

Ethical and Legal Challenges of AI-Based Surveillance in National Security Operations

CHIRANJEEVI KUNAPARAJU

Principal Site Reliability Engineer at Palo Alto Networks, Santa Clara
California, United States

Email: chiranjeevirajukr@gmail.com

ORCID NO: <https://orcid.org/0009-0004-0528-6973>

Abstract

The rapid integration of artificial intelligence into national security surveillance systems has transformed intelligence gathering, threat detection, and risk assessment practices. AI-based surveillance technologies, including biometric identification, predictive analytics, and large-scale data fusion, offer enhanced operational efficiency but simultaneously intensify concerns related to privacy, discrimination, transparency, and democratic accountability. These developments challenge existing legal safeguards and ethical norms governing state surveillance activities.

This article examines the ethical and legal challenges posed by the deployment of AI-based surveillance in national security operations. The study aims to assess whether current international human rights law, data protection regimes, and emerging AI governance frameworks provide adequate safeguards against rights violations while permitting legitimate national security objectives.

The study adopts a qualitative doctrinal and normative analysis. It synthesizes international human rights instruments, European data protection law, landmark judicial decisions, and established AI governance frameworks. These legal sources are examined alongside interdisciplinary literature on algorithmic accountability, bias, and surveillance ethics to develop an integrated analytical framework for evaluating AI-driven surveillance practices.

The analysis reveals that AI-based surveillance amplifies existing surveillance risks through scale, opacity, and automation, thereby increasing the likelihood of disproportionate interference with privacy and due process rights. While current legal frameworks articulate principles such as legality, necessity, and proportionality, they often lack operational clarity when applied to complex AI systems. Ethical challenges, particularly algorithmic bias, limited explainability, and accountability gaps, further undermine effective oversight in national security contexts.

The findings underscore the need for strengthened governance mechanisms that integrate legal compliance with ethical design and operational accountability. Policymakers should mandate risk-based deployment, robust human oversight, independent auditing, and meaningful oversight structures tailored to AI-enabled surveillance. Aligning national security practices with human rights standards is essential not only for legal legitimacy but also for sustaining public trust and long-term institutional effectiveness.

Keywords: *Artificial intelligence, AI-based surveillance, national security, privacy and data protection, algorithmic accountability, human rights, biometric identification, surveillance*

governance

1. Introduction

Expansion of AI-based surveillance in national security contexts

National security agencies increasingly rely on artificial intelligence–based surveillance systems to detect, prevent, and respond to security threats. Advances in machine learning, biometric identification, and largescale data analytics have enabled governments to process vast volumes of communications data, imagery, biometric records, and behavioral signals at unprecedented speed and scale (Brayne, 2017; Ferguson, 2017). AI-driven tools such as facial recognition, predictive analytics, and automated risk scoring are now embedded in intelligence gathering, border control, counterterrorism operations, and law enforcement–security coordination frameworks. These systems promise enhanced efficiency and accuracy compared to traditional surveillance methods, particularly in environments characterized by complex, fast-evolving threats (National Institute of Standards and Technology [NIST], 2023).

At the same time, the deployment of AI surveillance has expanded beyond narrowly targeted investigations toward more systematic and continuous monitoring practices. The capacity of AI systems to integrate data from multiple sources and to infer patterns not readily observable to human analysts raises profound questions about the scope, intensity, and permanence of state surveillance in democratic societies (Brayne, 2017).

Core ethical and legal tensions between security imperatives and fundamental rights

The expansion of AI-based surveillance intensifies longstanding tensions between national security objectives and the protection of fundamental rights. International human rights law recognizes the legitimacy of national security as a ground for limiting certain rights, yet it requires that any interference with privacy, expression, or related freedoms be lawful, necessary, and proportionate (United Nations General Assembly, 1966; UN Human Rights Committee, 1988, 2011). AI systems challenge these requirements by introducing opacity, automation bias, and probabilistic decision-making into surveillance practices.

From an ethical perspective, concerns arise regarding algorithmic bias, discrimination, lack of transparency, and diminished accountability. Empirical studies demonstrate that some AI surveillance tools, particularly facial recognition systems, exhibit differential accuracy across demographic groups, thereby increasing the risk of unjustified targeting and exclusion (Buolamwini & Gebru, 2018; Grother et al., 2019). Moreover, the complexity of AI models complicates individuals' ability to understand, challenge, or seek redress against surveillance-driven decisions, undermining procedural fairness and due process (Citron & Pasquale, 2014; Wachter et al., 2017).

Legally, courts and regulators have struggled to adapt existing surveillance and data protection frameworks to AI-enabled practices. Case law in Europe illustrates growing judicial concern about bulk and automated surveillance regimes that lack adequate safeguards, independent oversight, and effective remedies (Digital Rights Ireland Ltd v Minister for Communications, 2014; Roman Zakharov v Russia, 2015; Big

Brother Watch and Others v the United Kingdom, 2021). These developments underscore the need to reassess how ethical principles and legal standards apply in AI-driven security environments.

Research problem and objectives

The central problem addressed in this study is the absence of a coherent, integrated framework for evaluating the ethical and legal implications of AI-based surveillance in national security operations. While existing scholarship examines either legal constraints or ethical risks in isolation, fewer studies systematically analyze how international human rights law, data protection regimes, and AI governance frameworks jointly shape the permissible use of AI surveillance.

Accordingly, the objectives of this study are to:

- ✦ Examine the legal foundations governing AI-based surveillance under international human rights law and regional data protection frameworks.
- ✦ Identify the principal ethical challenges posed by AI surveillance in national

10.48047/jocaaa.2024.33.08.480

security contexts, including bias, transparency, and accountability.

- ✦ Analyze how courts, regulators, and governance frameworks have responded to AI-enabled surveillance practices.
- ✦ Propose policy-relevant insights for designing rights-respecting AI surveillance systems that remain operationally effective.

Contribution of the study

This study contributes to the literature in three key ways. First, it provides an integrated ethical–legal analysis that connects normative principles with concrete surveillance practices. Second, it synthesizes insights from human rights law, data protection law, and AI governance to clarify how existing standards apply to AI-based national security surveillance. Third, it offers a structured framework to support policymakers, regulators, and security institutions in balancing security imperatives with fundamental rights in the deployment of AI technologies.



