Lyseiuk, A. (2026). Artificial intelligence in intelligence and security services: Regulatory and ethical challenges of balancing operational efficiency with democratic oversight. Politics & Security, 16(2), 22–32.
https://doi.org/10.54658/ps.28153324.2026.16.2.pp.22-32

ARTIFICIAL INTELLIGENCE IN INTELLIGENCE AND SECURITY SERVICES: REGULATORY AND ETHICAL CHALLENGES OF BALANCING OPERATIONAL EFFICIENCY WITH DEMOCRATIC OVERSIGHT

Andrii Lyseiuk
National Academy of the Security Service of Ukraine
Kyiv, Ukraine
https://orcid.org/0000-0002-9026-1188
andrii_lyseiuk@sci-univ.com

Abstract. *Purpose. The article examines how democratic states can govern the use of artificial intelligence in intelligence and security services without undermining operational effectiveness. It addresses the tension between AI-enabled speed, scale, and pattern detection, on the one hand, and legality, accountability, privacy, and democratic oversight, on the other. Methods. The study applies qualitative comparative legal and policy analysis. It combines review of regulatory instruments, intelligence oversight models, and AI ethics literature with a structured comparison of the European Union, the United States, the United Kingdom, and Israel. The article also develops an original analytical model, the Security–Oversight Balance framework. Findings. AI is already reshaping intelligence work through open-source intelligence analysis, signals intelligence triage, biometric identification, predictive threat analytics, cyber intelligence, and decision-support tools. Existing governance regimes are uneven. The EU has adopted a rights-based risk model through the AI Act, but national security exemptions remain significant. The United States relies more heavily on executive policy, agency guidance, and national security memoranda. The United Kingdom has a mature investigatory powers regime but lacks a dedicated AI-specific intelligence statute. Israel demonstrates a security-driven and innovation-oriented model with more limited public transparency. Across all cases, the central risks are algorithmic bias, opacity, mission creep, weak auditability, and the displacement of responsibility from human decision-makers to technical systems. Conclusions. Democratic oversight of AI in intelligence should not be treated as an obstacle to security capacity. It is a condition of legitimate and sustainable security governance. The proposed Security–Oversight Balance framework recommends four linked safeguards: legal authorization, technical auditability, managerial accountability, and ethical proportionality. The article argues for classified but reviewable AI registers, mandatory human accountability for consequential decisions, independent technical audit capacity, and international minimum standards for AI use in security services.*

Keywords: artificial intelligence in security; intelligence services regulation; democratic oversight; algorithmic surveillance; AI ethics in public security; national security governance.

1. INTRODUCTION

Artificial intelligence has become a strategic technology for intelligence and security services. Machine learning systems can process satellite imagery, intercepted communications, financial flows, social media

content, cyber telemetry, biometric records, and multilingual open-source materials at a speed and scale unavailable to human analysts. In operational terms, this creates clear advantages: AI can detect patterns, prioritize weak signals, support early warning, reduce analytic overload, and accelerate decision cycles in counterterrorism, cyber defence, border security, and military intelligence. Intelligence services have always depended on information asymmetry, but AI changes both the volume of available information and the institutional capacity to convert data into actionable assessments.

This transformation creates a democratic dilemma. The same systems that improve intelligence capacity may also expand state surveillance, normalize predictive suspicion, and weaken traditional safeguards. AI-based systems are not neutral instruments. They are trained on historical data, embedded in institutional cultures, shaped by procurement choices, and often protected by secrecy, intellectual property claims, and national security classifications. In intelligence contexts, these features are intensified. A person affected by an AI-generated risk score, watchlist nomination, biometric match, or targeting recommendation may never know that such a system was used. Courts, parliaments, inspectors general, and data protection authorities may also lack the technical expertise or access needed to evaluate the system properly.

