

Article

Not peer-reviewed version

---

# A Mixed-Methods Governance Framework for Ethical Artificial Intelligence in Developing Public Sectors: Insights from Libya

---

Zirawani Binti Baharum and [Abdulbaset Salem Albaour](#)\*

Posted Date: 11 March 2026

doi: 10.20944/preprints202603.0838.v1

Keywords: artificial intelligence governance; ethical AI; public sector innovation; mixed-methods research; institutional readiness; fragile states; digital transformation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# A Mixed-Methods Governance Framework for Ethical Artificial Intelligence in Developing Public Sectors: Insights from Libya

Zirawani Binti Baharum and Abdulbaset Salem Albaour \*

University of Kuala Lumpur

\* Correspondence: [abdulbaset.albaour@s.unikl.edu.my](mailto:abdulbaset.albaour@s.unikl.edu.my)

## Abstract

The rapid diffusion of Artificial Intelligence (AI) across public sector institutions is reshaping governance practices and service delivery worldwide. However, most existing AI governance frameworks have been developed within technologically advanced and institutionally stable environments, limiting their applicability to developing and fragile states characterized by institutional volatility, regulatory gaps, and socio-political complexity. Addressing this gap, this study proposes an empirically grounded Ethical AI Governance Framework designed to support responsible AI adoption in public sector institutions operating within fragile governance contexts. The research adopts a sequential explanatory mixed-methods design, integrating quantitative and qualitative approaches to ensure methodological rigor and contextual validity. The quantitative phase involved a structured survey assessing organizational readiness for AI adoption, including dimensions of institutional capacity, technological infrastructure, governance preparedness, and stakeholder trust. The results provided baseline indicators of AI readiness across public sector entities. Building on these findings, the qualitative phase employed semi-structured interviews with policymakers, digital transformation experts, and institutional stakeholders to explore deeper governance challenges related to AI deployment, including ethical accountability, regulatory constraints, and institutional trust. The integration of both datasets enabled methodological triangulation and facilitated a comprehensive understanding of AI governance dynamics. Based on the empirical insights derived from this mixed-methods analysis, the study develops an Ethical AI Governance Framework structured around three interdependent pillars: Institutional Readiness, emphasizing organizational capacity, regulatory alignment, and technical infrastructure; Ethical Oversight, ensuring transparency, accountability, fairness, and responsible algorithmic governance; and Participatory Design, promoting citizen engagement, stakeholder inclusion, and trust-building mechanisms throughout the AI lifecycle. The proposed framework contributes to the emerging literature on AI governance in fragile and developing states by providing a context-sensitive governance model grounded in empirical evidence from the Libyan public sector. By bridging theoretical governance principles with practical policy considerations, the study offers a scalable blueprint for enabling responsible, trustworthy, and inclusive AI-driven digital transformation in complex institutional environments.

**Keywords:** artificial intelligence governance; ethical AI; public sector innovation; mixed-methods research; institutional readiness; fragile states; digital transformation

---

## 1. Introduction

### 1.1. Context and Problem

The global momentum behind AI in public service delivery is a key driver of digital transformation (DT). However, the underlying assumption of high institutional capacity for ethical