Figure 1. Conceptual Overview of AI-Based Surveillance in National Security Operations

This figure illustrates the end-to-end process of AI-based surveillance in national security contexts, beginning with the threat environment and progressing through AI-enabled surveillance tools, decision outputs, and operational actions. It also highlights the potential impacts on fundamental rights at each stage and identifies key legal and ethical safeguards, including oversight, proportionality, and accountability mechanisms, that are necessary to ensure rights-respecting deployment of AI surveillance systems.

2. Conceptual Background:
AIBased Surveillance Systems

2.1 Definition and Scope of AIBased Surveillance

AI-based surveillance refers to the use of artificial intelligence techniques to collect, process, and analyze large volumes of data for the purpose of monitoring, identifying, or predicting human behavior within security and intelligence contexts. Unlike traditional

surveillance, which relies heavily on manual review and predefined rules, AI-based surveillance systems operate through automated pattern recognition, probabilistic inference, and machinelearning models that continuously adapt based on new data inputs (Brayne, 2017; Ferguson, 2017).

In national security operations, the scope of AI-based surveillance typically includes facial recognition technologies, biometric identification systems, predictive analytics, and data fusion platforms. Facial recognition

10.48047/jocaaa.2024.33.08.480

systems apply computer vision and machine-learning algorithms to identify or verify individuals by analyzing facial features captured from images or video streams, often drawn from public spaces, border controls, or intelligence databases (Buolamwini & Gebru, 2018; Grother et al., 2019). Biometric systems extend beyond facial recognition to include fingerprints, iris scans, voice recognition, and gait analysis, enabling persistent identification across different environments and datasets.

Predictive analytics involves the use of statistical and machine-learning models to forecast potential security threats based on historical data, behavioral indicators, and network relationships. In intelligence and policing contexts, these tools are used to assess risk, prioritize targets, and allocate resources, often influencing decisions before any concrete offense has occurred (Citron & Pasquale, 2014; Brayne, 2017). Data fusion systems integrate heterogeneous data sources such as communications metadata, biometric databases, financial records, travel logs, and social media activity into unified analytical platforms, allowing security agencies to construct detailed profiles and infer associations at scale.

Together, these technologies significantly expand the reach, depth, and persistence of surveillance, raising complex ethical and legal questions regarding privacy, fairness, accountability, and democratic oversight.

2.2 Technical Characteristics Relevant to Ethics and Law

Several technical characteristics of AI-based surveillance systems are directly relevant to ethical evaluation and legal compliance. First, automation plays a central role. Decisions such as identity matching, risk scoring, or

anomaly detection are increasingly generated by automated systems with minimal human intervention. This raises concerns about automation bias, where human operators defer to algorithmic outputs even when those outputs may be uncertain or erroneous (Raji et al., 2020; Kroll et al., 2017).

Second, opacity is a defining feature of many AI surveillance systems. Machine learning models, particularly deep neural networks, often function as black boxes, making it difficult for affected individuals or oversight bodies to understand how specific conclusions were reached (Pasquale, 2015; Mittelstadt et al., 2016).

This lack of explainability complicates legal requirements related to transparency, due process, and the ability to challenge adverse decisions, especially in security contexts where secrecy is already prevalent (Wachter et al., 2017).

Third, scale distinguishes AI-based surveillance from traditional methods. AI systems enable continuous monitoring of vast populations rather than targeted surveillance of specific individuals. This scalability increases the risk of indiscriminate or mass surveillance, which may conflict with international human rights standards requiring targeted and proportionate interference with privacy (UN Human Rights Committee, 1988).

Finally, error propagation presents a significant risk. Errors or biases embedded in training data can be amplified as AI systems are deployed across interconnected security infrastructures. Empirical studies on facial recognition have demonstrated significant demographic disparities in error rates, particularly affecting racial and gender minorities, which may lead to discriminatory outcomes when such systems are used in national security or law enforcement decisionmaking (Buolamwini & Gebru, 2018; Grother, 2022).

Table 1: Types of AI Surveillance Technologies and Their Security Applications

Technology type	Function	Security use case	Key risk
Facial recognition	Automated identification or verification of individuals from images or video	Border control, watchlist matching, public space monitoring	Misidentification, demographic bias
Biometric systems	Persistent identification using biological traits	Identity management, access control, intelligence databases	Irreversible privacy ha
Predictive analytics	Forecasting potential threats based on patterns	Risk scoring, threat prioritization	Preemptive profiling, f positives
Data fusion platforms	Integration of multiple datasets into unified profiles	Intelligence analysis, network mapping	Mass surveillance, fu creep

3. International Legal Framework Governing AI Surveillance

3.1 Right to Privacy and Freedom of Expression Under International Law

The international legal framework governing surveillance is grounded primarily in the International Covenant on Civil and Political Rights (ICCPR). Article 17 of the ICCPR provides that no one shall be subjected to arbitrary or unlawful interference with their privacy, family, home, or correspondence, and that everyone has the right to the protection of the law against such interference (United Nations General Assembly, 1966). This provision establishes privacy as a fundamental human right applicable to all forms of state surveillance, including those justified on national security grounds.

The UN Human Rights Committee has clarified the scope of Article 17 through General Comment No. 16, emphasizing that surveillance measures must not be arbitrary and must be authorized by law that is precise, accessible, and foreseeable in its application (UN Human Rights Committee, 1988). The Committee further stressed that even lawful surveillance must be reasonable in the particular circumstances, thereby introducing an implicit proportionality requirement.

Article 19 of the ICCPR protects the right to freedom of expression, including the freedom to seek, receive, and impart information. In General Comment No. 34, the Human Rights Committee recognized that surveillance practices can have a chilling effect on expression, particularly when individuals fear that their communications are being

10.48047/jocaaa.2024.33.08.480 monitored or analyzed by state authorities (UN Human Rights Committee, 2011). AI-based surveillance systems that monitor online activity, communications metadata, or social networks may therefore interfere simultaneously with privacy and expressive freedoms.

