Nations International Journal of Multidisciplinary Studies, 2026, 1(2), 48–62
https://research.nations.gy/index.php/nijms
ISSN Print: 3005-656X
ISSN Online: 3104-7246

Artificial Intelligence, Ethics, and Accountability in Policing: A Systematic Review of Governance Frameworks and Operational Risks

Nicola D. Kendall

The Guyana Police Force Academy, Georgetown, Guyana
Email: nicolakendall82@gmail.com

How to cite this paper: Kendall, N. D., (2026). Artificial intelligence, ethics, and accountability in policing: A systematic review of governance frameworks and operational risks, *Nations International Journal of Multidisciplinary Studies, 1*(2), 48–62.
https://doi.org/10.59607/nuijms.v1i2.20

Received: March 29, 2026
Accepted: May 28, 2026
Published: June 30, 2026

Abstract

The increasing use of artificial intelligence (AI) in policing encompasses predictive analytics, surveillance systems, facial recognition, and other data-driven decision-support tools. While these technologies might facilitate operational efficiency and enhance crime-prevention capabilities, they also raise serious ethical and accountability issues. This study adopted a qualitative secondary desk study, using a structured, systematic literature review to examine the intersection of AI, ethics, and accountability in policing. The review synthesized evidence from peer-reviewed journal articles, institutional reports, and governance frameworks published between 2015–2025. The findings are organized into four major thematic areas: bias and discrimination in algorithms, transparency and explainability issues, privacy and surveillance risks, and gaps in accountability and oversight. The results indicated that policing AI exhibits the capacity to reproduce historical bias inherent in training data, serve as black boxes for decision-making processes and widen surveillance capabilities in ways that meaningfully threaten privacy and civil liberties. The study also found a chronic divide between high-level ethical values and their operationalization in policing institutions. This paper also contributed to the existing body of knowledge by integrating ethics and governance into a single analytic framework, before outlining implications for police legitimacy, institutional readiness, and responsible AI adoption, particularly in developing contexts.

Keywords

Artificial Intelligence in Policing; Ethics; Accountability; Predictive Policing; Governance

1. Introduction

This study defines artificial intelligence (AI) as a core transformative technology that reshapes the logics of global governance, public administration, and public decision-making in the 21st century. At present, the core goal of AI tool deployment by governments and public sectors across all countries is to improve the efficiency of public services, optimize the allocation of administrative resources, and enhance the precision of public decision-making. However, both the European Commission's 2020 report and the OECD's 2019 study point out that the rapid popularization of AI in the public sphere has also given rise to cross-sectoral universal ethical and legal governance challenges. These include accountability dilemmas stemming from a lack of fairness in algorithmic decision-making, opaque operational logics, and blurred responsibility boundaries, among other issues.

This study then narrows its analytical focus to policing and identifies the specific application scenarios of AI in this sector. Peer studies by Brayne (2017) and Perry et al. (2013) both confirm that the spread of policing AI has driven the global policing model to shift from a traditional passive response to a data-centered, proactive, intelligence-led model. The most widely used application, predictive policing, realizes the pre-deployment of policing resources by modeling and analyzing historical policing data to forecast potential risk events and high-incidence areas.

However, a 2016 study by Lum and Isaac simultaneously points out that the current deployment of policing AI faces two core social and institutional challenges. The first is algorithmic bias: historical law enforcement deviations embedded in training data will continuously amplify discrimination against marginalized groups. The second is insufficient algorithmic transparency and interpretability: the black-box nature of most commercial policing algorithms means that neither frontline law enforcement officers nor the public can understand the logic behind algorithmic decisions. The core problem of the lack of accountability mechanisms, which is directly linked to transparency, will be specifically unpacked and analyzed in the subsequent chapters of this study.

When AI-assisted policing decisions produce harmful, discriminatory, or even illegal outcomes, the dilemma of assigning liability remains an unresolved core challenge in this field. Prior studies by Diakopoulos (2016) and Kroll et al. (2017) both note that this challenge stems from the fragmentation of liability among multiple stakeholders, including software developers, technology vendors, and policing agencies, making it impossible to draw clear boundaries for holding parties responsible. Today, leading international frameworks such as the OECD Principles on Artificial Intelligence and the UNESCO Recommendation on the Ethics of Artificial Intelligence all prioritize the core values of accountability, transparency, and human oversight. However, implementation mechanisms for these frameworks remain far from mature in policing contexts across most jurisdictions worldwide.

