

AI in National Security: A Secure and Explainable Governance Framework

AMULYA C

DEPARTMENT OF COMPUTER
SCIENCE (CYBER SECURITY)

RV COLLEGE OF ENGINEERING

(AUTONOMOUS INSTITUTE AFFILIATED TO VTU,
BELAGAVI)

amulyac.cy23@rvce.edu.in

SNEHA PATEL

DEPARTMENT OF COMPUTER
SCIENCE (CYBER SECURITY)

RV COLLEGE OF ENGINEERING

(AUTONOMOUS INSTITUTE AFFILIATED TO VTU,
BELAGAVI)

snehapatel.cy23@rvce.edu.in

VIANCA CHOUDHARY

DEPARTMENT OF COMPUTER
SCIENCE (CYBER SECURITY)

RV COLLEGE OF ENGINEERING

(AUTONOMOUS INSTITUTE AFFILIATED TO VTU,
BELAGAVI)

viancac.cy23@rvce.edu.in

Dr. CHITRA B. T

DEPARTMENT OF INDUSTRIAL ENGINEERING
AND MANAGEMENT (IEM)

RV COLLEGE OF ENGINEERING

(AUTONOMOUS INSTITUTE AFFILIATED TO VTU,
BELAGAVI)

chitrabt@rvce.edu.in

Abstract— Artificial Intelligence (AI) is increasingly used in national security for surveillance, cyber defense, threat prediction, and intelligence analysis. However, the rapid growth of AI in defense systems has created major concerns related to privacy violations, biased decision-making, misuse of surveillance, lack of transparency, and theft of sensitive AI models. Existing research focuses on improving AI efficiency, while limited attention is given to governance, ethical accountability, explainability, and intellectual property protection of national-security AI systems. This paper proposes a Secure and Explainable AI Governance Framework that integrates human oversight, explainable AI, anti-manipulation security measures, and IPR protection to ensure responsible, transparent, and secure use of AI in national security operations.

Keywords—Artificial Intelligence (AI), National Security, Intelligence Analysis, Explainable AI (XAI), Intellectual Property Rights (IPR), Ethical Governance, Cybersecurity, Human-in-the-Loop, Surveillance Systems, AI Governance Framework

I. INTRODUCTION

Artificial Intelligence (AI) has become one of the most important technologies in modern national security and intelligence analysis, it is used everywhere. Governments across the world are increasingly leaning on AI-powered systems for surveillance, facial recognition, cyber defense, border monitoring, predictive risk analysis, and intelligence

gathering. In practice these tools help security agencies chew through massive amounts of data fast, and they also support better choices during those critical moments [1], [2].

AI has also improved operational effectiveness in defense and intelligence areas, a lot. Advanced machine learning algorithms can pick out odd or questionable activities, spot cyber threats, and work through intelligence data more efficiently than older or more traditional systems [10]. Because of this many countries are investing heavily in AI based national-security infrastructure, to strengthen defensive capacities and at the same time improve public safety [22].

Still, the quick expansion of AI in national security raises serious ethical, legal, and security matters. A lot of AI systems behave like black-box models where the logic behind decisions is not visible or transparent [7]. That kind of lack of explainability may end up fueling biased surveillance, wrongful targeting, privacy violations, and even misuse of citizen data [3], [11]. On top of that adversarial attacks, deep-fake intelligence, and unauthorized access to sensitive defense algorithms create major cybersecurity hazards for these national-security systems [12], [24].

An additional issue is there being no solid governance setup, plus not enough intellectual property protection for military AI systems. Most existing research tends to concentrate on boosting AI efficiency, automation, and performance. But there is comparatively less focus on ethical accountability, explainability, and ownership protection of the AI models that get used during defense operations [15], [16]. In India as

adoption of AI-enabled surveillance systems keeps rising, privacy protection, transparency, and responsible use of these technologies also turn into ongoing concerns.

So overall there is a clear demand for a secure and explainable governance framework, one that can ensure ethical AI use, human oversight, cybersecurity protection, and intellectual property rights (IPR) protection for AI-based national-security systems. This paper therefore proposes a Secure and Explainable AI Governance Framework for responsible deployment of AI in national security and intelligence analysis.

II. OBJECTIVES OF THE STUDY

- To investigate the part of Artificial Intelligence (AI) in national security, and intelligence analysis, you know the whole sphere.
- To sift through ethical, legal, and cyber security problems that come with AI powered surveillance and intelligence systems, how it all gets messy.
- To explore privacy violations, skewed decision making, deep-fake intelligence, and AI manipulation attacks, all of that.
- To spot research holes in governance, explainability, accountability, and intellectual property rights (IPR) protection for defense AI systems.
- To suggest a Secure and Explainable AI Governance Framework, so the use of AI in national security operations stays responsible and transparent.

III. RESEARCH METHODOLOGY

This present study really leans on secondary data, sourced from research papers, journals, government reports, various websites, cybersecurity publications, and articles touching Artificial Intelligence (AI) with national security, explainable AI, cybersecurity, and intellectual property rights (IPR). Overall, the work uses something like an analytical-and-descriptive research approach to look at ethical, legal, and governance type difficulties that come with AI driven intelligence systems, in practice.

The research is mostly aimed at concerns like AI surveillance, privacy intrusions, adversarial attacks, deepfake type intelligence, weak transparency or well, the missing visibility, and how to safeguard sensitive defense AI models. After going through what the literature already says and comparing it with today's national-security pressures, the study suggests a Secure and Explainable AI Governance Framework, for a more responsible and transparent rollout of AI technologies inside intelligence analysis plus defense operations, even when things are at high stakes.

