It is a process, as depicted in Fig. 1, that allows the gradual transformation of broad global standards to an indigenized form of localized approach.

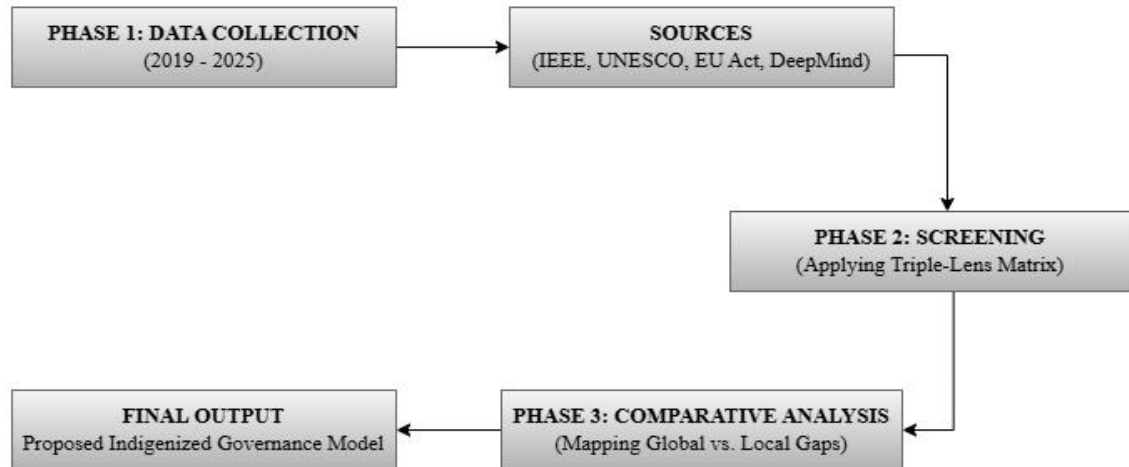


Figure 1. The flow chart of the Systematic Research Methodology which was used in this study was the data collection and screening (Triple-Lens Matrix) and comparative analysis process.applied in this study.

The plan was carried out in three chronological and thematic phases:

- Phase 1 (Global Normative Standards): The current phase involved the analysis of supporting structures (2019 to 2021) such as UNESCO and OECD. Those were the documents that determined the Control Group of early AI ethics, high-level normative principles.
- Phase 2 (Technological Disruption): This phase focused on technical reports and white papers (2023 to 2025) about risks specific to Large Language Models (LLMs). The DeepMind and Microsoft documentation, as well as other sources of this phase, were reviewed to locate the Implementation Gap, which is presented by the non-deterministic nature of Generative AI.
- Phase 3 (Regional and Cultural Synthesis): The final phase involved literature of 2025 to 2026 in response to Digital Colonialism, Trusteeship (I'timani) paradigm, and the Pakistani digital ecosystem-specific needs. The research was significant in determining the local governance requirements and strategy orientation of the SMEs.

3.2 Search Protocol and Databases

The search was conducted across IEEE Xplore, Google Scholar, ACM Digital Library, and SSRN to ensure the replicability.

- **Keywords:** AI Governance, Generative AI Ethics, Islamic AI Ethics, Digital Colonialism, and Algorithmic Auditing.
- **Timeline:** January 2019 to February 2026.
- **Document Count:** 32 primary sources were selected for deep critical analysis after an initial screening of 60+ relevant publications.

3.3 Chronological Evolution Matrix

This table graphically describes the input of the primary studies to the changing paradigm. of AI governance, starting with principles to the Digital Colonialism paradigm of 2026.

Table 2: Chronological Evolution of AI Governance Paradigms (2019–2026)

Year	Category	Reference	Key Finding / Contribution
2019	Foundation	Jobin et al. [1]	Global consensus on 5 principles but lack of implementation agreement.
2020	Auditing	Raji et al. [21]	Guidelines fail without mandatory internal algorithmic auditing.
2021	GenAI Risks	Weidinger et al. [4]	Identified LLM-specific harms (hallucinations) that pre-2020 rules ignored.
2023	Islamic Ethics	Elmahjub [8]	Proposed <i>Adl</i> and <i>Maslaha</i> as functional communal alternatives.
2025	Trusteeship	Ali et al. [28]	Introduced the Trusteeship (I'timāni) model for moral accountability.
2026	Colonialism	Ahmed [31]	Defined Digital Colonialism and the need for epistemic sovereignty.
2026	SMEs	Holland [32]	Identified Strategic Pathways to avoid Western compliance costs for SMEs.