Existing research on policing AI also has two critical limitations. First, it treats four core themes: predictive policing, algorithmic bias, surveillance technology, and AI governance, as independent, isolated topics for separate analysis. Second, its silos have ethical risks, accountability mechanisms, and governance structures, lacking an integrated analytical framework that covers the full operational chain. This gap is particularly acute in developing regions, where technological innovation often outpaces the evolution of regulatory frameworks, oversight mechanisms, and organizational capacity.

To address this unmet need, this study adopts a systematic literature review to fill this gap, focusing on the context of developing countries. Centering ethics, accountability, and governance as its core research dimensions, the study sets four objectives: to identify the main applications of policing AI, analyze related ethical risks, evaluate existing accountability and governance mechanisms, and explore their impacts on police legitimacy and institutional readiness. The rest of this paper is organized into the following sections: literature review, research methods, study findings, discussion of results, conclusions, future research directions, and policy recommendations.

2. Literature Review

2.1. AI Applications in Policing

AI in policing is typically instantiated through the following modalities: predictive policing, surveillance systems, facial recognition, and automated analysis systems. Predictive policing tools use historical and real-time data to predict where crimes are likely to occur, identify potential victims or offenders (see risk assessments), and inform patrol deployment. RAND's original framework illustrates how these systems direct resources and inform tactical decisions rather than providing forward intelligence (Perry et al., 2013).

Further, AI takes prediction-enhancing integrated practices of surveillance. Generalizing from this literature, we are entering the era of 'big data policing', in which officers and agencies combine criminal justice databases with administrative and digital traces to support monitoring, targeting, and risk scoring. Such developments increase the scope of information available to police institutions and further entrench analytics into quotidian practice (Brayne, 2017; Ferguson, 2017).

These capabilities are further broadened by facial recognition or other biometric tools that provide identification and matching at scale. Although such tools are correctly promoted as being less intensive and preventative, the literature warns that operational performance is inextricably tied to legal protections, data integrity, and institutional oversight. Hence, AI in policing should be understood as a socio-technical assemblage of software, data infrastructures, organizational routines, and governance systems rather than an independent technological solution (Brayne, 2017; Lum, Koper & Willis, 2017).

2.2. Ethical Risks and Bias

Arguably, a prominent ethics issue in literature on AI in policing is algorithmic bias. Research on predictive policing has shown that systems trained on arrest or enforcement data can simply replicate previous institutional patterns, such as the over-policing of already disenfranchised groups. However, in that way, while AI may predict crime, it also predicts patterns of past policing activities (Lum & Isaac, 2016; Richardson et al., 2019).

Discrimination and inequality in literature are also closely related to this. Ethical issues arise not just in cases of deliberate bias, but in those that might involve neutral-seeming models built on corrupted proxies, bad data, or an infraction history riddled with inequity. Such limitations can produce different outcomes even in the absence of discriminatory design. These risks are especially acute in policing, where interventions bear on liberty and dignity; mobility; and non-citizen trust- the constant hope that bestows standing upon account (O'Neil, 2016; Benjamin, 2019).

The third significant ethical issue involves privacy risks associated with data integration. As policing systems combine multiple data integrations, police institutions improve their capabilities to collect data, derive information about people and communities, and act. The literature frames this issue not only in terms of technical data governance but also as a matter of rights, proportionality, and democratic control (Brayne, 2017; Council of Europe, 2019).

2.3. Transparency and Explainability

Ethical AI governance is grounded in transparency and explainability, since real people - the ones who should be protected from its harm, must understand how consequential decisions are made; regulators overseeing the systems need to operate at arm's length from this complexity. However, the literature suggests that machine learning systems are arguably ipso facto opaque given that technical complexity, proprietary espionage, and institutional barriers to disclosure create their own black boxes (Burrell, 2016; Pasquale, 2015).

Such opacity is especially troublesome in policing, as decision-support tools can be used to determine patrol routes, suspicion, case prioritization, and the intensity of surveillance. Without interpretability, independent auditing becomes impossible; courts, oversight agencies, and members of the public are disempowered to challenge bad outcomes. Henceforth, it is important to distinguish between superficial transparency and meaningful explainability. Therefore, revealing the existence of a system is not sufficient for accountability or governability, because systems must be sufficiently interpretable to enable informed and accurate oversight and evaluation. (Ananny & Crawford, 2018; Doshi-Velez & Kim, 2017).