3.2 Lawful Limitations for National Security

International human rights law permits certain limitations on privacy and expression for legitimate national security objectives, but such limitations are subject to strict conditions. Surveillance measures must satisfy the principles of legality, necessity, and proportionality, and must be accompanied by effective oversight and remedies.

Legality requires that surveillance activities be grounded in clear and publicly accessible laws that define the scope, purpose, and limits of state powers. Necessity demands that surveillance measures address a pressing security need and that less intrusive alternatives are insufficient. Proportionality requires that the degree of interference with individual rights be balanced against the concrete security benefit pursued (UN Human Rights Committee, 1988).

Oversight is an essential component of lawful surveillance. Independent judicial or quasi-judicial mechanisms must supervise the authorization, implementation, and review of surveillance activities to prevent abuse and ensure accountability. In the context of AI-based surveillance, these principles require particular attention due to the scale, opacity, and automated nature of the technologies involved, which can

10.48047/jocaaa.2024.33.08.480
obscure individual responsibility and weaken
traditional safeguards (Council of Europe, 1981;
De Terwangne, 2021).

Table 2: International Legal Principles Applicable to AI Surveillance

Legal principle	Source	Legal requirement	Implication for AI systems
Legality	ICCPR Article 17	Surveillance must be clearly authorized by law	Transparent system design and documented use
Necessity	General Comment No. 16	Interference only when strictly required	Limits on mass or continuous monitoring
Proportionality	ICCPR Articles 17 and 19	Balance between rights and security aims	Targeted deployment and minimization
Oversight	ICCPR, Council of Europe Convention 108	Independent supervision and remedies	Auditability and human review mechanisms

10.48047/jocaaa.2024.33.08.480

4. European Data Protection and Surveillance Law

European data protection law provides one of the most developed legal frameworks governing surveillance activities, including those involving artificial intelligence. While national security remains primarily within the competence of Member States, European legal instruments impose binding constraints on how surveillance technologies may be designed, deployed, and overseen. These constraints are particularly relevant where AI systems enable large-scale, automated, or biometric surveillance.

4.1 Convention 108 and Convention 108+

The **Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108)** represents the first binding international treaty dedicated to data protection. Convention 108 establishes core principles that directly apply to AI-based surveillance systems, including lawfulness, fairness, purpose limitation, data minimization, and security safeguards.

The modernised **Convention 108+** extends these protections by explicitly addressing contemporary data-processing risks associated with algorithmic systems. It strengthens requirements for transparency, accountability, and independent supervision, and introduces enhanced protections against automated decisionmaking that significantly affects individuals. Importantly, Convention 108+ clarifies that even in national security contexts, personal data processing must remain subject to legal safeguards and independent oversight.

For AI-based surveillance, these instruments are particularly relevant because they restrict indiscriminate or bulk data processing and require that automated systems operate within clearly defined legal mandates. They also emphasize the need for effective remedies, even where secrecy considerations limit full disclosure.

4.2 GDPR and Law Enforcement Directive (Directive 2016/680)

The General Data Protection Regulation (GDPR) establishes comprehensive data protection obligations across the European Union, including rules on automated decision-making, profiling, and biometric data processing. Although the GDPR contains exemptions for national security, its principles remain influential in shaping acceptable standards for AI system design, especially where surveillance systems rely on biometric identification or predictive analytics.

Complementing the GDPR, the **Law Enforcement Directive (Directive (EU) 2016/680)** applies specifically to data processing by competent authorities for law enforcement purposes, including crime prevention, investigation, and prosecution.

The Directive introduces safeguards particularly relevant to AI-based surveillance, such as requirements for data accuracy, strict purpose limitation, logging of processing activities, and special protection for sensitive data categories.

The Directive also places limits on solely automated decision-making that produces adverse legal effects, reinforcing the principle that meaningful human involvement must be retained in security decision processes. These provisions directly challenge the deployment

10.48047/jocaaa.2024.33.08.480

of fully automated AI surveillance systems without effective human review mechanisms.

4.3 Judicial Oversight through EU and ECHR Case Law

European courts have played a central role in defining the legal boundaries of surveillance, particularly where technological capabilities

effective oversight mechanisms. The Court underscored that the mere existence of unchecked surveillance powers, even without proven abuse, could constitute a violation of the right to privacy. This judgment is highly relevant to AI-based surveillance systems that operate with limited transparency.

Similarly, in **Big Brother Watch v United**

Table 3: Key European Legal Instruments and Case Law

Instrument or case	Scope	Surveillance constraint	Relevance to AI system
Convention 108	Council of Europe member states	Lawfulness, purpose limitation, data minimization	Limits large-scale automated data processing
Convention 108+	Modernised data protection treaty	Enhanced accountability and safeguards against automated decision-making	Addresses algorithmic and AI-driven processing
GDPR	EU general data protection	Restrictions on profiling and biometric data	Influences design of AI surveillance systems
Directive 2016/680	Law enforcement data processing	Accuracy, logging, and human oversight	Limits fully automated security decisions
Digital Rights Ireland	EU data retention law	Prohibits indiscriminate mass data retention	Challenges bulk AI surveillance measures
Roman Zakharov v Russia	Secret surveillance regimes	Requires legal clarity and oversight	Applies to opaque AI surveillance practices

risk undermining fundamental rights.

In **Digital Rights Ireland**, the Court of Justice of the European Union invalidated the Data Retention Directive on the grounds that indiscriminate data retention violated the principles of necessity and proportionality. The Court emphasized that large-scale data collection without individualized suspicion or effective safeguards was incompatible with fundamental rights protections. This reasoning is directly applicable to AI-driven mass surveillance systems that rely on generalized data harvesting.

In **Roman Zakharov v Russia**, the European Court of Human Rights held that secret surveillance regimes must be accompanied by robust legal safeguards, including clear legal bases, independent authorization, and

Kingdom, the Grand Chamber found that bulk interception regimes require stringent safeguards to prevent abuse, including oversight of selection criteria, storage, access, and sharing of data. The Court highlighted the heightened risks posed by automated and large-scale surveillance, reinforcing the need for legal controls that keep pace with technological advancement.

Together, these cases establish that AI-enabled surveillance must be subject to strict legal constraints, independent oversight, and demonstrable necessity, regardless of national security justifications.