The core problem is therefore not whether intelligence services should use AI. Democratic states cannot ignore technologies that adversaries, criminal networks, and authoritarian regimes already exploit. The real question is how AI can be integrated into intelligence operations while preserving democratic legitimacy. As Gill and Phythian (2018) argue, intelligence in democratic societies requires both effectiveness and control. Security services must be capable enough to protect the state, but constrained enough not to become a threat to the constitutional order they are meant to defend.

This article addresses the following research question: how can democratic states operationalize effective and legitimate control over the use of AI in intelligence operations without undermining security capacity? The article proceeds in nine steps. First, it explains the methodology. Second, it develops a typology of AI applications in intelligence and security operations. Third, it analyses the regulatory landscape in the EU, the United States, the United Kingdom, and Israel. Fourth, it identifies the main ethical challenges. Fifth, it compares democratic oversight mechanisms. Sixth, it proposes the Security–Oversight Balance framework. Seventh, it offers policy recommendations. The conclusion argues that the future of democratic intelligence governance depends on transforming oversight from a reactive legal check into a continuous socio-technical process.

2. METHODOLOGY

The article uses qualitative comparative analysis of legal, policy, and institutional materials. The research design combines three methods: normative document analysis, comparative institutional analysis, and conceptual framework development. Normative document analysis is applied to binding and non-binding instruments regulating AI, surveillance, data protection, and intelligence oversight. These include Regulation (EU) 2024/1689, the General Data Protection Regulation, the UK Investigatory Powers Act 2016, U.S. intelligence community AI principles, U.S. federal AI policy documents, the Council of Europe Framework Convention on Artificial Intelligence, the UNESCO Recommendation on the Ethics of Artificial Intelligence, NATO AI strategies, and the NIST AI Risk Management Framework.

The comparative component focuses on four democratic or democratic-allied jurisdictions: the European Union, the United States, the United Kingdom, and Israel. These cases were selected because each combines advanced security capacity with distinct legal and institutional approaches to AI governance. The EU represents a rights-based and risk-based regulatory model. The United States represents an executive-policy and agency-guidance model, complicated by changes across administrations. The United Kingdom represents a mature surveillance-law model structured around warrants, commissioners, and parliamentary review. Israel represents a security-intensive innovation model in which AI policy has developed alongside strong national security imperatives.

The article does not conduct empirical testing of classified AI systems. That limitation is unavoidable because intelligence AI deployments are often not publicly disclosed. Instead, the article evaluates governance design: how legal rules, oversight institutions, technical safeguards, and ethical principles could be combined to control high-risk AI uses. The proposed Security–Oversight Balance framework is therefore conceptual and prescriptive. It is intended as a tool for lawmakers, oversight bodies, and security institutions rather than as an empirical measurement instrument.

3. AI IN INTELLIGENCE OPERATIONS: A TYPOLOGY

AI in intelligence services should not be treated as a single technology. It is better understood as a family of systems embedded across the intelligence cycle: collection, processing, analysis, dissemination, operational planning, and post-operation review. The governance problem differs depending on whether AI merely filters large datasets, identifies persons, predicts conduct, recommends operational action, or autonomously triggers consequences.

A first major field is open-source intelligence. OSINT now includes social media posts, geolocation data, commercial satellite imagery, leaked databases, public procurement records, videos, messaging platforms, and digital traces from conflict zones. AI tools support entity recognition, translation, sentiment analysis, image classification, geospatial matching, network analysis, and disinformation detection. Recent OSINT research shows that AI can improve the speed of extracting signals from public data, but it also increases the risk of misinterpretation, source contamination, and overconfidence in pattern recognition (Browne et al., 2024). In conflicts such as Russia's war against Ukraine, OSINT has become operationally relevant, but the boundary between intelligence, journalism, activism, and information warfare has become blurred.

A second field is signals intelligence and communications analysis. Intelligence agencies collect and process vast volumes of metadata and content. AI can support language identification, voice recognition, anomaly detection, clustering of communication patterns, and prioritization of intercepts. These systems may reduce analytic overload, but they also create significant proportionality problems. The more automated the filtering process becomes, the more difficult it is to determine whether collection remains targeted or becomes functionally indiscriminate.