3.4 The Analytical Framework: Triple-Lens Matrix

A specialized Triple-Lens Matrix was used to evaluate all the chosen documents in the context of the Pakistani environment to identify their viability. This framework will make sure that the intended model is technically sound, economically viable, and culturally aligned.

Table 3: Triple-Lens Evaluation Matrix

Dimension	Key Evaluation Question	Core Reference Alignment
1. Operational Viability	Does the framework offer "Strategic Pathways" for SMEs to avoid prohibitive compliance costs?	Holland [32], Raji [21]
2. GenAI Readiness	Can it mitigate non-deterministic risks like hallucinations and reasoning-aware safety failures?	Ferdaus [30], Taelhagh [5]
3. Socio-Cultural Inclusivity	Does it integrate Amanah (Trust) and Adl (Justice) to fight Digital Colonialism?	Ali [28], Ahmed [31], Elmahjub [8]

4. Comparative Analysis of Global Frameworks

The section will compare international AI governance models in the framework of a comparative approach, to compare developed (EU, UNESCO) and agile (Singapore, Big Tech) models to find out their readiness to welcome the Generative AI age.

4.1 The Governance Matrix

Table 1 provides the key differences between Hard Law, Soft Law and Self-Regulation. Critical thought reveals that it is unanimity of the world on the What (Principles), but the world is widely divided on the How (Implementation) [1, 3].

Table 4: Comparative Matrix of Global AI Governance Models

Framework	Origin & Type	Primary Focus	Enforcement Mechanism	Key Limitation

EU AI Act	Regional (Legislation)	Risk Management & Rights	Hard Law: €35M or 7% of world turnover.	SMEs have high compliance cost; is inflexible.
UNESCO	Global (UN) [3]	Human Rights & Diversity	Soft Law: Reporting by the state and moral suasion.	Lacks "teeth"; purely voluntary adoption.
IEEE EAD	Technical (Standard)	Engineering Ethics	Demonstrations: Technical certifications.	Too complex for non-technical policymakers.
Singapore Model	National (Guidance) [10]	Innovation & Pragmatism	Voluntary Tools: A testing toolkit, AI Verify.	Ethics washing risks that are not even legally required.
Corporate Rules	Industry (Self- Reg) [13]	Safety & Reputation	Internal: Red-teaming & audits.	Social ethics is usually subdued by profit.

4.2 Core Divergences (With Concrete Evidence) Analysis.

The matrix demonstrates three key findings, which have been reduced to evidence-based hints according to research gaps:

- **Enforcement vs. Innovation (The SME Burden):** AI is viewed as a high-risk product by EU AI Act. This will help in securing the citizens although there is a major obstacle, the compliance cost. To be more concrete, the cost of a small business to meet the EU standards may be hundreds of thousands of Euros which in the real sense would stifle the emerging Pakistani startup culture, with small sums of money. The Singapore model of Testing-first [10] on the other hand allows business start ups to be innovative and provides solutions to voluntary safety testing.
- **The Teeth Problem and Vulnerability of Global South** The systems like UNESCO [3] are cultural inclusive and not enforced. When it comes to a country like Pakistan, the concept of utilising Soft Law alone is not secure because multinational technology companies are likely to concentrate on business development and forget about the moral aspects of the

residents. International guidelines lack Hard Law teeth and thus are mere recommendations that cannot stop the Digital Colonialism [6].

- The Corporate Audit Gap: The change in corporate structures (ex: Microsoft [15]) is faster than state legislations. However, Internal Audits usage is a conflict of interest. As it has been pointed out, the lack of external control might give rise to the situation when companies are concentrated on the speed-to-market rather than on safety, and that is why the mechanism of the independent technical audit should be state-led [21].