10.48047/jocaaa.2024.33.08.480

5. Ethical Challenges of AI-Based Surveillance

Beyond legal compliance, AI-based surveillance raises profound ethical concerns that challenge the legitimacy and social acceptability of national security practices. These challenges stem from the technical characteristics of AI systems and their interaction with power, secrecy, and scale.

5.1 Algorithmic Bias and Discrimination

One of the most documented ethical risks of AI surveillance is algorithmic bias. Empirical studies of facial recognition systems demonstrate significantly higher error rates for women and racial minorities compared to white male subjects. These disparities raise serious concerns when such systems are used in security contexts, where false positives can result in wrongful suspicion, detention, or surveillance escalation.

In national security operations, biased AI systems risk reinforcing existing structural inequalities and disproportionately targeting marginalized communities. The ethical problem is amplified by the difficulty individuals face in detecting, challenging, or correcting algorithmic errors within secretive security frameworks.

5.2 Transparency, Explainability, and Contestability

AI-based surveillance systems often rely on complex models that are difficult to interpret, even by their developers. This opacity undermines transparency and makes it challenging for affected individuals to understand why they were flagged, monitored, or subjected to further investigation.

From an ethical standpoint, the lack of explainability weakens due process and erodes trust in public institutions. Where individuals cannot contest AI-driven decisions or access meaningful explanations, the right to an effective remedy becomes largely illusory. This tension is particularly acute in national security settings, where secrecy further restricts access to information.

5.3 Automation Bias and Accountability Gaps

Automation bias occurs when human operators place undue trust in algorithmic outputs, leading them to defer to AI systems even when errors or inconsistencies are present. In security environments, this bias can result in reduced critical scrutiny and the normalization of flawed or unjustified surveillance decisions.

Accountability gaps emerge when responsibility for harm becomes diffused across technical developers, security agencies, and oversight bodies. Without clear accountability structures, ethical responsibility is displaced, and meaningful redress becomes difficult to achieve. These issues highlight the necessity of maintaining robust human oversight and clearly defined responsibility chains in AI-enabled security operations.

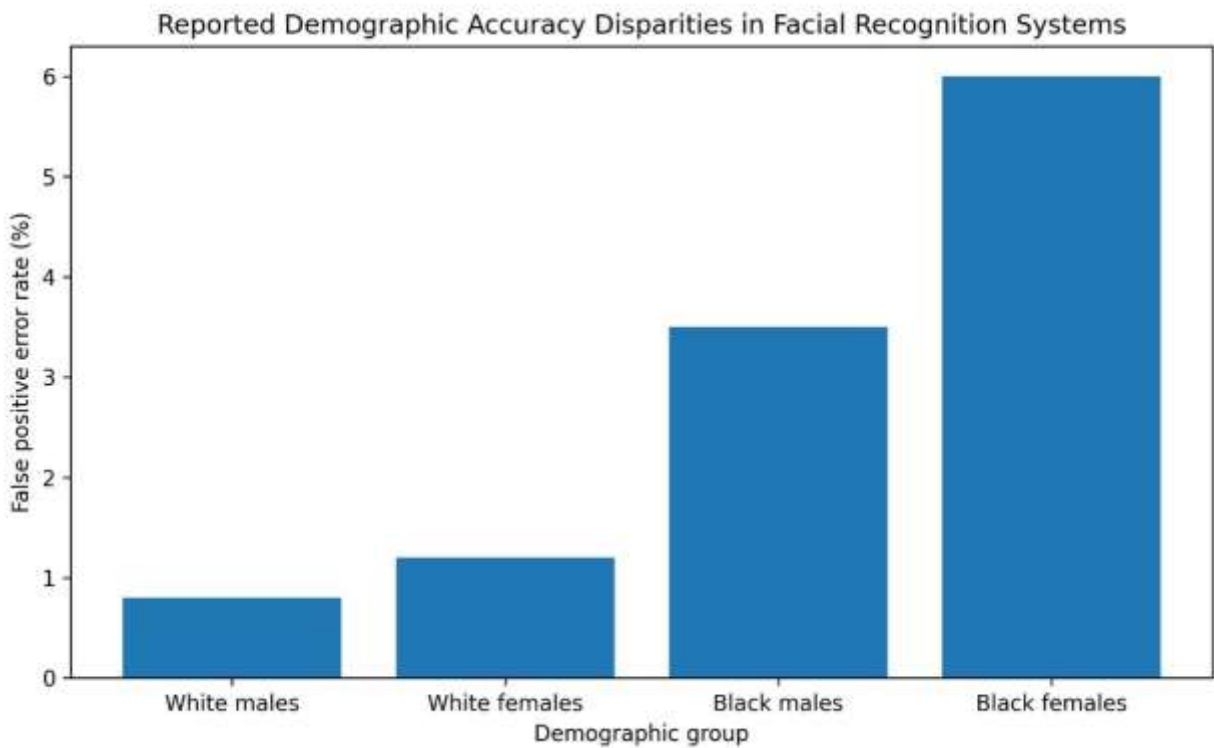


Figure 2. Bar Chart Reported Demographic Accuracy Disparities in Facial Recognition Systems

Figure 2 illustrates reported false positive error rates across demographic groups in facial recognition systems, showing higher error rates for women and racial minorities. These disparities highlight ethical risks associated with deploying biometric AI technologies in national security surveillance contexts.

10.48047/jocaaa.2024.33.08.480

Zakharov v. Russia, 2015; Big Brother Watch and Others v. United Kingdom, 2021).

6. Governance, Accountability, and Oversight Mechanisms

The deployment of AI-based surveillance in national security operations necessitates robust governance structures to mitigate risks of rights violations, abuse of power, and unaccountable decision-making. Given the opacity, scale, and automation inherent in AI systems, traditional oversight mechanisms designed for human-led surveillance are often insufficient. Effective governance therefore requires a combination of institutional oversight, technical accountability tools, and internationally recognized AI governance frameworks that collectively ensure legality, proportionality, and democratic accountability (Brayne, 2017; Kroll et al., 2017).

6.1 Institutional oversight models

Institutional oversight remains the cornerstone of accountability for national security surveillance. Courts, parliamentary bodies, and independent regulators play distinct but complementary roles in constraining executive power and safeguarding fundamental rights.