2.4. Accountability and Governance

Accountability in AI encompasses answerability, auditability, reviewability, and enforceable mechanisms of responsibility that operate throughout the entire lifecycle of AI systems. These mechanisms extend across all stages of development and use, including design, procurement, testing, deployment, monitoring, and redress processes (Kroll et al., 2017; Diakopoulos, 2016). Principles at a high level, internationally consistent standards require transparency, fairness, accountability, and human-centered oversight. Several increasingly authoritative initiatives on the governance of trustworthy AI, such as OECD's AI Principles and UNESCO's Recommendation on the Ethics of Artificial Intelligence, have proven that human rights, fairness, and accountability are at the heart of AI governance (OECD, 2019; UNESCO, 2021).

However, the literature shows that mechanisms of institutional oversight are often slow to adapt to the pace of technological change. Law enforcement agencies may use tools without adequate audit trails, independent validation, public reporting, or reliable procedures for challenging outcomes. As a result, the reality of governance is often reactive and sporadic rather than proactive and systematic (Raji et al., 2020; Veale et al., 2018).

2.5. Gaps in Literature

The literature identifies several gaps. To begin with, regulatory frameworks are still patchy and tend to be too broad-brushed to address operational realities in policing. Second, even after releasing numerous calls for explainability and transparency, the explainability problem persists. Third, a wide range of governance principles remains aspirational and has not been translated into specific institutional controls. Fourth, there is a prevalence of inadequate empirical evidence suggesting that technology can address intractable collective social and crime issues without perpetuating existing structural harm, often resulting in excessive over-belief (Leslie, 2019; Floridi et al., 2018; Mittelstadt et al., 2016).

Several of these gaps underpin the rationale for an integrated review that considers ethics, accountability, and governance within a single analytical framework, given that governance structures have historically failed to keep pace with the rapid adoption of new technologies in public institutions (Winfield & Jirotka, 2018).

3. Materials and Methods

3.1. Research Design

This study utilized a qualitative secondary desk research design based on a structured systematic literature review. The systematic review was chosen as the methodology because it provides a clear, valid, and reproducible method for synthesizing literature dispersed across several knowledge domains (law, ethics, criminology, public policy, and information systems) (Snyder, 2019; Tranfield et al., 2003). This design was particularly appropriate given the interdisciplinary nature of AI in policing, which requires integrating conceptual, empirical, and governance-based resources.

3.2. Data Sources

The study encompasses a wide range of data sources, including influence reports, peer-reviewed journal articles, scholarly legal analyses, conference papers, institutional documents, and international governance documents. Academic literature was identified from searches through Scopus, Web of Science, and Google Scholar to identify core academic materials; policy and governance materials were sourced from authoritative institutional sources, including the OECD, UNESCO, and the European Commission (Kitchenham et al., 2009; Okoli, 2015). Such an academic and institutional mix was required, as AI in policing research is influenced by empirical scholarship as well as by new governance norms, regulatory discussions, and policy agendas about public-sector agencies (OECD, 2019; UNESCO, 2021).

3.3. Search Strategy

Searches were conducted using combinations of the following terms: “AI in policing”, “artificial intelligence policing ethics”, “predictive policing ethics”, “facial recognition policing accountability,” and “algorithmic bias law enforcement, AI accountability law enforcement ". Governance, surveillance, and explainability issues were included in combinations as necessary. The search and screening process adhered to the principles of systematic review reporting logic in line with the transparency embodied in PRISMA (Moher et al., 2009; Page et al., 2021). This technique improved methodological transparency and enhanced the reproducibility of the search process.

3.4. Inclusion and Exclusion Criteria

All sources used in this study were published between 2015 and 2025, were peer-reviewed or produced by reputable institutions, and had direct ties to policing, law enforcement, or governance issues most relevant to policing. Pre-2015 seminal sources were employed as needed for conceptual or operational background, only in instances where the relevant attached guidance itself called for their inclusion (for instance, foundational work on predictive policing (Perry et al., 2013). Sources were excluded from the final review if they lacked sufficient relevance to policing, were duplicated within the identified thematic areas, or made an insufficiently authoritative contribution to the analysis of ethics, governance, or accountability. Setting these criteria at the outset also helped to minimize selection bias and enhance consistency across reactions (Petticrew & Roberts, 2006; Booth et al., 2016).