4.3 Can Singapore be a Transferable Model of Pakistan?

An important question is this, can Pakistan just emulate the Singaporean model?

Partial Inspiration: Singapore is founded on Pragmatism and its AI Verify toolkit would find very useful in the case of the Ministry of IT (MoITT) attempting to encourage local AI development without the legal penalties as severe in the EU model.

The Trust Gap: Singapore is an institutional trust-based country. But Pakistan must adopt a Hybrid Approach. We need to possess the technical elasticity of Singapore (so as to support the local developers), and a Pluralist Ethical Benchmarking (Islamic Ethics) so that the AI can emulate the local societal values [8].

4.4 Generative AI Gap: Paradigm Shift.

Pre-GenAI GenAI Frameworks developed prior to the 2022 surge were more focused on predictive AI (e.g., credit scoring), but in essence they are unprepared to deal with the novel disruptions of Generative AI. The previous emphasis was on algorithms and openness of non-interactive decision-making, whereas the new reality of GenAI is full of previously unknown dangers of hallucinations, deepfakes, and massive copyright theft. There is a serious gap because the concept of content provenance (watermarking) is not mentioned in most of the guidelines that were published in 2019.

In this sense, the lack of liability systems towards AI-produced lies is becoming obsolete in the era of LLM unless the regulation frameworks expressly provide the responsibility of autonomous agents. To illustrate the intersection of these technical gaps with the local socio-cultural requirements of the Global South, Fig. 2 graphs the ethical and structural difference that requires a local response.

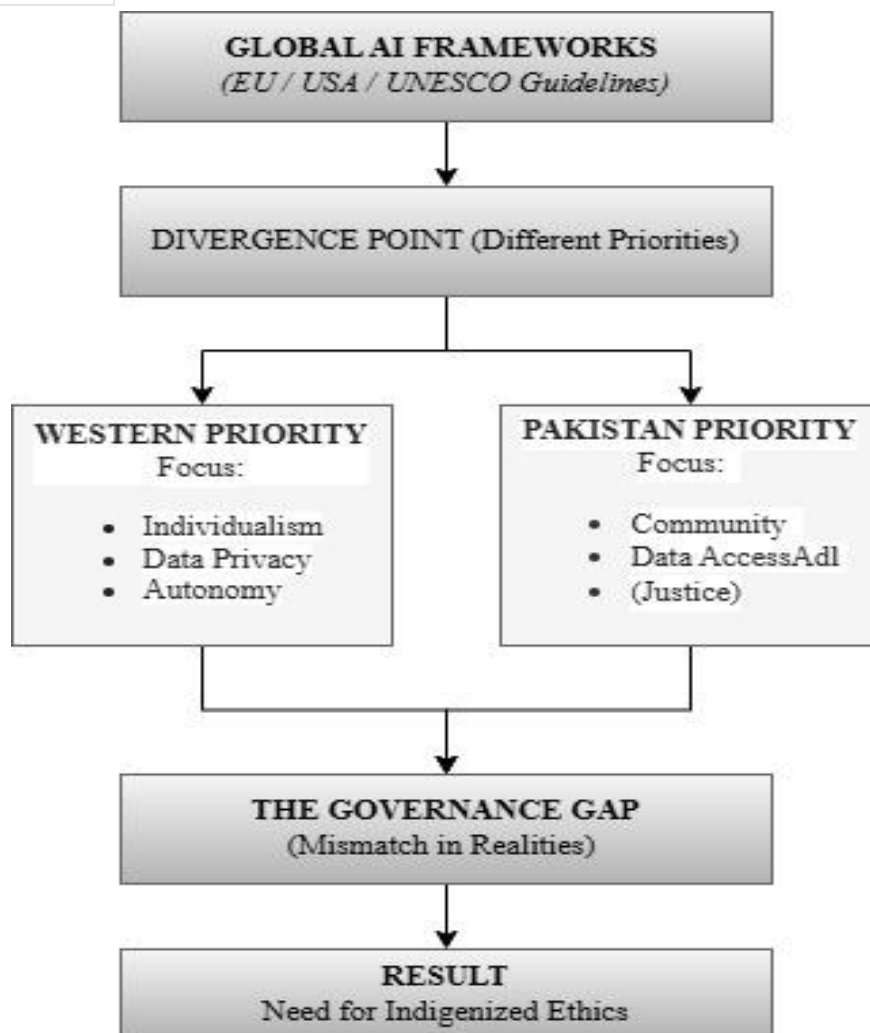


Fig. 2. The Ethical Divergence Map:

A visual depiction comparing Western priorities (Individualism) with Indigenized/Pakistani priorities (Community Welfare), and the aspect of the gap in governance.