Judicial oversight is primarily responsible for authorizing surveillance activities and reviewing their legality. Courts assess whether AI-enabled surveillance measures meet established legal standards of necessity, proportionality, and foreseeability, particularly under human rights law (United Nations General Assembly, 1966). Case law from the European Court of Human Rights demonstrates that surveillance regimes lacking effective judicial control are likely to violate the right to privacy, even when justified on national security grounds (Roman

Parliamentary oversight bodies provide democratic scrutiny over intelligence agencies by reviewing surveillance legislation, budgets, and operational compliance. Parliamentary committees may conduct inquiries, demand classified briefings, and issue public reports, thereby enhancing transparency without compromising operational secrecy (Zalnieriute, 2022). However, the effectiveness of parliamentary oversight is often constrained by information asymmetries and limited technical expertise in AI systems.

Independent regulators and supervisory authorities serve as specialized oversight institutions with technical and legal expertise. Data protection authorities, for example, are empowered under European law to supervise the processing of personal data by law enforcement and intelligence bodies, including algorithmic systems (European Union, 2016a; European Union Agency for Fundamental Rights, 2018). Their independence from political branches is critical for ensuring impartial enforcement, although their access to classified systems may be restricted in practice.

6.2 Technical accountability tools

In addition to institutional oversight, technical accountability tools are essential for governing AI-based surveillance systems at the operational level. These tools embed accountability directly into system design and deployment.

Algorithmic audits are systematic evaluations of AI systems to assess compliance with legal, ethical, and technical standards. Audits may

10.48047/jocaaa.2024.33.08.480

examine training data, model performance, bias, and decision logic, and they are increasingly recognized as a mechanism for detecting discriminatory outcomes and systemic errors (Raji et al., 2020). In national security contexts, audits are often conducted internally or by trusted third parties due to classification constraints.

Logging and traceability mechanisms record system inputs, outputs, and decision pathways, enabling post hoc review and investigation. Comprehensive logging is critical for reconstructing how an AI system contributed to a surveillance decision, particularly where adverse outcomes occur. Without such records, accountability is undermined and meaningful redress becomes impossible (Kroll et al., 2017).

Impact assessments, including data protection impact assessments and algorithmic impact assessments, provide ex ante evaluation of risks before deployment. These assessments identify potential harms to privacy, non-discrimination, and due process, and they require authorities to justify system necessity and adopt mitigation measures (Leiser & Custers, 2019). While impact assessments are widely mandated in civilian data protection law, their application in national security remains uneven and often voluntary.

6.3 AI governance frameworks International AI governance frameworks provide normative guidance for responsible AI deployment, including in securitysensitive environments. Although not legally binding in all cases, these frameworks increasingly shape regulatory expectations and institutional practices.

The OECD AI principles emphasize humancentered values, transparency, robustness, and accountability across the AI

lifecycle (Yeung, 2020). These principles stress that public-sector AI systems should be subject to oversight mechanisms appropriate to their risk level, making them particularly relevant for surveillance applications.

The NIST Artificial Intelligence Risk Management Framework provides a structured approach to identifying, assessing, and mitigating AI-related risks (National Institute of Standards and Technology, 2023). The framework encourages continuous monitoring, documentation, and governance processes, offering practical tools that can be adapted to classified environments while maintaining accountability.

The UNESCO Recommendation on the Ethics of Artificial Intelligence establishes a global ethical baseline grounded in human rights, rule of law, and democratic values (UNESCO, 2021). It explicitly warns against mass surveillance and disproportionate data collection, calling for transparency, oversight, and remedies even in public security contexts.

Table 4: Governance mechanisms for AI surveillance accountability

Mechanism	Responsible body	Purpose	Limitations
Judicial authorization	Courts	Ensure legality and proportionality of surveillance	Limited technical insight in models
Parliamentary oversight	Legislative committees	Democratic scrutiny and policy review	Information asymmetry and secrecy constraints
Regulatory supervision	Independent regulators	Enforce data protection and compliance	Restricted access to classified systems
Algorithmic audits	Internal or third-party auditors	Detect bias, errors, and non-compliance	Often non-public and discretionary
Impact assessments	Deploying agencies	Identify risks prior to deployment	Variable depth and enforcement

7. Comparative Regulatory Approaches

Regulatory responses to AI-based surveillance vary significantly across jurisdictions, reflecting different legal traditions, threat perceptions, and human rights commitments. A comparative perspective highlights both convergence around core principles and divergence in implementation.

In the United Kingdom, the Investigatory Powers Act 2016 establishes a comprehensive statutory framework governing surveillance activities, including bulk data collection and interception. The Act combines executive authorization with judicial approval through a “double-lock” mechanism, aiming to balance operational flexibility with legal oversight (Investigatory Powers Act 2016). While the Act predates widespread AI deployment, its oversight architecture has been interpreted by courts in light of emerging technologies (Glover, 2023).

The European Union adopts a rights-centric regulatory approach rooted in data protection and human rights law. Instruments such as the General Data Protection Regulation and the Law Enforcement Directive impose strict requirements on data processing, transparency, and accountability, even for public authorities (European Union, 2016a; European Union, 2016b). Judicial decisions, including Digital Rights Ireland and Big Brother Watch, reinforce the principle that generalized or indiscriminate surveillance is incompatible with fundamental rights.

Emerging global standards increasingly draw on soft-law instruments and ethical frameworks. OECD, NIST, and UNESCO guidelines promote convergence around risk-based governance, human oversight, and accountability mechanisms, although their implementation in national security remains largely voluntary. These standards nonetheless influence legislative reforms, procurement policies, and international cooperation.