governance is often violated in fragile or developing countries [1,12]. The Libyan public sector, as the contextual core of the doctoral study, exemplifies an environment of profound institutional fragility, characterized by fragmented data infrastructure, a fractured legal environment, and compromised public trust following prolonged instability [10,11]. Implementing complex AI systems in such settings, without a tailored framework, risks accelerating existing political and ethical deficits [3,7]. The urgency of DT in government agencies globally is mostly influenced by the high rate of AI penetration into the system of service-delivery, policy-making, and management. With unprecedented efficiency improvements, amplified responsiveness, and insights based on data, AI technologies introduce themselves as a key lever to modernize governance which makes the implementation of AI technologies a significant value in the context of the modernization of contemporary governance. However, this global trend covers the fact that there is a long-standing issue: the belief that state institutions have enough resources, adequate regulatory frameworks, and a high level of trust among people who are required to ensure ethical and responsible AI control. The assumption is habitually abrogated in weak or underdeveloped situations, where the very fundamentals of good governance are naturally undermined. The Libyan state apparatus, which is the center of interest of this doctoral study, is an example of a highly institutionally weak environment. The history of political insecurity, civil war, and divided authority has developed a culture of a grossly weakened state capacity over the decades. This vulnerability manifests in a fragmented national data infrastructure, which disrupts the acquisition, standardization and safe management of information of value in AI applications. At the same time, the legal and policy landscape is still splintered, as it is characterized by the lack of unified regulations, unclear mandates, and weak enforcement procedures about emerging technologies. Above all, the long-term lack of stability has undermined the trust of citizens in government organizations, creating an atmosphere of distrust that negatively impacts the credibility and acceptance of any technological intervention initiated by the government. The blind implementation of advanced AI systems in high-risk and low-capacity settings, where a context-specific and empirically based governance system does not exist, is a high risk. Not only can such implementation fail to deliver its intended benefits, but it can also contribute to already existing political failures (e.g. by leading to opaque decision-making and power imbalances) and speed up ethical dilemmas (e.g. privacy violations, algorithmic bias with no appeals). In turn, these changes continue to jeopardize weak prospects of stability and equal development. Overcoming the acute lack of governance facing AI implementation in fragile situations is thus the immediate concern of this manuscript.

This study contributes to the literature in three ways. First, it provides one of the few empirical analyses of AI governance challenges in fragile states. Second, it introduces a mixed-methods validated governance framework tailored to low-capacity public sector environments. Third, it operationalizes ethical AI principles into actionable governance mechanisms applicable to developing countries.

### *1.2. Research Gap and Objective*

Current global AI governance frameworks (e.g., EU AI Act, UNESCO Principles), as critiqued in Manuscript 1, are largely normative and fail to offer operational guidance for low-resource, high-risk contexts where institutional stability cannot be assumed [1,2,12]. This creates a critical research gap: the absence of empirically validated, bottom-up governance models specifically designed for fragile states [10]. The objective of this Manuscript 2 is to address this gap by developing and justifying a data-grounded Ethical AI Governance Framework derived directly from the integrated quantitative and qualitative evidence of the PhD research [13].

### *1.3. Link to Thesis and Contribution*

This manuscript synthesizes the core methodological contribution of the thesis (Chapter VI). It translates the quantitative deficits in readiness and the rich thematic narratives on institutional and ethical barriers into a cohesive, generalized framework. The key contribution is the operationalized

framework, which acts as a practical methodological blueprint for integrating global AI ethics principles with the specific political and institutional realities of fragile states [5,6].

## 2. Methodological Rationale

### 2.1. Mixed-Methods Justification

A sequential explanatory mixed-methods design (Quan - Qual) was employed to fully capture the multidimensional nature of AI governance in Libya. This approach was necessary because quantitative data (Phase I) could measure the *extent* of readiness and trust, while the qualitative data (Phase II) was indispensable for *explaining* and *contextualizing* the underlying institutional and political *reasons* for these scores. This two-phased approach ensured the final governance framework was grounded in comprehensive stakeholder experience.

### 2.2. Phase I: Quantitative Component

The initial phase involved a comprehensive stakeholder survey across government employees, private-sector experts, and citizens (N=15) a small exploratory stakeholder sample. The survey measured perceived AI awareness, organizational readiness, and digital trust. Key descriptive findings revealed a critical tension: while overall public trust was moderate, perceived Institutional Readiness was strikingly low (M=1.60). This disparity highlighted the need for a deeper, qualitative investigation to explore the specific political and institutional factors contributing to this low readiness.

### 2.3. Phase II: Qualitative Component

The qualitative phase involved in-depth, semi-structured interviews analyzed via thematic analysis (Braun & Clarke method). This phase yielded rich narratives that clustered into critical themes, including Ethical Issues, Political and Institutional Fragmentation, and Skills Capacity Deficits. These thematic patterns were essential for refining the ethical and participatory components of the new governance framework.