5. The Proposed Adaptive Governance Model for Pakistan (Three-Pillar Model)

The combination of the global governance environment and the socio-technical gaps it entails results in the main findings of the present work the Three-Pillar Adaptive Governance Model. The model is aimed to transform Pakistan out of the state of Algorithmic Dependence to an epistemic Sovereignty through a balance between technical robustness and communal welfare.

It transcends the toothless normative rules of the past to offer a working enforceable model based on the Pakistani context.

5.1 Pillar 1: Regulatory Innovation and strategic directions of SMEs.

Instead of following the strict "Hard Law" frameworks of the Global North, this pillar aims at establishing an environment based on flexibility to enable local innovation without compromising safety requirements.

Regulatory Sandboxes: Sandboxes that allow startups to experiment with AI-based solutions like gamified educational applications without the immediate compliance burden or legal liability of high compliance costs and legal penalties.

Tiered Compliance: To create a risk-based liability system whereby compliance requirements rise corresponding to the size and the volume of the business without imposing an innovation tax on the local SMEs. **Strategic Incentives:** The government subsidies to award technical assistance to the SMEs to acquire international safety certifications and enable the products made locally to compete in the international market.

5.2 Pillar 2: Technical Sovereignty & Indigenous Data

To reduce the risks of Digital Colonialism, this pillar will emphasize building localized AI infrastructure which reflects the language and cultural specificities of Pakistan.

- **Urdu-Centric LLMs:** In contrast to English-centric methods, they should focus on collecting and enhancing quality and native Urdu material to develop models that reduce the impact of cultural hallucinations and address the local knowledge systems.
- **Data Sovereignty Protocols:** Moving to localized data storage and processing norms such that Pakistani citizen data is a communal resource instead of a resource to be excavated by technology giants around the world.
- **Algorithmic Auditing:** Bias and failure mode testing using localized, reasoning-conscious safety benchmarks to the Pakistani socio-technical direction.

5.3 Pillar 3: Ethical Oversight and the Trusteeship (I'timani) Paradigm.

This pillar substitutes Western procedural individualism with an Islamic Amanah (Divine Trust) based moral system, which makes the individual rights less significant and the communal protection more significant.

This pillar replaces Western procedural individualism with what the Islamic Amanah (Divine Trust) based moral system, reducing the importance of individual rights and increasing the importance of the communal protection.

- Trusteeship (I'timani) Council: Establishing a multi-stakeholder oversight agency that is independent and has technical specialists and ethicists to evaluate AI systems on ethical accountability.
- The Maslaha (Public Interest) Test: Finding out the positive impact of an AI system on the community as a legal measure to deploy to ensure technology is used in the public good.
- Adl (Absolute Justice) Filters: Adding measures of fairness to diminish the structural power distance between the creators and users of AI to equally grant access to and equally distribute the benefits of technology.

6. Analysis and Discussion

The section assesses the possible role of the Three-Pillar Adaptive Governance Model in meeting the systemic gaps that are revealed in world AI frameworks. The discussion will answer the primary research questions (RQs) that were outlined at the start of this study by synthesizing technical strength of the Global North to the ethical paradigms of the Global South.

6.1 Reducing the risks of GenAI with Technical Sovereignty (Addressing RQ1)

The major shortcoming of current frameworks (EU, UNESCO) is that they do not regulate the non-deterministic aspect of Large Language Models (LLMs) [4, 30].

