Available on: https://jsshrs.ep-journals.org/ E-ISSN: 3051-3561 P-ISSN: 3051-3553

Journal of Social Science and Human Research Studies

Volume: 01 Issue: 06 December-2025

Auditing Artificial Intelligence in Public Management: A Framework for Accountability, Transparency, and Ethical Governance

Juan Carlos Rodríguez Luna

EsSalud – Seguro Social de Salud, Perú, Universidad Continental, Peru

Published Date: December 22, 2025 **Article DOI: 10.65150/EP-jsshrs/V1E6/2025-04**

ABSTRACT: The integration of Artificial Intelligence (AI) in public management has transformed governmental operations, decision-making processes, and service delivery mechanisms. However, this technological adoption introduces unprecedented challenges in accountability, transparency, bias mitigation, and ethical governance. This article proposes a comprehensive framework for auditing AI systems in public sector contexts, addressing both technical and governance dimensions. Through systematic literature review and analysis of emerging regulatory frameworks, we identify critical audit domains including algorithmic transparency, data governance, decision explicability, bias detection, and compliance verification. The study examines case implementations in public procurement, social services, and administrative decision-making, drawing insights from international experiences and regulatory developments. Key findings reveal that effective AI auditing requires multi-dimensional approaches combining technical validation, ethical assessment, legal compliance verification, and stakeholder engagement. The proposed framework establishes methodological guidelines for internal auditors, external oversight bodies, and governance structures, emphasizing risk-based auditing strategies aligned with public sector accountability principles. This research contributes to the emerging discourse on AI governance in public administration by providing practical audit methodologies adaptable to diverse governmental contexts while maintaining democratic values and citizen protection standards.

KEYWORDS: Artificial Intelligence Auditing, Public Management, Algorithmic Accountability, AI Governance, Public Sector Technology, Transparency

1. INTRODUCTION

The proliferation of Artificial Intelligence (AI) technologies in public management represents one of the most significant transformations in contemporary governance. Governments worldwide are deploying AI systems across diverse functions including public procurement evaluation, benefit eligibility determination, risk assessment for regulatory compliance, predictive policing, resource allocation, and citizen service delivery (Wirtz et al., 2019; Veale et al., 2018). These implementations promise enhanced efficiency, data-driven decision-making, cost reduction, and improved service quality. However, the opacity of AI algorithms, potential for embedded biases, accountability gaps, and limited explainability raise fundamental questions about democratic governance, citizen rights, and public trust (O'Neil, 2016; Eubanks, 2018).

Recent high-profile cases have demonstrated the risks of inadequately governed AI in public sector applications. Algorithmic systems deployed for welfare benefit assessments have been found to perpetuate discriminatory outcomes, predictive policing algorithms have reinforced racial profiling patterns, and automated decision systems have produced unexplained denials of citizen applications (Eubanks, 2018; Richardson et al., 2019).

This article develops a comprehensive framework for auditing AI systems deployed in public management contexts. The research objectives include: (1) identifying critical dimensions requiring audit scrutiny; (2) proposing methodological approaches for technical and governance auditing; (3) examining regulatory frameworks and compliance requirements; (4) analyzing practical implementation challenges and risk mitigation strategies; and (5) establishing guidelines adaptable to diverse governmental contexts.

2. METHODOLOGY

This research employs a multi-method approach combining systematic literature review, comparative regulatory analysis, and framework development.

2.1 Literature Review Protocol

A systematic literature review was conducted following PRISMA guidelines covering publications from January 2015 to December 2024. The initial search yielded 847 potentially relevant sources. After screening, final inclusion involved 127 sources demonstrating direct relevance to AI auditing frameworks.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

2.2 Comparative Regulatory Analysis

Regulatory frameworks were analyzed across multiple jurisdictions including the European Union, United Kingdom, United States, Canada, Peru (Law 32069), and emerging Latin American frameworks.

3. THEORETICAL FOUNDATIONS

Understanding AI system characteristics is fundamental to developing appropriate audit methodologies. AI technologies exhibit distinct properties demanding specialized scrutiny including algorithmic opacity, dynamic learning, data dependencies, and probabilistic decision-making.