3.5. Analytical Approach

The analysis was undertaken using a combination of thematic coding and cross-study comparison. In line with the analytical framework, initial coding was conducted across four broad thematic areas: algorithmic bias and discrimination; challenges to transparency and explainability; privacy and surveillance risks; accountability and oversight gaps. Evidence from the included studies was extracted descriptively and subsequently grouped into more aggregate themes for interpretative purposes within the discussion section (Braun & Clarke, 2006). This was not only appropriate, as it allowed the detection of patterns across different source types, but also more iterative, as it provided conceptual depth and analytical coherence.

3.6. PRISMA Screening and Study Selection

This study presents a systematic review of the themes of artificial intelligence (AI) ethics and accountability in policing. To improve the transparency and reproducibility of the research, we completed the study design in compliance with the PRISMA 2020 framework proposed by Page et al. in 2021. We searched three major databases: Scopus, Web of Science, and Google Scholar, using six sets of search terms closely aligned with the core issues of AI ethics, accountability, and algorithmic bias in policing contexts. The initial search yielded 250 academic works, which underwent multiple rounds of screening: 40 duplicates were removed, leaving 210 works to enter the title and abstract screening stage. After excluding 140 works with insufficient thematic relevance, 70 works remained for full-text evaluation. We then eliminated 28 unqualified works that only focused on technical development, lacked governance analysis, or were thematically duplicated, and finally included 42 works from diverse sources. The full process is summarized in Figure 1.

Figure 1

PRISMA 2020 Flow Diagram

IDENTIFICATION

Records identified through database searching

(n = 250)

|
▼

Duplicate records removed

(n = 40)

|
▼

SCREENING

Records screened

(n = 210)

Records excluded

(n = 140)

ELIGIBILITY

Full-text articles assessed for eligibility

(n = 70)

Full-text articles excluded

(n = 28)

Reasons:

- Not directly related to policing (n = 12)
- Focused primarily on technical AI development (n = 7)
- Limited governance or accountability analysis (n = 6)
- Duplicate thematic contribution (n = 3)

INCLUDED

Studies included in final review (n = 42)

3.7. Ethical Considerations

This study did not involve human participants. All data were derived from publicly available academic, legal, policy, and institutional sources. Consequently, the primary ethical considerations concern appropriate sourcing and attribution, accurate representation of the literature, and responsible synthesis of existing knowledge, rather than compliance with research ethics protocols governing human subjects (Resnik, 2018; Wiles, 2013). Throughout the review process, care was taken to avoid overstating findings or drawing conclusions beyond what the evidence supported. The analysis remained faithful to the original arguments, findings, and limitations reported in the reviewed sources, thereby promoting scholarly integrity and methodological rigor.

4. Findings

4.1. Theme 1: Algorithmic Bias and Discrimination

The literature review reveals normatively that predictive policing systems can produce biased historical policing patterns. Research has shown different things; for example, when models are trained on police data, the outputs reinforce prior distributions of arrests and patrols rather than serving as a separate, independent measure of underlying offending (Lum & Isaac, 2016; Richardson et al., 2019). A common result is situational feedback loops, whereby repeated deployments to the same area led to more incidents recorded there and thus further deployment (Ensign et al., 2018). This suggests that, across the studies surveyed, algorithmic bias should not be viewed as an isolated technical failure but rather as recurrent outputs connected to past critical data and the systemic reinforcement of outcomes.

4.2. Theme 2: Challenges Related to Transparency and Explainability

The literature highlights that AI tools in high-stakes settings often continue to be opaque or un-auditable technical opacity, institutional opacity, and the limits of transparency are characteristics that frequently emerge in studies reviewed (Burrell, 2016; Ananny & Crawford, 2018). Even when systems are made known, external reviewers may struggle to understand how outputs are generated or how (and with what) model logic should be rebutted in the field. Consequently, the studies show that transparency does not necessarily lead to accountability, especially in significant decision-making settings (e.g., criminal justice and policing) where black-box models are utilized as they lack explainability (Ananny & Crawford, 2018; Rudin, 2019).