- Beyond Transparency: Although the rules of 2019 were aimed at transparency, the Pillar 2 of the proposed model requires localized Algorithms Auditing to identify the cultural hallucinations that are particular to the Pakistani setting [30, 31].
- Rationality-conscious Safety: The framework enables attainment of a greater level of safety against automated disinformation and deepfakes by incorporating technical testing toolkits, which are based on the Singapore model but tailored to the Urdu-centric data [10, 30].

6.2 Reclaiming Data Sovereignty and Countering Colonialism (Addressing RQ2)

Digital Colonialism 2026 paradigm outlines the process of extracting data of the Global South to train models of Western standards [31].

- **Indigenous Infrastructure:** This is directly opposed by the model proposed, which obligates the use of Urdu-Centric LLMs and localized data storage protocols, such that the Pakistani data has become a communal asset [3, 31], instead of a commodity of international tech giants.
- **Epistemic Empowerment:** The framework avoids the so-called Algorithmic Dependence as local AI systems can be used to address the specific linguistic and socio-economic requirements of the community by focusing on its local knowledge systems [31].

6.3 Economic Viability and Ethical Accountability (Addressing RQ3)

This analysis demonstrates that Western compliance is inflexible, and it is an innovation tax on local developers [32].

- **SME-Centric Governance:** In Pillar 1, Pakistan SMEs can innovate without the crippling expenses of international "Hard Law" such as the EU AI Act when offered the functional alternative of Regulatory Sandboxes [6, 10, 32].
- **Moral vs. Procedural Ethics:** In contrast to the procedural rights of the individual of EU AI Act, the Trusteeship (I'timāni) Paradigm of Pillar 3 transfers responsibility to the developers of a moral Amanah [28, 29].
- **Structural Justice (Adl):** The implementation of Adl Filters will prevent the AI application to extend the already existing power imbalances and make governance a means of social provision instead of corporate domination [28, 29].

Table 5: Mapping Solution to Identified Gaps

Identified Challenge	Supporting Reference	Proposed Solution (Pillar)
LLM Hallucinations	Ferdaus [30]	Pillar 2: Localized Auditing
Data Extraction	Ahmed [31]	Pillar 2: Data Sovereignty
Compliance Cost	Holland [32], EU Act [6]	Pillar 1: Regulatory Sandboxes
Cultural Mismatch	Ali [28], Rahman [29]	Pillar 3: Trusteeship Paradigm
Identified Challenge	Supporting Reference	Proposed Solution (Pillar)

7. Findings and Strategic Recommendations

This paper evaluated the applicability of AI governance plans across the globe in the 2026 Generative AI generation in the developing world. The review confirms that western-biased policies tend to overlook the socio-economic fact, data paucity, and cultural sensibilities of the Global South. Moreover, models created in 2019-2021 do not have the systems to resolve contemporary Generative AI threats of hallucinations and deepfakes.

7.1 Key Findings

Based on the Comparative Gap Analysis and Triple-Lens Matrix, the following findings are consolidated, as follows:

- Compliance-Innovation Paradox: Hard Law-style solutions (e.g., the EU AI Act) lead to prohibitive compliance costs that put the formation of AI startups in developing nations such as Pakistan at risk [6, 32].
- The Representation Gap: Digital Colonialism The quality of Urdu datasets used to create a Global AI model is characterized by the absence of high-quality ones, which leads to culturally incorrect or offensive outputs [31].
- Governance Obsolescence: Generative AI is non-deterministic, which means that its use should be supported by more than the outdated ethical principles formulated in the 2019-2021 period [4, 5, 14].
- Need of Pluralist Ethics: There is a functional need to consider Islamic values like Adl (Justice) and Maslaha (Public Interest) in Muslim-majority nations in order to win the trust of the people and attain distributive justice [28, 29].

7.2 Strategy Recommendations for Pakistan

In order to fill these gaps, this paper provides a Three-Pillar Strategic Model of AI governance in Pakistan.