Comparative Regulatory Model for AI Surveillance Governance			
	Legal Safeguards <i>(Legality, Necessity, Proportionality)</i>	Oversight Mechanisms <i>(Courts, Regulators, Parliamentary Review)</i>	Accountability Tools <i>(Audits, Logging, Impact Assessments)</i>
United Kingdom	 Robust Legal Tests	 Judicial & IPCO Oversight	 Limited Auditing & Review
European Union	 Strict Data Protection	 Data Regulators & Courts	 Comprehensive Audits
Global Standards	 Variable Protections	 Inconsistent Oversight	 Voluntary & Emerging

Figure 3. Comparative regulatory model for AI surveillance governance

This matrix diagram compares regulatory approaches to AI-based surveillance across three governance dimensions: legal safeguards (legality, necessity, and proportionality), oversight mechanisms (courts, regulators, and parliamentary review), and accountability tools (audits, logging, and impact assessments). The rows represent selected jurisdictions and frameworks, namely the United Kingdom, the European Union, and global soft-law standards. The columns reflect core governance dimensions, allowing visual comparison of relative strengths and gaps in how AI surveillance systems are regulated and overseen across different legal and institutional contexts.

8. Policy and Operational Implications

The increasing deployment of AI-based surveillance tools in national security operations requires carefully designed policy and operational safeguards to ensure alignment with legal obligations and ethical principles. While AI technologies may enhance intelligence collection and threat detection capabilities, their deployment without adequate governance mechanisms risks undermining fundamental rights, democratic legitimacy, and public trust. This section outlines key policy and operational implications for designing, governing, and deploying AI-driven surveillance systems within national security frameworks.

8.1 Designing rights-respecting AI surveillance systems

Rights-respecting AI surveillance systems must be grounded in the principles of legality, necessity, proportionality, and accountability as articulated under international human rights law. Article 17 of the International Covenant on Civil and Political Rights establishes a clear prohibition against arbitrary or unlawful interference with privacy, even in contexts involving national security (United Nations General Assembly, 1966; UN Human Rights Committee, 1988). AI systems that automate or scale surveillance activities therefore require explicit statutory authorization and narrowly defined operational purposes.

From an ethical perspective, system design must address well-documented risks of algorithmic bias, particularly in biometric and facial recognition technologies. Empirical evidence demonstrates that commercially

10.48047/jocaaa.2024.33.08.480

deployed facial recognition systems exhibit differential error rates across gender and racial groups, raising concerns of discriminatory impact when such tools are used in security operations (Buolamwini & Gebru, 2018; Grother et al., 2019). Rights-respecting design thus requires pre-deployment bias testing, continuous performance monitoring, and clearly defined error thresholds beyond which systems may not be operationally used.

In addition, explainability and contestability must be incorporated at the design stage. While full transparency may be constrained by classified operational requirements, decision-support systems should allow for internal explanation, auditability, and meaningful human review. The absence of such safeguards risks undermining due process and accountability, particularly when AI outputs influence surveillance targeting or enforcement decisions (Citron & Pasquale, 2014; Mittelstadt et al., 2016).

8.2 Balancing operational secrecy and democratic accountability

National security operations necessarily involve a degree of secrecy to protect sources, methods, and ongoing investigations. However, secrecy cannot operate as a blanket justification for the erosion of democratic oversight. Judicial and comparative legal analysis demonstrates that effective surveillance governance requires layered accountability mechanisms, even where public disclosure is limited (Roman Zakharov v. Russia, 2015; Big Brother Watch and Others v. the United Kingdom, 2021).

Democratic accountability can be maintained through institutional mechanisms such as independent judicial authorization, parliamentary intelligence committees, and data protection authorities with classified oversight mandates. These bodies serve as proxies for public scrutiny by ensuring that surveillance practices comply with statutory limits and human rights standards. The European Court of Human Rights has consistently emphasized that independent oversight and effective remedies are essential components of lawful surveillance regimes (Zalnieriute, 2022).

Ethically, accountability also requires resisting automation bias within security institutions. When AI outputs are treated as inherently objective or superior to human judgment, there is a heightened risk that decision-makers defer responsibility to algorithmic systems. Such deference weakens individual accountability and may normalize excessive surveillance practices over time (Raji et al., 2020; Kroll et al., 2017). Policies must therefore mandate documented human review and clear responsibility chains for AI-assisted surveillance decisions.

8.3 Implications for procurement and deployment decisions

Procurement decisions play a critical role in shaping the ethical and legal outcomes of AI-based surveillance. Security agencies increasingly rely on private vendors for AI tools, creating challenges related to transparency, vendor accountability, and system interoperability. Procurement frameworks must therefore include mandatory requirements for ethical compliance, data protection safeguards, and

10.48047/jocaaa.2024.33.08.480
audit access, rather than focusing solely on technical performance.

International and transnational governance instruments provide useful guidance for procurement standards. The OECD Recommendation on Artificial Intelligence and the NIST AI Risk Management Framework emphasize risk-based assessment, lifecycle governance, and continuous monitoring as prerequisites for trustworthy AI deployment (Yeung, 2020; National Institute of Standards and Technology, 2023). Similarly, UNESCO's Recommendation on the Ethics of Artificial Intelligence underscores the responsibility of public authorities to ensure that AI systems procured for security purposes remain aligned with human rights norms (UNESCO, 2021).

Operational deployment decisions must also consider whether AI surveillance systems are proportionate to the specific security threat being addressed. Broad or indefinite deployment risks creating mission creep, where technologies initially justified for exceptional threats become normalized tools of routine surveillance. Legal and ethical safeguards must therefore include sunset clauses, periodic review requirements, and evidence-based assessments of effectiveness relative to rights impact.

Table 5: Recommended safeguards for national security AI surveillance

Safeguard	Ethical rationale	Legal basis	Implementation requirements
Clear statutory authorization	Prevents arbitrary surveillance	ICCPR Article 17	Explicit legislation defining scope and purpose
Purpose limitation	Limits mission creep	GDPR; Directive 2016/680	Narrow operational mandates
Human-in-the-loop review	Preserves accountability	ECHR jurisprudence	Mandatory documented review and validation
Bias and accuracy testing	Prevents discrimination	Equality and nondiscrimination principles	Pre- and post-deployment testing
Independent oversight	Ensures democratic control	Big Brother Watch judgment	Judicial or parliamentary oversight bodies
Auditability and logging	Enables accountability	Convention 108+	Secure, reviewable system logs
Periodic proportionality review	Aligns surveillance with necessity	ICCPR jurisprudence	Regular reassessment of necessity and justification

9. Discussion

Synthesis of legal and ethical findings

The analysis demonstrates that AI-based surveillance in national security operations operates at the intersection of evolving technological capability and long-standing legal and ethical norms. International human rights law provides a stable normative foundation centered on legality, necessity, proportionality, and accountability, while ethical scholarship highlights risks related to bias, opacity, and institutional overreliance on automated systems. Together, these perspectives reveal that AI surveillance challenges are not solely technical but deeply governance-oriented.