### 2.4. Data Triangulation Strategy

The robustness of the framework relies on the formal triangulation of these two data types, as outlined in Table 1. The framework's components were directly shaped by the cross-synthesis of quantitative deficits (the *what*) and qualitative explanations (the *why*), ensuring the framework is both empirically sound and contextually relevant.

**Table 1. Mixed-Methods Design Summary.** The sequential explanatory approach grounds the framework in empirical evidence from the Libyan stakeholder context.

Phase	Methodology	Data Type	Purpose	Key Output
I	Quantitative Survey	Perception Scores	Assess readiness & trust across groups	Statistical profiles (means, correlations)
II	Qualitative Interviews	Rich Narratives	Explore barriers & ethical dimensions	Thematic patterns (7 key themes)
Integration	Cross-Synthesis	Mixed Evidence	Develop framework components & validity	<b>Ethical AI Governance Framework</b>

### 3. Framework Development Process

#### 3.1. Integrating Quantitative and Qualitative Findings

The framework emerged directly from the **meta-inferences** of data integration. The quantitative finding of low organizational readiness converged perfectly with the qualitative themes describing fragmented infrastructure and political uncertainty. This convergence dictated that Institutional Readiness must be the foundational pillar of the framework. Furthermore, the qualitative data's emphasis on citizen disenfranchisement and data privacy concerns directly informed the strength and scope of the Ethical Oversight and Participatory Design pillars.

#### 3.2. Modeling Approach

The modeling process involved synthesizing the integrated themes into three normative, non-negotiable categories for governance [6]. This translation ensured the conceptual validity of the final framework, moving from the purely descriptive empirical data to a practical prescriptive model for policy implementation.

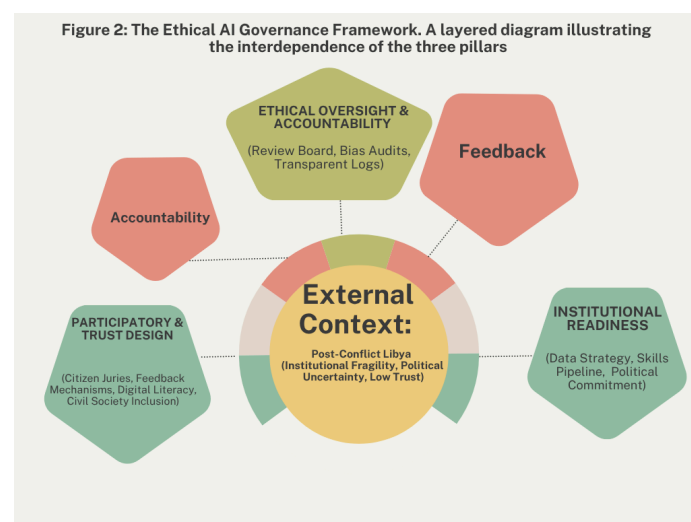
#### 3.3. Validation Through Triangulation

The framework's high validity stems from the empirical justification of every pillar. Its core concepts are not abstract; they are the direct conceptual solution to the specific deficits and challenges statistically measured and qualitatively explored within the Libyan public sector context [11].

## 4. The Ethical AI Governance Framework

#### 4.1. Framework Overview

The proposed Ethical AI Governance Framework is designed as a layered, dynamic structure (See Figure 2 below). It is anchored by the recognition of post-conflict constraints[10]. The model operates on three interconnected pillars that guide responsible implementation.



**Figure 1.** The Ethical AI Governance Framework.

A layered diagram illustrating the interdependence of the three pillars (Institutional Readiness, Ethical Oversight, Participatory Design), all operating within the external context of Post-Conflict Institutional Fragility.