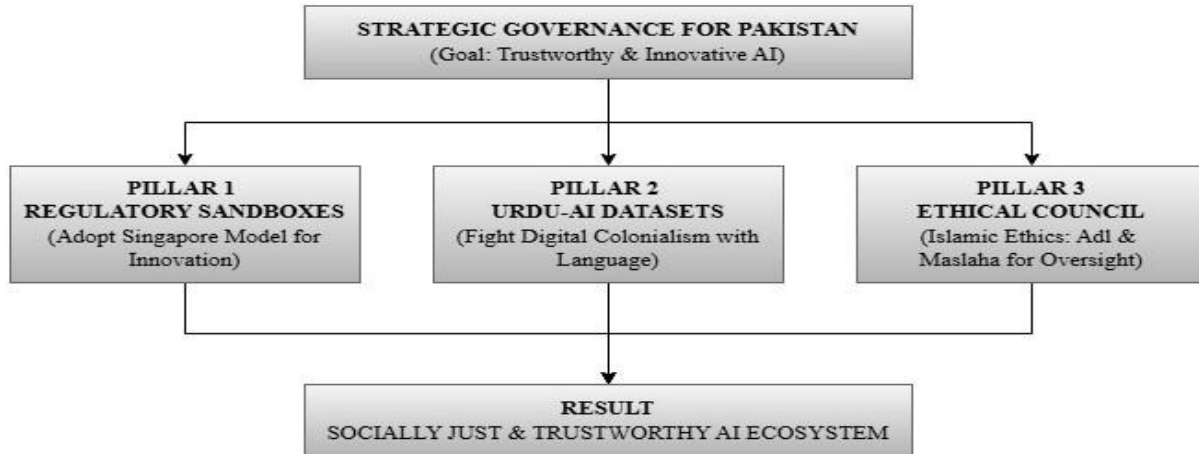


Fig. 3. The Proposed Strategic Governance Model of Pakistan

- **Pillar 1:** Adopt Regulatory Sandboxes (The Singapore Model): Pakistan should adopt government-regulated AI Sandboxes as a substitute to bans [10, 32]. This will enable startups to experiment on Gen-AI products without the risk of paying huge fines, a trade-off between innovation and safety. Pillar 2:
- **Pillar 2 Countering Digital Colonialism with Urdu-AI:** The Ministry of IT & Telecom (MoITT) should emphasize development of local Urdu datasets [31]. Pakistan can lessen the dependency of foreign-trained models by creating National Language Models (NLMs) at the university level.
- **Pillar 3:** The Adl Ethical Audit: A special advisory of technologists and Islamic scholars should be created with the purpose of overseeing the high-risk application of AI [28]. This renders AI systems audited to distributive justice and collective good [29].

7.3 Policy Impact and Global South Transferability

The framework is universal and can be used in other developing nations with majority of Muslims and where there are similar tensions between western technology and local values [31].

- **Short-term Impact:** The implementation of Sandboxes will have the direct impact of reducing entry barriers of local tech SMEs and result in the local AI application boom [32].

- **Long-term Consequence:** Indigenous datasets and Ethical Councils will ensure Data Sovereignty, ensuring that the digital future of the country will be placed under the authority of the local moral and cultural forces, and not the algorithms of other nations [31].

7.4 Future Research Directions

Further investigation of the future should consider encoding of ethics technically i.e. how to inject such things as Maslaha (Public Interest) into the reward functions or loss functions of Neural Networks. The environmental cost of optimization of large models of such energy-constrained countries is also a critical area of interest.

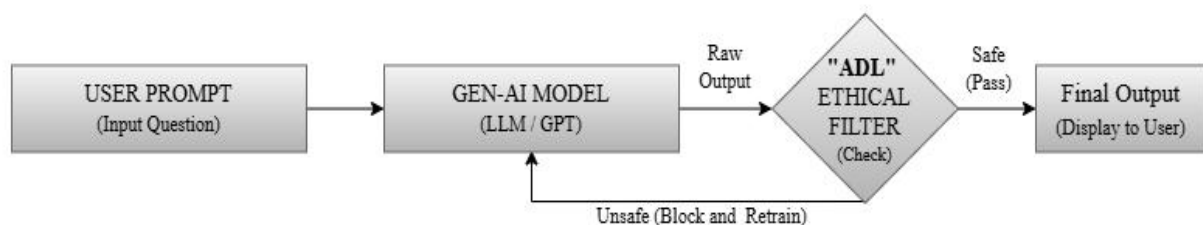


Fig. 4. Technical Architecture of the Adl Ethical Filter

7.5 Final Verdict

AI is not only a tool, but also a cultural phenomenon. The age of 2026 requires Pakistan not to copy and paste the foreign law; rather, it should pursue indigenized solutions. Based on technical auditing and the wisdom of the Islamic ethic, Pakistan may set the vision of the AI-driven future that will be more technologically advanced and acceptable by the society.