A third field is biometric intelligence. Facial recognition, gait analysis, voice identification, iris recognition, and other biometric tools allow security services to identify persons across physical and digital environments. Biometric systems can be valuable for identifying terrorism suspects, locating missing persons, or verifying identities at borders. Yet biometric AI is among the most rights-sensitive applications because it links bodies to state databases. Studies of commercial facial analysis have shown intersectional accuracy disparities, especially for darker-skinned women (Buolamwini & Gebru, 2018). In security contexts, such error disparities can produce wrongful suspicion, discriminatory targeting, or unlawful watchlisting.

A fourth field is predictive threat analytics. These systems seek to forecast possible threats by analysing behavioural, geographic, social, or transactional data. They may be used to prioritize investigations, identify radicalization risks, allocate patrols, detect insider threats, or predict cyberattacks. Predictive systems are attractive because they promise anticipation rather than reaction. However, predictive policing literature warns that such models may reproduce historical enforcement patterns and create feedback loops: areas or communities previously subject to more policing generate more recorded data, which then justifies further policing (Brayne, 2017; Ferguson, 2017).

A fifth field is cyber intelligence. AI can detect malware patterns, classify phishing campaigns, identify anomalies in network traffic, map adversary infrastructure, and support automated response to cyber incidents. Compared with biometric or predictive policing systems, cyber AI may seem less directly intrusive. Nevertheless, cyber intelligence often involves large-scale monitoring, cross-border data flows,

and cooperation with private infrastructure providers. It therefore raises oversight questions about authorization, necessity, and collateral access to personal or commercial data.

A sixth emerging field is AI-assisted decision support. Generative and agentic systems can summarize intelligence reports, generate hypotheses, simulate adversary behaviour, draft threat assessments, or recommend operational options. These tools may improve productivity, but they also introduce risks of hallucination, automation bias, and hidden dependence on systems that cannot reliably explain their outputs. In intelligence work, a plausible but false AI-generated synthesis can be more dangerous than an obviously incomplete report because it may appear authoritative.

Table 1. Artificial intelligence applications in intelligence and security services: operational benefits and main oversight risks

AI application	Operational benefit	Main oversight risk
OSINT analysis	Rapid processing of public and semi-public data	Misclassification, source contamination, weak verification
Signals intelligence triage	Prioritization of large intercept datasets	Expansion of bulk collection and opaque filtering
Biometric identification	Faster person identification and identity verification	Discrimination, false matches, continuous public surveillance
Predictive threat analytics	Early warning and resource allocation	Feedback loops, profiling, predictive suspicion
Cyber intelligence	Faster anomaly detection and threat attribution	Overbroad monitoring and unclear authorization
AI decision support	Faster synthesis and scenario generation	Automation bias, hallucination, diluted responsibility

This typology shows that governance cannot rely on a generic statement that "AI is supervised by humans." Oversight must be matched to function. A language translation model used for OSINT triage does not require the same control as a biometric watchlist system or a predictive threat model. The democratic challenge is to classify AI uses by operational intensity and rights intrusiveness, then attach appropriate safeguards.

4. THE REGULATORY LANDSCAPE

The regulatory landscape is fragmented. No major democracy has yet adopted a complete AI-specific intelligence statute. Instead, AI in intelligence is governed through overlapping layers: AI regulation, data protection law, surveillance law, constitutional rights, procurement rules, classified directives, and internal agency policies.

The European Union has adopted the most comprehensive horizontal AI regulation. Regulation (EU) 2024/1689 establishes a risk-based model, including prohibited practices, obligations for high-risk systems, transparency duties, and rules for general-purpose AI. Law enforcement and biometric systems appear in high-risk categories, and certain forms of real-time remote biometric identification in publicly accessible spaces are restricted. However, the AI Act also contains exclusions for military, defence, and national security purposes. This creates a central tension: the EU has a strong rights-based AI framework, but some of the most security-sensitive AI uses may fall partly outside its scope.