Tensions between effectiveness and legitimacy

A central tension identified in this study lies between operational effectiveness and democratic legitimacy. While AI technologies may enhance surveillance efficiency, their legitimacy depends on compliance with legal safeguards and ethical standards. Surveillance practices that deliver short-term intelligence gains but undermine privacy, equality, or accountability ultimately risk eroding institutional trust and public confidence, which are themselves essential to effective national security governance.

Risks of unchecked AI deployment in security operations

Unchecked deployment of AI-based surveillance systems poses significant risks, including normalization of mass surveillance, discriminatory targeting, and dilution of human responsibility. Empirical

evidence from policing and biometric research indicates that technological accuracy alone cannot justify deployment absent robust oversight and accountability mechanisms (Brayne, 2017; Grother, 2022). Without clear limits and governance structures, AI surveillance risks shifting security policy toward technologically driven decision-making divorced from democratic and legal control.

10. Limitations and Future Research Directions

10.1 Access to classified data

A primary limitation of research on AI-based surveillance in national security operations is restricted access to classified or sensitive operational data. Intelligence and security agencies rarely disclose detailed information on system design, deployment parameters, error rates, or decision-making processes due to national security considerations. This opacity constrains the ability of scholars to empirically assess compliance with legal principles such as necessity, proportionality, and effective oversight, which are central to international human rights law (United Nations General Assembly, 1966; UN Human Rights Committee, 1988).

As highlighted in European human rights jurisprudence, the secrecy inherent in surveillance activities complicates meaningful external scrutiny and judicial review, even where formal oversight mechanisms exist (Roman Zakharov v. Russia, 2015; Zalnieriute, 2022). Consequently, much of the existing literature relies on doctrinal legal analysis, policy documents, or isolated case studies rather than systematic empirical evidence. Future research would benefit from controlled access arrangements, anonymized audit data,

10.48047/jocaaa.2024.33.08.480
or oversight body reports that enable independent evaluation without compromising legitimate security interests.

10.2 Rapid technological evolution

A second limitation arises from the rapid pace of technological change in artificial intelligence systems used for surveillance. Advances in machine learning, biometric identification, and data fusion techniques frequently outpace the development of legal frameworks and ethical guidelines designed to regulate them. Regulatory instruments such as the GDPR and the Law Enforcement Directive were drafted before the widespread deployment of contemporary AI surveillance tools, resulting in interpretative challenges when applying existing principles to novel technical contexts (Leiser & Custers, 2019; Wachter et al., 2017).

Similarly, governance frameworks such as the OECD AI principles and the NIST AI Risk Management Framework provide high-level guidance but may struggle to anticipate future capabilities or emergent risks associated with increasingly autonomous and adaptive systems (National Institute of Standards and Technology, 2023; Yeung, 2020). Future research should therefore adopt adaptive and forward-looking approaches, including scenario analysis and regulatory sandbox evaluations, to assess how ethical and legal safeguards can remain effective in the face of continual technological evolution.

10.3 Need for empirical evaluation of security outcomes

A further limitation concerns the limited availability of empirical evidence

demonstrating the actual security benefits of AI-based surveillance. While such systems are frequently justified on grounds of enhanced efficiency, threat detection, or crime prevention, there is comparatively little peer-reviewed research evaluating whether these outcomes are consistently achieved in practice (Brayne, 2017; Ferguson, 2017). Moreover, few studies systematically weigh security gains against documented risks such as false positives, discriminatory impacts, or erosion of public trust.

Existing technical evaluations, particularly in the context of facial recognition, have demonstrated significant demographic performance disparities, raising questions about reliability and fairness in high-stakes security environments (Buolamwini & Gebru, 2018; Grother et al., 2019). Future research should prioritize interdisciplinary empirical studies that combine legal analysis, technical audits, and outcomebased assessments to evaluate whether AI surveillance systems deliver measurable security improvements that justify their ethical and legal costs.

11. Conclusion

11.1 Summary of key arguments

This article has examined the ethical and legal challenges posed by the use of AI-based surveillance in national security operations. It has shown that while such technologies offer potential operational advantages, their deployment raises profound concerns relating to privacy, data protection, due process, transparency, and accountability. International human rights law, European data protection regimes, and judicial case law collectively establish that surveillance measures, including

10.48047/jocaaa.2024.33.08.480

those powered by AI, must comply with principles of legality, necessity, proportionality, and effective oversight (United Nations General Assembly, 1966; Council of Europe, 1981; Big Brother Watch v. United Kingdom, 2021).

From an ethical perspective, the analysis highlights persistent risks associated with algorithmic bias, automation bias, and accountability gaps, particularly in contexts where individuals may be subject to adverse consequences without meaningful avenues for explanation or redress (Citron & Pasquale, 2014; Mittelstadt et al., 2016; Raji et al., 2020). These challenges are amplified in national security settings, where secrecy and urgency often limit transparency and public scrutiny.

11.2 Contributions to AI governance and security law scholarship

The article contributes to AI governance and security law scholarship by integrating legal doctrine, ethical theory, and technical evidence into a unified analytical framework. It demonstrates how established legal standards governing surveillance can be meaningfully applied to contemporary AI systems, while also identifying areas where existing frameworks require reinterpretation or reinforcement. By drawing on international instruments, European regulatory models, and emerging AI governance standards, the study provides a structured basis for evaluating AI surveillance beyond purely technical or policy-oriented discussions.

Additionally, the article advances scholarly debate by emphasizing the importance of

accountability mechanisms, such as algorithmic auditing and risk management frameworks, as essential complements to traditional legal oversight in national security contexts (Kroll et al., 2017; National Institute of Standards and Technology, 2023).