4.3. Theme 3: Privacy and Surveillance Risks Associated with the Use of AI in Law Enforcement.

The literature reviewed indicates that AI-enabled policing is increasingly characterized by integrated surveillance systems that combine automated monitoring of public and private spaces with data derived from multiple sources. The adoption of big-data policing has expanded law enforcement agencies' capacity to collect, aggregate, and analyze vast datasets, enabling the generation of behavioral insights and risk assessments through advanced analytical techniques (Brayne, 2017). These developments have significantly broadened the operational and informational capabilities available to police institutions.

However, the literature also highlights that the expansion of surveillance capabilities carries substantial privacy implications. The integration of AI-driven analytics into routine policing practices facilitates more extensive monitoring of individuals and communities, raising concerns about proportionality, civil liberties, and the protection of personal data (Brayne, 2017; Perry et al., 2013). Furthermore, the normalization of surveillance through the routine use of predictive and analytical technologies may alter decision-making processes within law enforcement agencies and increase the risk of intrusive oversight of citizens' activities. Consequently, privacy risks extend beyond the mere collection of data and encompass the broader institutionalization of data analytics and surveillance technologies within everyday policing operations

4.4. Theme 4: The Lack of Accountability and Oversight

There are persistent accountability gaps across the AI lifecycle in the literature. A recurring theme across reviewed sources is the identification of issues such as the absence of audit trails, limited institutional oversight, unclear assignment of responsibility, and limited lifecycle governance (Kroll et al., 2017; Raji et al., 2020; Veale et al., 2018). More evidence for accountability being notably present when systems are used alongside documentation, testing, monitoring, and review structures. However, it is questionable whether many deployed systems are found within the cohesive governance arrangements described (Raji et al., 2020). Institutional design and governance practice — rather than the simple use of AI technology — therefore inform how accountability is defined in the reviewed literature.

Table 1

Thematic Summary of Results

Theme	Core findings from the reviewed studies	Representative sources
Algorithmic bias and discrimination	Predictive systems can reproduce biased policing patterns and generate reinforcing feedback loops.	Lum and Isaac (2016); Richardson, Schultz & Crawford (2019); Ensign et al. (2018)
Transparency and explainability challenges	Many systems remain opaque, and transparency alone is often insufficient for meaningful accountability.	Burrell (2016); Ananny & Crawford (2018); Rudin (2019)
Privacy and surveillance risks	Big-data policing expands surveillance capacity and increases privacy concerns.	Brayne (2017); Perry et al. (2013)
Accountability and oversight gaps	Effective governance requires structured auditing and clearer accountability mechanisms across the AI lifecycle.	Kroll et al. (2017); Raji et al. (2020); Veale et al. (2018)

Source: Author's synthesis of reviewed literature.

Table 2

AI Risks vs Accountability Measures

AI application	Ethical risk	Accountability mechanism
Predictive policing	Bias, discriminatory targeting, feedback loops	Dataset auditing, impact assessments, independent validation, periodic review
Facial recognition	Misidentification, privacy intrusion, disproportionality	Legal thresholds, judicial oversight, accuracy testing, public reporting
Surveillance analytics	Privacy erosion, function creep, over-monitoring	Data governance rules, retention limits, authorization procedures, and external oversight
Risk scoring / decision-support tools	Opacity, unfair outcomes, unclear responsibility	Explainability standards, documentation, human-in-the-loop review, and appeal mechanisms

Source: Synthesized from the reviewed literature on algorithmic bias, transparency, surveillance, and accountability frameworks.

5. Discussion

5.1. Bias and Structural Inequality Interpretation

The results indicate that the bias of police AI systems is not simply a technical problem but rather a matter of structural governance. When police data show unequal histories of enforcement, these same data can structure AI systems that inherently perpetuate (and even mask) inequality (Barocas & Selbst, 2016; Benjamin, 2019; Eubanks, 2018). This is important for policing legitimacy because communities are not going to see a system that seems technologically advanced but just replicates familiar exclusionary patterns and disproportionality as worthy of their trust. This evidence shows that responsible AI governance cannot be limited to downstream model performance but must engage with upstream social and institutional conditions (Benjamin, 2019; Eubanks, 2018).