4. CRITICAL AUDIT DIMENSIONS

Based on literature synthesis and regulatory analysis, we identify six critical dimensions: (1) Algorithmic fairness and bias detection; (2) Legal compliance and regulatory conformity; (3) Transparency and explainability; (4) Accountability structures and human oversight; (5) Data governance and privacy protection; and (6) Performance monitoring and risk management.

5. PROPOSED AUDIT FRAMEWORK

The comprehensive framework is structured in five phases: (1) Pre-implementation assessment evaluating purpose legitimacy and impact; (2) Technical validation examining algorithms and training data; (3) Governance compliance verification assessing legal compliance and accountability; (4) Operational monitoring tracking performance and bias; and (5) Periodic comprehensive review conducting holistic reassessment.

6. IMPLEMENTATION CHALLENGES

Practical implementation encounters multiple obstacles including technical complexity exceeding auditor expertise, proprietary algorithm opacity limiting access, resource constraints, rapidly evolving technology, and organizational resistance. Strategic mitigation includes specialized training, contractual disclosure requirements, risk-based prioritization, and flexible framework design.

7. CASE ANALYSIS

Framework application is illustrated in three public sector domains: (1) Public procurement evaluation systems where Peru's Law 32069 establishes comprehensive regulations; (2) Social benefit eligibility determination involving fairness analysis and due process compliance; and (3) Regulatory compliance risk assessment requiring calibration and fairness verification.

8. DISCUSSION

The analysis reveals fundamental tensions between AI technological capabilities and public sector accountability requirements. Key governance implications include redefining transparency in algorithmic contexts, balancing innovation and precaution through risk-based approaches, institutional capacity development, dynamic governance frameworks, and ethical dimensions beyond legal compliance.

9. CONCLUSIONS

This article developed a comprehensive framework for auditing AI systems in public sector contexts. Key conclusions include: First, effective AI auditing requires multi-dimensional approaches integrating technical analysis, governance assessment, ethical evaluation, and stakeholder engagement. Second, AI's distinctive characteristics demand specialized audit methodologies. Third, risk-based prioritization is essential. Fourth, transparency requirements must be reframed for AI contexts emphasizing outcome-oriented accountability. Fifth, clear accountability structures maintaining human authority remain essential. Sixth, continuous monitoring supplements periodic comprehensive audits. Seventh, institutional capacity development represents an essential prerequisite.

The proposed framework offers adaptable guidance applicable across diverse governmental contexts. As AI deployment accelerates, robust auditing mechanisms become increasingly critical for ensuring technological advancement aligns with democratic values, legal requirements, and public interest.

REFERENCES

1) Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society, 20(3), 973-989. https://doi.org/10.1177/1461444816676645
2) Arrieta, A. B., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012
3) Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104(3), 671-732. https://doi.org/10.15779/Z38BG31
4) Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society, 3(1), 1-12. https://doi.org/10.1177/2053951715622512
5) Coglianese, C., & Lehr, D. (2019). Transparency and algorithmic governance. Administrative Law Review, 71(1), 1-56.
6) Diakopoulos, N. (2016). Accountability in algorithmic decision making. Communications of the ACM, 59(2), 56-62. https://doi.org/10.1145/2844110
7) Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/

8) Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1
9) Government of Canada. (2021). Algorithmic Impact Assessment Tool. Treasury Board of Canada Secretariat.
10) Kearns, M., & Roth, A. (2019). The ethical algorithm: The science of socially aware algorithm design. Oxford University Press.
11) Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. Philosophical Transactions of the Royal Society A, 376(2133), 20180089.
12) O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown Publishing Group.
13) Richardson, R., Schultz, J., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. New York University Law Review Online, 94, 192-233.
14) Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1-14.
15) Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—Applications and challenges. International Journal of Public Administration, 42(7), 596-615.

AUTHOR BIOGRAPHY

CPC Juan Carlos Rodríguez Luna holds an MBA specializing in Public Procurement with over 20 years of experience in Peru's public sector. He currently serves at EsSalud (Peru's Social Security System) and teaches public procurement at Universidad Continental. His research focuses on risk management in public contracting, artificial intelligence applications in government procurement, and regulatory compliance frameworks. He is a member of the Centro Internacional de Investigación y Desarrollo Latinoamericano (CIID) and author of "Gestión de riesgos en la contratación pública: Claves para una ejecución efectiva" (2025).

License: This is an open access article under the CC BY 4.0 license: https://creativecommons.org/licenses/by/4.0/