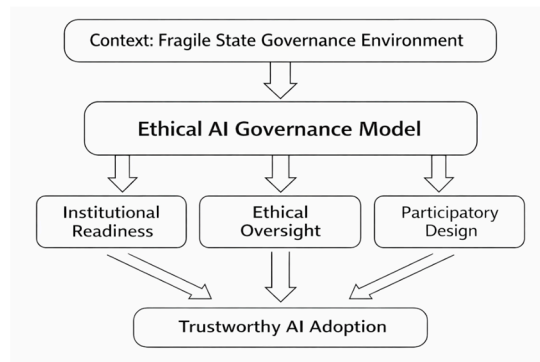
#### 4.2. Pillar Explanations

1. **Institutional Readiness:** (Addressing the Quantitative Deficit) This pillar focuses on structural and capacity-building requirements, grounded by the mean readiness score of 1.60.
  - *Key Components:* Implementation of a unified national data strategy (addressing fragmentation), a dedicated national digital skills pipeline for public servants, and securing high-level political commitment for inter-agency DT coordination.
2. **Ethical Oversight & Accountability:** (Addressing Qualitative Risk) This pillar ensures moral sustainability, derived from the thematic findings on human rights and risk of bias.
  - *Key Components:* Establishment of an independent Ethical Review Board (ERB) with multi-stakeholder representation, mandatory bias auditing and mitigation protocols for all high-risk AI applications, and clear, transparent decision-making logs.
3. **Participatory & Trust Design:** (Addressing the Trust Deficit) This pillar is rooted in the essential need to empower citizens and rebuild trust in public institutions.
  - *Key Components:* Mandatory inclusion of civil society groups in the design phase, establishing accessible citizen feedback and appeal mechanisms for algorithmic decisions, and prioritizing digital literacy programs to foster informed consent and trust.

#### 4.3. Integration Mechanism

The dynamic interaction among the three pillars is the mechanism for adaptive governance [7]. Ethical Oversight must continuously audit and feed back into Institutional Readiness (e.g., bias monitoring informs data standard updates). Crucially, Participatory Design functions as the accountability layer, ensuring that all policies and systems developed under the first two pillars are legitimate and responsive to citizen needs, thereby enhancing the overall resilience of the DT process.

Based on the integrated findings of the quantitative and qualitative phases, this study proposes an Ethical AI Governance Framework designed for fragile public sector environments. The framework conceptualizes responsible AI governance through three interdependent pillars: institutional readiness, ethical oversight, and participatory design. These components collectively enable trustworthy AI adoption within the constraints of fragile governance systems.



**Figure 2.** Ethical AI Governance Framework for Fragile Public Sector Environments.

The framework illustrates how AI governance in fragile states is structured around three interconnected pillars. Institutional readiness represents the foundational capacity required for AI deployment, ethical oversight ensures responsible and accountable use of AI technologies, while participatory design emphasizes citizen engagement and trust-building mechanisms. Together, these pillars enable the responsible and trustworthy adoption of AI in public sector institutions.

## 5. Theoretical and Policy Implications

### 5.1. Theoretical Validation and Extension

The framework aligns with human-centric AI models but significantly extends theory by operationalizing governance within a fragile-state context [6]. By defining and integrating Institutional Resilience and Ethical Viability as core governance dimensions, the model provides a necessary counter-theory to the stability assumptions of current global standards [12]. This theoretical contribution moves the field closer to developing genuinely universal and context-aware AI governance theory.

### 5.2. Novelty and Contribution

The primary novelty of this Manuscript 2 is the creation of a fully specified, mixed-methods-validated governance blueprint [13]. The framework is not an abstract theoretical model but a structured instrument whose components have been stress-tested against both statistical evidence and thematic insights from the ground. This offers a unique contribution to the sparse literature on AI policy in the developing world.

### 5.3. Policy Applicability

The framework provides immediate, actionable guidance for policymakers, not only in Libya but across other fragile states [10]. The phased structure, addressing foundational readiness first, then imposing ethical and participatory oversight-offers a prioritized implementation roadmap for public sector DT initiatives supported by development partners [12].