5.2. Transparency Versus Accountability

One of the key findings of this review is that transparency and accountability, while closely related, are distinct concepts within AI governance. Transparency alone does not guarantee accountability, particularly in the context of complex or "black box" AI systems where decision-making processes remain difficult to interpret or scrutinize. In policing, institutional legitimacy depends not only on the existence of publicly disclosed systems but also on the extent to which their outputs can be understood, questioned, challenged, and, where necessary, corrected (Ananny & Crawford, 2018; Pasquale, 2015).

The findings further suggest that meaningful accountability requires more than the disclosure of information. It depends on the presence of explainability, auditability, and enforceable review mechanisms that enable oversight and redress. These safeguards are especially important in policing contexts, where AI-assisted decisions may affect individual rights, freedoms, and interactions with state authority (Kroll et al., 2017). Consistent with this perspective, the OECD AI Principles distinguish between transparency, explainability, and accountability, emphasizing the need for traceability across datasets, processes, and decisions throughout the AI system lifecycle. Together, these mechanisms provide the foundation for responsible and trustworthy AI governance in law enforcement.

5.3. Impacts on Expansion of Surveillance and Rights

The surveillance results would lead to larger rule-of-law issues. AI increases the police's ability to collect, integrate, and act on data at scale, but these efficiencies may be purchased at the cost of privacy, proportionality, and civil liberties (Brayne, 2017; Lyon, 2018; Zuboff, 2019). Zuboff notes, "If left unchecked, these capabilities threaten to erode the very democratic legitimacy of policing itself. As such, the expansion of surveillance cannot only be an operational matter — it is also a question of constitutional and human rights. The same is true for UNESCO, which emphasizes transparency, explainability, and human oversight as essential to AI governance and to upholding human rights and fundamental freedoms.

5.4. Gaps in Governance & Institutional Readiness

A comparison with global frameworks exposes a continuing implementation gap. With the OECD AI Principles reaffirmed as an international standard for trustworthy AI based on three principles: transparency, robustness and accountability, the UNESCO Recommendation positions human rights, fairness, and human oversight at the heart of its principles around AI governance. However, the literature review reveals that in practice, policing deployments generally fail to meet these substantive benchmarks of documentation and auditability (and redress) (Raji et al., 2020; Kroll et al., 2017). Hence, not a lack of principled values, but weak operationalization. This reading is also in line with the accountability literature, which argues that governance failures tend to occur when institutions deploy AI tools before establishing lifecycle controls, review mechanisms, and clear lines of responsibility (Raji et al., 2020; Leslie, 2019).

5.5. Implications for Setting Contexts in Analytical Perspectives, Including Guyana

The digital transformation has accelerated faster than the evolution of associated institutional safeguards, resulting in governance challenges that may be particularly acute within developing contexts. UNDP's Digital Strategy argues for inclusive, ethical, and sustainable digital societies. However, recent World Bank digital governance materials emphasize that rapid digitalization also poses significant risks to data protection, cybersecurity, and institutional capacity. This implies that countries such as Guyana should approach AI in policing with caution, ensuring it is based on legal certainty, procurement standards, data governance and use protocols, training and oversight capacity, and public accountability. In the absence of these protections, AI may merely exacerbate pre-existing institutional fragility rather than improve police efficacy. So to that extent, institutional readiness is not tangential to AI governance; it is a precondition for legitimate and rights-respecting deployment.

6. Conclusion

This study is an original systematic review focused on the application of artificial intelligence (AI) in policing. Taking widely adopted AI technologies in contemporary policing as its subject, this paper centers on three core challenges arising from these technologies: ethics, accountability, and governance. It first outlines their dual impacts: on one hand, these technologies can improve the operational efficiency of policing, support intelligence-led policing, and optimize resource allocation; on the other hand, they pose tangible risks including algorithmic bias, insufficient transparency, privacy infringements, and gaps in accountability and oversight.

When it comes to policing AI, people think that how well it works and whether it is fair depends on two main things: technical abilities and the rules that govern its use. However, having a strong set of rules that covers everything from creating to using AI is even more important. This study brings together three key areas - ethics, accountability, and governance - into one framework to analyze policing AI. By doing this, it aims to address problems in current research, which often examines these areas separately.