The GDPR remains relevant because AI systems often process personal data. Article 23 allows restrictions on certain rights and obligations when necessary and proportionate for national security, defence, public security, or criminal law purposes. This means that data protection rights may be limited in intelligence contexts, but not eliminated. The key legal test is necessity and proportionality. For AI-enabled intelligence, this implies that secrecy cannot become a blanket justification. Even where individual notification is impossible, independent bodies must be able to examine whether data processing is lawful, limited, and auditable.

The United States follows a more fragmented model. The Office of the Director of National Intelligence adopted Principles of Artificial Intelligence Ethics for the Intelligence Community and an AI Ethics Framework in 2020. These documents emphasize lawfulness, accountability, transparency, reliability, security, and human judgment. The NIST AI Risk Management Framework provides voluntary, rights-preserving guidance for identifying and managing AI risks. U.S. policy has also changed significantly since the 2023 Executive Order 14110. That order was rescinded in January 2025 and replaced by a more innovation- and competitiveness-oriented policy direction under Executive Order 14179. The 2025 America's AI Action Plan and OMB Memorandum M-25-21 emphasize accelerated federal AI adoption, governance, and public trust, while the 2026 National Security Presidential Memorandum on AI in the National Security Enterprise focuses directly on military and intelligence adoption.

The U.S. model therefore combines strong technical capacity with executive volatility. Agency-level AI safeguards may be sophisticated, but many are policy-based rather than statutory. Oversight occurs through congressional intelligence committees, inspectors general, the Foreign Intelligence Surveillance Court, privacy and civil liberties offices, and internal compliance mechanisms. However, classified AI systems may remain difficult for external actors to evaluate, especially when model performance, training data, or vendor contracts are not transparent.

The United Kingdom has a more consolidated surveillance-law regime. The Investigatory Powers Act 2016 governs interception, equipment interference, communications data acquisition, bulk powers, and oversight mechanisms. It introduced safeguards such as the "double lock" model, involving ministerial authorization and approval by a Judicial Commissioner for certain warrants. The Investigatory Powers Commissioner's Office oversees the use of investigatory powers, while the Intelligence and Security Committee of Parliament scrutinizes intelligence agencies. This structure is valuable for AI governance because it already contains mechanisms for classified review. However, the UK still lacks a comprehensive AI-specific framework for intelligence systems. Existing warrant and oversight models must therefore be adapted to evaluate algorithmic tools, not only human decisions to collect or intercept data.

Israel represents a different model. It is a technologically advanced security state with strong AI innovation capacity and high exposure to terrorism, cyber threats, and regional military conflict. Israel's 2023 AI regulation and ethics policy emphasizes responsible innovation, sectoral regulation, and risk management. However, the public legal framework for intelligence AI remains less transparent than in the EU or UK. This is partly a function of Israel's security environment and partly a feature of institutional design. The Israeli case illustrates a broader challenge: democracies facing acute security threats may prioritize operational flexibility, but this makes independent oversight even more important.

Table 2. Comparative regulatory approaches to AI in intelligence and security services: the European Union, the United States, the United Kingdom, and Israel

Jurisdiction	Regulatory model	Strength	Weakness
EU	Horizontal AI Act plus data protection law	Strong rights-based framework	National security exclusions and enforcement complexity
United States	Executive policy, agency guidance, courts, congressional oversight	Technical sophistication and agency frameworks	Policy volatility and fragmented statutory basis
United Kingdom	Surveillance statute with commissioners and parliamentary oversight	Mature warrant and review architecture	Limited AI-specific legal standards
Israel	Sectoral AI policy and security-driven governance	Innovation capacity and operational integration	Lower public transparency in intelligence AI

Across these models, the central regulatory gap is not the absence of principles. Lawfulness, accountability, transparency, fairness, reliability, and human oversight appear repeatedly in AI governance instruments. The gap is operationalization. Intelligence agencies need procedures that translate principles

into authorization thresholds, logs, audit trails, bias testing, red-team evaluations, procurement conditions, and reviewable records.