Conclusion

As it has been critically evaluated in this paper, the governance of AI has changed in 2026 towards underlying principles and the complex needs of the Generative AI generation. In this research, the conclusion is that global standards are presented as a normative framework, although the socio-economic and cultural unique aspects of the Global South and Pakistan in particular are not addressed [3].

It was discussed that current AI regulation has a gap in implementation that is caused by technological limitations of the old models [4, 30], local SMEs economic barriers [32], and institutional threat of Digital Colonialism [31].

These gaps can also be addressed with the help of the Three-Pillar Adaptive Governance Model proposed that can offer a practical framework to guide Pakistan towards Epistemic Sovereignty. Institutionalization of Regulatory Sandboxes, focusing on Urdu-centric datasets, codifying the Trusteeship (I'timania) Paradigm [28, 29, 31] culturally representative [28, 29, 31] can make Pakistan one of the first countries to build a technologically ambitious, economically inclusive and culturally reflective digital future.

References

1. A. Jobin, M. Ienca, and E. Vayena, "Artificial Intelligence: the global landscape of ethics guidelines", *Nat Mach Intell* 1, 389–399 (2019). <https://doi.org/10.1038/s42256-019-0088-2>.
2. T. Hagendorff, "The ethics of AI ethics: An evaluation of guidelines," *Minds and Machines*, vol. 30, no. 1, pp. 99–120, 2020. <https://doi.org/10.1007/s11023-020-09517-8>.
3. UNESCO, "Comparing Governance Mechanisms for AI: Hard Law vs. Soft Law," UNESCO Digital Library, Paris, 2025. [link](#)
4. L. Weidinger et al., "Ethical and social risks of harm from language models," *arXiv preprint, arXiv:2112.04359*, 2021.
5. A. Taeihagh, "Governance of generative AI," *Policy & Society*, vol. 44, no. 1, pp. 1–22, Jan. 2025, <https://doi.org/10.1093/polsoc/puaf001>.
6. ASEAN Secretariat, *Expanded ASEAN Guide on AI Governance and Ethics—Generative AI*. Jakarta, Indonesia: ASEAN, 2025.
7. A. Vijayakumar, *AI Ethics for the Global South: Perspectives, Practicalities, and India's Role*. New Delhi, India: Research and Information System for Developing Countries (RIS), 2024.
8. E. Elmahjub, "Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI," *Philosophy & Technology*, vol. 36, Art. no. 73, 2023, <https://doi.org/10.1007/s13347-023-00668-x>.