This new approach aims to provide a more complete understanding of what is needed to make policing AI effective and fair. It is proposed that the responsible implementation of policing AI requires five binding governance mechanisms: a formal AI governance framework, regular algorithmic audits, independent oversight, explainability requirements, and substantive human oversight throughout the entire decision-making process. These mechanisms are particularly critical in developing contexts where regulatory and institutional capacities are still evolving.

Future research should shift its focus to evaluating the real-world implementation outcomes of governance frameworks, while conducting cross-jurisdictional comparative studies to examine different regions' institutional readiness, regulatory approaches, and governance achievements, to advance the development of responsible, transparent policing AI systems.

Acknowledgement

The author acknowledges the teachers who participated in this study.

AI Statement

Tools such as ChatGPT were employed to assist with aspects of this manuscript, including language editing and structuring ideas. Article content has been sourced, reviewed, and verified by the author.

Conflict of Interest Statement

The author declares no conflicts of interest.

References

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

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review, 104*(3), 671–732. https://doi.org/10.15779/Z38BG31

Benjamin, R. (2019). Assessing risk, automating racism. *Science, 366*(6464), 421–422. https://doi.org/10.1126/science.aaz3873

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). SAGE. https://doi.org/10.4135/9781473920466

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review, 82*(5), 977–1008. https://doi.org/10.1177/0003122417725865

Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1), Article 2053951715622512. https://doi.org/10.1177/2053951715622512

Diakopoulos, N. (2016). Accountability in algorithmic decision-making. *Communications of the ACM, 59*(2), 56–62. https://doi.org/10.1145/2844110

Doshi-Velez, F., & Kim, B. (2017). *Towards a rigorous science of interpretable machine learning* (arXiv:1702.08608). arXiv. https://arxiv.org/abs/1702.08608

Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C., & Venkatasubramanian, S. (2018). Runaway feedback loops in predictive policing. *Proceedings of Machine Learning Research, 81*, 1–12. http://proceedings.mlr.press/v81/ensign18a.html

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

European Commission. (2020). *White paper on artificial intelligence: A European approach to excellence and trust*. European Commission. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

Kitchenham, B., Brereton, P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering: A systematic literature review. *Information and Software Technology, 51*(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review, 165*(3), 633–705. https://doi.org/10.2139/ssrn.2903468

Leslie, D. (2019). *Understanding artificial intelligence ethics and safety*. The Alan Turing Institute. https://doi.org/10.5281/zenodo.3240529

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance, 13*(5), 14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Lyon, D. (2018). *The culture of surveillance: Watching as a way of life*. Polity Press.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), Article e1000097. https://doi.org/10.1371/journal.pmed.1000097

Organization for Economic Co-operation and Development. (2019). *OECD principles on artificial intelligence*. https://oecd.ai/en/ai-principles

Okoli, C. (2015). A guide to conducting a systematic literature review of information systems research. *Journal of the Association for Information Systems, 16*(9), 879–910. https://doi.org/10.17705/1CAIS.03743

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Perry, W. L., McInnis, B., Price, C. C., Smith, S. C., & Hollywood, J. S. (2013). *Predictive policing: The role of crime forecasting in law enforcement operations* (RR-233). RAND Corporation. https://www.rand.org/pubs/research_reports/RR233.htm

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 33–44). ACM. https://doi.org/10.1145/3351095.3372873

Resnik, D. B. (2018). *The ethics of research with human subjects: Protecting people, advancing science, promoting trust*. Springer. https://doi.org/10.1007/978-3-319-68756-8

Richardson, R., Schultz, J., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *NYU Law Review Online, 94*, 15–55. https://www.nyulawreview.org

Rudin, C. (2019). Stop explaining black-box machine learning models for high-stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

United Nations Development Program. (2021). *Digital strategy 2022–2025*. https://www.undp.org/digital/Digital-Strategy

United Nations Educational, Scientific and Cultural Organization. (2021). *Recommendation on the ethics of artificial intelligence*. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Paper 440, pp. 1–14). ACM. https://doi.org/10.1145/3173574.3174014

Wiles, R. (2013). *What are qualitative research ethics?* Bloomsbury Academic. https://doi.org/10.5040/9781849666558

World Bank. (2024). *Digital progress and trends report 2023*. https://www.worldbank.org/en/publication/digital-progress-and-trends-report

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.