5. ETHICAL CHALLENGES

The first ethical challenge is bias and discrimination. AI systems learn from data generated by unequal societies and institutions. In policing and intelligence, historical data often reflect enforcement priorities, surveillance patterns, reporting biases, and political decisions. If such data are used to train predictive models, the system may treat past suspicion as evidence of future risk. Barocas and Selbst (2016) describe how data-driven systems can produce disparate impact even without explicit discriminatory intent. Selbst et al. (2019) further argue that fairness cannot be solved by technical metrics alone because AI systems operate within socio-technical contexts.

The second challenge is accountability. Intelligence work already involves secrecy and compartmentalization. AI can multiply this opacity. When an analyst acts on a model output, responsibility may be distributed among software developers, data engineers, procurement officers, commanders, analysts, and legal advisers. Kroll et al. (2017) argue that algorithmic accountability requires mechanisms beyond simple transparency, including technical design choices that allow systems to be tested and governed. In intelligence, this means that every consequential AI output should be traceable to a system version, dataset, operational purpose, authorization, and human decision-maker.

The third challenge is privacy. AI makes it possible to infer sensitive facts from apparently ordinary data. Location patterns, online behaviour, metadata, and social networks can reveal religion, political views, health conditions, intimate relationships, or professional associations. The OHCHR has warned that AI-enabled profiling and automated decision-making can seriously affect the right to privacy and associated rights. For intelligence services, the ethical issue is not only whether data were lawfully obtained, but whether AI use changes the nature of the intrusion. A dataset collected for one purpose may become far more intrusive when combined with other datasets and analysed by machine learning systems.

The fourth challenge is mission creep. Security tools introduced for exceptional threats often migrate into ordinary governance. Facial recognition deployed for counterterrorism may later be used for public order, immigration enforcement, or minor criminal investigations. Predictive analytics designed for national security may be repurposed for welfare fraud detection or protest monitoring. Lyon (2018) describes surveillance as a cultural and institutional process, not merely a set of devices. AI intensifies this process because the marginal cost of reusing models and datasets is low.

The fifth challenge is explainability. Some AI models produce outputs that cannot be easily explained in legal or operational terms. Intelligence agencies may not need full public transparency, but they do need reviewability. An oversight body should be able to ask why a person was flagged, what variables mattered, whether protected characteristics were directly or indirectly used, what error rates apply, and whether the model was valid for the relevant population. Wachter et al. (2017) argue that explanations must be meaningful to affected persons and institutions, not merely technical descriptions of code.

The sixth challenge is automation bias. Analysts and commanders may over-trust AI outputs, especially when systems are presented as objective or statistically advanced. In high-pressure security environments, speed can become a substitute for judgment. AI should support human reasoning, not replace it in decisions that affect liberty, life, reputation, or political rights. Human-in-the-loop governance is insufficient if the human merely rubber-stamps machine output. The relevant standard should be meaningful human responsibility.

6. DEMOCRATIC OVERSIGHT MECHANISMS

Democratic oversight of intelligence traditionally combines parliamentary, judicial, executive, and independent review. AI does not replace these mechanisms, but it requires them to evolve.

Parliamentary oversight provides democratic legitimacy. Committees can review budgets, mandates, legal authorities, strategic priorities, and major failures. The UK Intelligence and Security Committee and U.S. congressional intelligence committees illustrate this function. However, parliamentary bodies often lack technical capacity and depend on information supplied by the agencies they oversee. For AI governance, parliamentary oversight should include access to independent technical advisers, classified AI inventories, procurement contracts, and aggregate performance metrics.

Judicial oversight is essential where surveillance interferes with rights. Courts or judicial commissioners can authorize intrusive powers, assess necessity and proportionality, and review complaints. The U.S. Foreign Intelligence Surveillance Court and the UK double-lock warrant system provide examples of judicialized intelligence control. Yet AI creates new questions for judicial review. A judge may authorize interception, but not know that AI tools will later search, score, or correlate the collected data in ways that expand the intrusion. Warrants should therefore specify not only collection powers but also AI-enabled processing when such processing materially changes the risk to rights.