Online ISSN

Print ISSN

<http://amresearchreview.com/index.php/Journal/a>

3007-3197

3007-3189

9. E. Ormond, "Governance of AI ethics: Perspective from the Global South (Africa)," SSRN Electronic Journal, Feb. 23, 2023, Available at SSRN: <http://dx.doi.org/10.2139/ssrn.4368020>
10. IMDA Singapore, "Model AI Governance Framework for Generative AI," Infocomm Media Development Authority, Singapore, Jan. 2024.
11. J. G. Allen, J. Loo, and J. Luna, "Governing intelligence: Singapore's evolving AI governance framework," SMU Centre for Digital Law Research Paper, no. 1/2025, Oct. 2024, Available at SSRN: <http://dx.doi.org/10.2139/ssrn.5097834>
12. W. Hoffmann-Riem, "Artificial intelligence as a challenge for law and regulation," in *Regulating Artificial Intelligence*, Cham, Switzerland: Springer, 2019, pp. 1–29.
13. D. Li, "Governing machines, governed by giants: Mapping big tech's influence on global AI governance landscapes," *Chinese Public Administration Review*, Art. no. 15396754251359961, 2025.
14. P. R. A. Puchakayala, "Responsible AI: Ensuring ethical, transparent, and accountable artificial intelligence systems," *Journal of Computational Analysis and Applications*, vol. 30, no. 1, 2022.
15. Microsoft Corp., "Responsible AI: Ethical Policies and Practices," Microsoft AI Transparency Report, Redmond, WA, 2024.
16. Nishtiman Othman Mohammed. Ethical and legal challenges of ai in healthcare: A global perspective. *Int J Law Justice Jurisprudence* 2025;5(1):29-35, doi: <https://dx.doi.org/10.22271/2790-0673.2025.v5.i1a.160>.
17. S. Gerke, T. Minssen, and G. Cohen, "Ethical and legal challenges of artificial intelligence-driven healthcare," in *Artificial Intelligence in Healthcare*, 1st ed., 2020, pp. 295–336, doi: 10.1016/B978-0-12-818438-7.00012-5.
18. Lancet Digital Health, "Patient data privacy in the age of generative AI," *The Lancet*, vol. 6, no. 4, 2024.
19. UNESCO, "Generative AI in Education: The Ethics of Personalized Learning vs. Cheating," UNESCO Education Sector, Paris, 2024, [Link](#).

23. Future of Life Institute, "Governing AI Agents: Risks and Frameworks for Autonomous Action," FLI Reports, 2025.
24. I. D. Raji, A. Smart, R. N. White, M. Mitchell, T. Gebru, B. Hutchinson, et al., "Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing," in Proc. 2020 Conf. on Fairness, Accountability, and Transparency, 2020, pp. 33–44.
25. A. J. Mökander, J. Schuett, H. R. Kirk, and L. Floridi, "Auditing large language models: A three-layered approach," *AI and Ethics*, vol. 4, no. 4, pp. 1085–1115, 2024, <https://doi.org/10.1007/s43681-023-00289-2>.
26. Harvard Law Review, "The Black Box Problem: Liability for Autonomous Systems," Harvard Law Review, [Link](#).
27. T. Wang, Empowering Minds: A Round Table on Generative AI and Education in Asia-Pacific, 2023.
28. I. Pinho, A. P. Costa, and C. Pinho, "Generative AI governance model in educational research," *Frontiers in Education*, vol. 10, p. 1594343, Jul. 2025.
29. N. K. Corrêa, C. Galvão, J. W. Santos, C. Del Pino, E. P. Pinto, C. Barbosa, et al., "Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance," *Patterns*, vol. 4, no. 10, 2023.
30. V. Bolón-Canedo, L. Morán-Fernández, B. Cancela, and A. Alonso-Betanzos, "A review of green artificial intelligence: Towards a more sustainable future," *Neurocomputing*, vol. 599, p. 128096, 2024.
31. M. M. Ferdaus, M. Abdelguerfi, E. Loup, K. N. Niles, K. Pathak, and S. Sloan, "Towards trustworthy AI: A review of ethical and robust large language models," *ACM Computing Surveys*, vol. 58, no. 7, Art. no. 176, pp. 1–43, Jan 2026, doi: 10.1145/3777382.
32. A. S. A. Ahmed, "Algorithmic dependence and digital colonialism: A conceptual framework for artificial intelligence in education and knowledge systems of the Global South," *Frontiers in Education*, vol. 11, 2026, doi: 10.3389/feduc.2026.1720563.

33. F. Ali, K. Bouzoubaa, F. Gelli, *et al.*, "Islamic ethics and AI: An evaluation of existing approaches to AI using trusteeship ethics," *Philosophy & Technology*, vol. 38, Art. no. 120, 2025, doi: 10.1007/s13347-025-00922-4.
34. Rahman, Mohamad. "The Algorithmic Trust: Navigating the Ethical and Practical Dimensions of AI Business through an Islamic Lens".
35. J. Holland, "Artificial Intelligence Adoption for Sustainable Development in SMEs: Challenges, Strategic Pathways, and Lessons for Developing Economies", *IJSBE*, vol. 2, no. 1, Jan. 2026, doi: [10.51137/wrp.ijbsbe.502](https://doi.org/10.51137/wrp.ijbsbe.502).