11.3 Final policy message

The central policy message of this study is that the legitimacy and effectiveness of AI-based surveillance in national security operations depend on robust ethical and legal governance. AI systems should not be treated as exceptional technologies exempt from established rights protections. Instead, they must be embedded within clear legal mandates, subject to independent oversight, and supported by technical safeguards that ensure fairness, transparency, and accountability.

Without such measures, the expansion of AI surveillance risks undermining public trust, weakening democratic accountability, and eroding the very values that national security institutions are intended to protect. Conversely, rights-respecting governance frameworks offer a pathway to harnessing the potential benefits of AI while maintaining compliance with fundamental legal and ethical principles.

References

1. AI, N. (2023). Artificial intelligence risk management framework (AI RMF 1.0). URL: <https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1>.
2. Brayne, S. (2017). Big data surveillance: The case of policing. *American sociological review*, 82(5), 977-1008.
3. Buolamwini, J., & Gebru, T. (2018, January). Gender shades: 10.48047/jocaaa.2024.33.08.480 Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR.
4. Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. *Wash. L. Rev.*, 89, 1.
5. de l'Europe, C. (1981). *Convention for the protection of individuals with regard to automatic processing of personal data* (Vol. 108). Council of Europe.
6. De Terwangne, C. (2021). Council of Europe convention 108+: A modernised international treaty for the protection of personal data. *Computer Law & Security Review*, 40, 105497.
7. Directive - 2016/680 - EN - Law Enforcement Directive; LED - EUR-Lex. (2016). Europa.eu. <https://eurlex.europa.eu/eli/dir/2016/680/oj/eng>
8. Eskens, S. (2022). The Ever-Growing Complexity of the Data Retention Discussion in the EU: An In-Depth Review of La Quadrature du Net and Others and Privacy International Joined Cases C-511/18, C-512/18 and C-520/18 La Quadrature du Net and Others [2020] ECLI: EU: C: 2020: 791; Case C-623/17 Privacy International [2020] ECLI: EU: C: 2020: 790. *Eur. Data Prot. L. Rev.*, 8, 143.
9. European Union Agency for Fundamental Rights. (2018). *Handbook on European data protection law: 2018 Edition*. Council of Europe.
10. Nirdesh Pachoriya, (2022). *Explainable AI Based Reliability*

- Analytics for Performance Optimization in Large Scale Cloud Services. *Computer Fraud And Security*, 2022 (12), pp. 38-49.
11. Ferguson, A. G. (2017). The rise of big data policing: Surveillance, race, and the future of law enforcement. In *The rise of big data policing*. New York University Press.
 12. Glover, P. (2023). Regulating Strategic Communications Investigation: The Investigatory Powers Act 2016 (UK) as an Exemplar for India. In *Comparative Approaches in Law and Policy* (pp. 337-360). Singapore: Springer Nature Singapore.
 13. Grother, P. (2022). Face recognition vendor test (FRVT) part 8: Summarizing demographic differentials. *National Institute of Standards and Technology (NIST)*, 8429, 8.
 14. Grother, P., Ngan, M., & Hanaoka, K. (2019). Face recognition vendor test (fvrt): Part 3, demographic effects. Gaithersburg, MD: National Institute of Standards and Technology.
 15. Investigatory Powers Act 2016. (2016). [Legislation.gov.uk. https://www.legislation.gov.uk/ukpga/2016/25/contents](https://www.legislation.gov.uk/ukpga/2016/25/contents)
 16. Joined Cases, C. (2014). 293/12 and C-594/12 *Digital Rights Ireland Ltd v Minister for Communications, Marine and Natural Resources and Others and Kärntner Landesregierung and Others* [2014] para, 47.
 17. Kroll, J. A. (2015). *Accountable algorithms* (Doctoral dissertation, Princeton University).
 18. Niradesh Pachoriya, (2023). *Autonomous Performance Engineering Framework Using Artificial Intelligence for Resilient* 10.48047/jocaaa.2024.33.08.480
 - Cloud Native Systems. *Membrane Technology*, 2023(6), pp. 18-29.
 19. Leiser, M., & Custers, B. (2019). The law enforcement directive: Conceptual challenges of EU directive 2016/680. *Eur. Data Prot. L. Rev.*, 5, 367.
 20. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 2053951716679679.
 21. No, G. C. (1988). 16: Article 17 (Right to Privacy), *The Right to Respect of Privacy, Family, Home and Correspondence, and Protection of Honour and Reputation*. UN Human Rights Committee (HRC), 8, 3-4.
 22. Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. In *The black box society*. Harvard university press.
 23. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., ... & Barnes, P. (2020, January). 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).
 24. Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems - Freedom of Expression - [www.coe.int](https://www.coe.int/en/web/freedom-expression/committee-ofministers-adopted-texts/asset_publisher/aDXmrol0vvsU/content/recommendation-cm-rec-). (2020). Freedom of Expression. https://www.coe.int/en/web/freedom-expression/committee-ofministers-adopted-texts/asset_publisher/aDXmrol0vvsU/content/recommendation-cm-rec-

20201-of-the-committee-of-ministers-
to-member-states-on-the-human-rights-
impacts-of-algorithmic-systems

25. Regulation, P. (2018). General data protection regulation. *Intouch*, 25, 1-5.

26. Taorui Guan, "Evidence-Based Patent Damages," 28 *Journal of Intellectual Property Law* (2020), 1-61.

27. UN Human Rights Committee. (2011). General comment No. 34: Article 19: Freedoms of opinion and expression. *International Covenant on Civil and Political Rights*.

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

29. Uppuluri, V. (2019). The Role of Natural Language Processing (NLP) in Business Intelligence (BI) for Clinical Decision Support. *ISCSITR INTERNATIONAL JOURNAL OF BUSINESS INTELLIGENCE (ISCSITR-IJBI)*, 1(2), 1-21.

30. v Russia, R. Z. (2015). Application no 47143/06. Concurring opinion of Judge Dedov.

31. Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International data privacy law*, 7(2), 76-99.

32. Yeung, K. (2020). Recommendation of the council on artificial intelligence (OECD). *International legal materials*, 59(1), 27-34.

33. Zalnieriute, M. (2022). Big brother watch and others v. the United Kingdom. *American Journal of International Law*, 116(3), 585-592.