## 6. Conclusion

This paper adopts a sequential explanatory mixed-method design to elaborate and construct an Ethical AI Governance Framework systematically. Combining both quantitative measures of organizational preparedness and the qualitative studies on the obstacles of situation, the research offers an empirically based solution to the problem of AI usage in the context of the public sector [10]. The resulting model is anchored on three pillars that are interrelated namely: Institutional Readiness that focuses on core organizational and infrastructural potentials, Ethical Oversight that guarantees moral sustainability and accountability, and Participatory Design that incorporates stakeholder participation and building trust processes. These pillars are based directly on and certified by data gathered in the Libyan public sector. The framework represents a substantive input in that it closes the knowledge gap between abstract global ethics principles of AI and the operationalities of small-capacity governance settings [11]. It goes beyond prescriptive normativity to provide a situational and practical description of how to appropriately adopt AI. The framework is based on empirical evidence, which is why it can be used as a pragmatic model by the Libyan public sector and guide future research and implementation in similar settings of developing and post-conflict [13]. It introduces a reverse-thesis to the prevailing idea of institutional stability, highlighting that a fundamental preparedness, strict moral responsibility and genuine citizen engagement are essential requirements to deal with trustworthy AI. The overall quantitative study of the opinions of the stakeholders that will define a basic level of preparedness and trust is elaborated of this dissertation. provides very thematic stories based on qualitative interviews with in-depth insights into institutional, political, and ethical barriers faced by stakeholders in practice. Collectively, the above manuscripts provide solid empirical support, which underlies and builds upon the Ethical AI Governance Framework, which in turn supports its practical usefulness and theoretical soundness.

The framework, therefore, is illustrative of the potential of a mixed-method study to produce policy-receptive information in regard to intricate technological regulation dilemmas.

## References

1. P. Plantinga et al., "Responsible artificial intelligence in Africa: towards policy learning," *Data & Policy*, vol. 6, 2024.
2. V. Noordt, Colin, and G. Misuraca, "Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union," 10171st ed., 2022.
3. V. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York, NY: St. Martin's Press, 2018.
4. A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 389–399, 2019, doi: 10.1038/s42256-019-0088-2.
5. L. Floridi et al., "AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations," *Minds and Machines*, vol. 28, pp. 689–707, 2018, doi: 10.1007/s11023-018-9482-5.
6. J. COWLS et al., "Toward trustworthy AI: mechanisms for ethical governance of AI systems," *Nature Machine Intelligence*, vol. 3, pp. 522–532, 2021, doi: 10.1038/s42256-021-00396-x.
7. M. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, pp. 1–21, 2016, doi: 10.1177/2053951716679679.
8. J. Mišić, R. van Est, and L. Kool, "Good governance of public sector AI: a combined value framework for good order and a good society," *AI and Ethics*, vol. 7, pp. 1–16, 2025, doi: 10.1007/s43681-025-00751-3.
9. Rahmani, Ehsan, S. Zandi, and F. Kiani, "AI in Public Administration: Legal Implications for Automated Government Decision-Making," *Legal Stud. Digit. Age*, vol. 3.1, pp. 35-43, 2024.
10. Wynn, Martin, A. Bakeer, and Y. Forti, "E-government and digital transformation in Libyan local authorities," *Int. J. Teaching Case Stud.*, pp. 119-139, 2024.
11. Anshari, Muhammad, M. Hamdan, N. Ahmad, and E. Ali, "Public service delivery, artificial intelligence and the sustainable development goals: trends, evidence and complexities," *J. Sci. Technol. Policy Manage.*, vol. 16, no. 1, pp. 163-181, 2025.
12. Y.-F. Wang, Y.-C. Chen, and S.-Y. Chien, "Citizens' intention to follow recommendations from a government-supported AI-enabled system," *Public Policy and Administration*, vol. 40, no. 2, pp. 372–400, 2025.
13. A. Albaour, "A Relative Study of Artificial Intelligence Impact on the Digital Transformation of Government Services in Libya," *PhD Dissertation*, Universiti Utara Malaysia, 2024.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.