Independent oversight bodies are especially important because they can combine continuity, expertise, and classified access. Inspectors general, data protection authorities, privacy commissioners, and investigatory powers commissioners can examine internal systems more deeply than parliaments or courts. The challenge is technical capacity. Oversight bodies need staff who understand model validation, dataset bias, cybersecurity, adversarial manipulation, and audit logs. Without this capacity, oversight risks becoming formal rather than substantive.

Internal agency governance also matters. Democratic control does not occur only outside security services. Intelligence agencies should maintain AI review boards, legal compliance units, model risk committees, red teams, and ethics officers. However, internal governance cannot substitute for external oversight. It should generate the records and technical evidence that external bodies can inspect.

Civil society and media oversight are limited in intelligence contexts because secrecy is unavoidable. Still, investigative journalism, academic research, strategic litigation, and NGO reports often reveal abuses or governance gaps. Privacy International, Amnesty International, and other organizations have played a significant role in identifying the risks of surveillance technologies. Democratic systems should protect responsible disclosure and avoid using secrecy to shield unlawful or unethical AI practices.

The most effective model is layered oversight. No single institution can govern AI in intelligence. Parliaments can authorize and scrutinize; courts can approve intrusive powers; independent bodies can audit; internal units can manage risk; civil society can challenge public failures. AI governance should connect these layers through documentation, reporting duties, and escalation procedures.

7. THE SECURITY–OVERSIGHT BALANCE FRAMEWORK

The Security–Oversight Balance framework is designed to help democratic states evaluate AI uses in intelligence and security services. It rests on a simple premise: the level of oversight should increase with both operational intensity and rights intrusiveness. A low-risk translation tool used by analysts does not require the same controls as a biometric identification system deployed in public space. A predictive model used for strategic assessment does not require the same controls as a model that nominates individuals for coercive action.

The framework has four dimensions.

The first dimension is legal authorization. Each AI use should have a clear legal basis, defined operational purpose, authorized user group, data source limits, retention rules, and review procedure. National security secrecy may justify withholding information from the public, but it should not justify legal ambiguity. If an AI system changes the practical scope of a surveillance power, the authorization should reflect that change.

The second dimension is technical auditability. AI systems used in intelligence should generate records that allow later review. These include model version, training and validation data summaries, known limitations, performance metrics, bias testing, prompts or queries where relevant, output logs, human interventions, and override decisions. Technical auditability does not require publishing classified models. It requires making systems inspectable by authorized oversight bodies.

The third dimension is managerial accountability. Every consequential AI use should have an accountable human owner. Agencies should define who approved the system, who validated it, who may use it, who monitors performance, and who is responsible when harm occurs. Calo and Citron (2021) warn that automation can undermine the legitimacy of administrative action when agencies outsource or obscure judgment. Intelligence agencies must therefore avoid "the model decided" reasoning. AI may inform decisions, but it should not become an institutional shield.

The fourth dimension is ethical proportionality. AI use should be evaluated by necessity, suitability, proportionality, discrimination risk, and reversibility. The more difficult it is for affected persons to know, challenge, or correct an AI-generated consequence, the stronger the ex ante safeguards must be. Ethical proportionality also requires attention to group harms. A community may be disproportionately surveilled even if no single individual can prove a rights violation.

Table 3. The Security–Oversight Balance framework: four oversight tiers for AI use.

Tier	AI use	Required safeguards
Tier 1: Administrative support	Translation, summarization, document search	Internal validation, user training, data security
Tier 2: Analytic support	OSINT clustering, cyber anomaly detection, strategic forecasting	Model logs, analyst verification, periodic audit
Tier 3: Person-affecting support	Watchlist support, biometric search, risk scoring	Legal authorization, bias testing, independent review, human justification
Tier 4: Coercive or operational action support	Targeting, arrest prioritization, intrusive surveillance triggers	Prior authorization, strict necessity test, technical audit, post-operation review

The value of the SOB framework is that it avoids two extremes. It does not prohibit all AI in intelligence, which would be unrealistic and potentially harmful to democratic security. It also rejects unrestricted technological adoption, which would erode legitimacy. The framework treats oversight as a design requirement. AI systems should be built from the beginning to be reviewable, contestable where possible, and accountable at institutional level.

8. POLICY RECOMMENDATIONS

First, legislatures should require classified AI registers for intelligence and security services. These registers should list AI systems, operational purposes, legal bases, data categories, vendors, risk tiers, and responsible officials. Public disclosure may be limited, but parliamentary and independent oversight bodies should have access to the full register.

Second, warrants and authorizations should address AI-enabled processing. If AI materially expands the ability to search, correlate, identify, or predict from collected data, this should be treated as a distinct rights-relevant step. Judicial or commissioner approval should not be limited to the initial act of collection.

Third, intelligence agencies should establish AI assurance units. These units should test models before deployment, monitor drift, conduct bias and robustness evaluations, review incidents, and maintain audit records. They should include legal, technical, operational, and ethics expertise.

Fourth, procurement contracts should include audit rights. Governments should not purchase black-box systems that cannot be inspected by authorized public bodies. Vendors should be required to provide

documentation, performance evidence, cybersecurity assurances, and information on training data limitations.

Fifth, human accountability should be mandatory for consequential decisions. AI outputs should not be sufficient grounds for watchlisting, arrest, targeting, intrusive surveillance, or adverse administrative action. Human decision-makers should record their reasoning when relying on AI outputs.

Sixth, NATO and the EU should develop minimum standards for AI in intelligence cooperation. Intelligence sharing increasingly involves shared platforms and interoperable systems. If one partner uses weakly governed AI outputs, the risk may spread across alliances. Minimum standards should include auditability, proportionality, reliability, and bias mitigation.

Seventh, oversight bodies need technical capacity. Parliamentary committees, courts, commissioners, and inspectors general should be supported by independent AI experts with security clearances. Without technical expertise, oversight cannot evaluate whether AI systems actually perform as claimed.

9. CONCLUSIONS

AI is transforming intelligence and security services by increasing the speed, scale, and predictive ambition of state security operations. These capabilities can strengthen democratic resilience against terrorism, cyberattacks, hostile intelligence activity, organized crime, and foreign interference. Yet they can also intensify surveillance, reproduce discrimination, obscure responsibility, and weaken the rule of law.

The answer to the research question is that democratic states can operationalize legitimate control over AI in intelligence only by making oversight continuous, technical, and risk-sensitive. Traditional legal authorization remains necessary, but it is no longer sufficient. Oversight must extend into datasets, models, procurement, validation, deployment, human-machine interaction, and post-operation review.

The proposed Security–Oversight Balance framework offers one way to structure this process. It links legal authorization, technical auditability, managerial accountability, and ethical proportionality. Its central principle is proportional governance: the more intrusive and operationally consequential the AI system, the stronger the oversight requirements must be.

The article has limitations. It relies on publicly available materials and cannot evaluate classified deployments directly. It also focuses on democratic and democratic-allied systems, leaving authoritarian AI surveillance for separate study. Future research should examine how oversight bodies actually evaluate AI systems in practice, how classified technical audits can be designed, and how international intelligence cooperation can prevent rights laundering through less regulated partners.

Democratic oversight should not be seen as a burden on intelligence effectiveness. In the long term, legitimacy is a security asset. AI systems that cannot be explained, audited, or controlled may produce short-term operational gains, but they create strategic risks: public distrust, legal challenge, diplomatic friction, and institutional overreach. The task for democratic states is not to choose between security and oversight, but to design intelligence AI so that security remains accountable to democracy.

AI Use Declaration. *This article was originally written by the author in Ukrainian. AI-based tools were used to assist in translating the manuscript into English, in language editing and proofreading of the translated text, and in restructuring and reorganizing the material to meet the journal's requirements. The author reviewed and revised the resulting text, verified all sources and citations against the original materials, and confirms that the research, analysis, and conclusions presented are entirely the author's own. The author takes full responsibility for the accuracy, originality, and integrity of the final text.*

REFERENCES

Amoore, L. (2020). Cloud ethics: Algorithms and the attributes of ourselves and others. Duke University Press.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104(3), 671–732.

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

Brayne, S. (2021). Predict and surveil: Data, discretion, and the future of policing. Oxford University Press.

Brayne, S., & Christin, A. (2021). Technologies of crime prediction: The reception of algorithms in policing and criminal courts. Social Problems, 68(3), 608–624.

Browne, T. O., Abedin, M., & Chowdhury, M. J. M. (2024). A systematic review on research utilising artificial intelligence for open source intelligence applications. *International Journal of Information Security*, 23(4), 2911–2938. https://doi.org/10.1007/s10207-024-00868-2

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of Machine Learning Research, 81, 77–91.

Calo, R., & Citron, D. K. (2021). The automated administrative state: A crisis of legitimacy. Emory Law Journal, 70(4), 797–846.

Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. Washington Law Review, 89(1), 1–34.

Council of Europe. (2018). Protocol amending the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data. Council of Europe.

Council of Europe. (2024). Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. Council of Europe.

Deibert, R. J. (2023). Subversion Inc.: The age of private spycraft. MIT Press.

European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. Official Journal of the European Union.

European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Official Journal of the European Union.

Feldstein, S. (2021). The rise of digital repression: How technology is reshaping power, politics, and resistance. Oxford University Press.

Ferguson, A. G. (2017). The rise of big data policing: Surveillance, race, and the future of law enforcement. New York University Press.

Gill, P., & Phythian, M. (2018). Intelligence in an insecure world (3rd ed.). Polity Press.

Hildebrandt, M. (2015). Smart technologies and the end(s) of law: Novel entanglements of law and technology. Edward Elgar.

Investigatory Powers Act 2016, c. 25. (UK).

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.

Leslie, D., Burr, C., Aitken, M., Katell, M., Briggs, M., & Rincon, C. (2022). Human rights, democracy, and the rule of law assurance framework for AI systems: A proposal. The Alan Turing Institute.

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

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. https://doi.org/10.1177/2053951716679679

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

North Atlantic Treaty Organization. (2021). Summary of the NATO Artificial Intelligence Strategy. NATO.

North Atlantic Treaty Organization. (2024). Summary of NATO's revised Artificial Intelligence Strategy. NATO.

Office of the Director of National Intelligence. (2020a). Principles of artificial intelligence ethics for the intelligence community. ODNI.

Office of the Director of National Intelligence. (2020b). Artificial intelligence ethics framework for the intelligence community. ODNI.

Office of Management and Budget. (2025). Memorandum M-25-21: Accelerating federal use of AI through innovation, governance, and public trust. Executive Office of the President.

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.

Organisation for Economic Co-operation and Development. (2024). OECD AI principles. OECD.

Pasquale, F. (2015). The black box society: The secret algorithms that control money and information. Harvard University Press.

Privacy International. (2014). Who's watching the watchers? A comparative study of intelligence organisations oversight mechanisms. Privacy International.

President of the United States. (2025). Executive Order 14179: Removing barriers to American leadership in artificial intelligence. The White House.

President of the United States. (2026). National Security Presidential Memorandum/NSPM-11: Artificial intelligence in the national security enterprise. The White House.

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68. https://doi.org/10.1145/3287560.3287598

Slobogin, C., & Brayne, S. (2023). Surveillance technologies and constitutional law. Annual Review of Criminology, 6, 219–240.

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. UNESCO.

United Nations High Commissioner for Human Rights. (2021). The right to privacy in the digital age: Report of the United Nations High Commissioner for Human Rights (A/HRC/48/31). United Nations.

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. https://doi.org/10.1093/idpl/ipx005

White House. (2025). America's AI Action Plan. Executive Office of the President.