IV. RELATED WORK

Artificial Intelligence (AI) is getting used more in national security and intelligence analysis, like surveillance, cyber defense, facial recognition, border monitoring, and predictive threat detection [1], [2]. In many cases, governments and defense organizations depend on AI systems to sort through huge volumes of data, flag strange patterns, and help with quicker decision making during critical scenes [10], [22]. In fact, a lot of existing research still concentrates on boosting AI efficiency, automation, and accuracy in intelligence workflows.

Some researchers have also gone into transparency related problems, plus ethical worries. Doshi-Velez and Kim [7] mention that many AI systems act like black box models, which makes their decisions tricky to understand. DARPA's Explainable AI program [8] points out that transparent AI decision making matters a lot, especially in defense and intelligence where risk is higher. Work by Wachter et al. [3] along with UNESCO [15] discusses things such as biased surveillance, privacy violations, and the misuse of citizen data when AI systems are not properly controlled.

Then there are security threats targeting AI systems, which is another key point. Adversarial attacks can tweak surveillance data, and deepfake intelligence may end up undermining national security operations [12], [24]. In India, the increasing deployment of AI enabled surveillance and facial recognition has triggered worries about privacy, transparency, and ethical governance. Yet, even with these discussions, the current body of research still gives limited attention to intellectual property rights (IPR), governance structures, and accountability for AI systems that are used in national security contexts. So, in a way, a secure plus explainable AI governance framework becomes necessary for using AI in intelligence analysis and defense operations responsibly, and not just in theory but in a measurable, explainable manner.

V. RESEARCH GAP

Existing studies on Artificial Intelligence (AI) in national security mostly zoom in on boosting surveillance accuracy, cyber defense mechanisms, predictive threat detection, and automated intelligence analysis [1], [10]. In general, a lot of the work leans toward technical improvements, and operational efficiency of AI based systems in military and intelligence contexts [2], [22]. Still, not much attention is paid to governance, explainability, ethical accountability, and even intellectual property protection for AI systems when they get used in national security operations [15], [16].

An extensive portion of AI driven intelligence systems still behave like black-box models, where the whole decision process is not visible to security agencies or analysts [7], [8]. This kind of opacity can bring specific dangers, like skewed surveillance, misguided targeting, wrong threat predictions, and misuse of citizen data [3], [11]. On top of that, the

current literature seems to discuss too little about human supervision, and the accountability setups around AI assisted national security decisions [25].

Another major research gap is straightforward, but it keeps showing up—there is no solid framework for intellectual property rights (IPR) protection, or reliable ownership mechanisms for defense AI models. Many studies barely touch on patent safeguarding, secure licensing, cyber hardening, and ways to stop AI model theft or adversarial tampering [16], [17]. And meanwhile, AI technology is moving forward at a speed that legal frameworks, ethical policies, and privacy rules cannot keep up with [4], [15].

In India, the growing adoption of AI-enabled surveillance tools and facial recognition technologies has pushed privacy questions, transparency demands, and ethical governance concerns more into the open. Cybersecurity incidents that target critical infrastructure also make it clearer that AI manipulation, and unauthorized access to sensitive intelligence systems, can create serious risks. So, there is really a need for a secure and explainable AI governance framework that folds in ethical monitoring, human oversight, anti-manipulation security defenses, and intellectual property protection for responsible national security operations.

VI. AI IN NATIONAL SECURITY AND INTELLIGENCE ANALYSIS

Artificial Intelligence (AI) has become important technology in modern national security, and intelligence operations too. In practice, AI systems are now widely used for surveillance, facial recognition, cyber defense, border monitoring, predictive threat analysis, and for intelligence data processing [1], [2]. These approaches tend to help security agencies sift through huge volumes of information fast, and they also improve how decisions get made during urgent situations.

AI-driven surveillance systems and predictive analytics are, by now, increasingly used by governments to spot suspicious activities, and to bolster public security in general. In the area of cybersecurity, AI supports the detection of malware, phishing attacks, network intrusions, and weird or abnormal digital patterns [18], [24]. Beyond that, AI is applied in intelligence analysis to handle data gathered from social media, communication networks, satellite imagery, and surveillance systems [10].

In India, AI-enabled surveillance along with facial recognition technologies are being adopted for public security and law-enforcement needs. Even if these tools can improve monitoring efficiency and threat detection, worries about privacy infringement, biased surveillance, misuse of citizen data, and a lack of transparency are still growing [3], [11]. This rising reliance on AI in national-security environments makes it clear that ethical governance matters, explainable decision-making should be insisted on, cybersecurity safeguards are necessary, and intellectual property rights

(IPR) protection must be in place for sensitive defense AI technologies.

VII. PROPOSED FRAMEWORK

This paper proposes a Secure and Explainable AI Governance Framework, for responsible deployment of Artificial Intelligence (AI) in national security and intelligence analysis, kind focused on keeping the whole thing controllable. The framework brings together explainable AI, human oversight, anti-manipulation security mechanisms, intellectual property protection, and ethical governance to raise transparency, accountability, and security for AI-driven defense systems.

A. Explainable AI Layer

The proposed framework puts Explainable AI (XAI) mechanisms in the center, so the transparency in AI generated intelligence decisions goes up, and yes, in a practical sense. The system provides interpretable outputs for threat predictions surveillance alerts, and intelligence recommendations, allowing security analysts to grasp why AI produced a certain result [7], [8]. That, in turn, helps cut down the risks linked to biased analysis and wrongful targeting, even when the model “looks confident.”

B. Human-in-the-Loop Intelligence

The framework makes sure AI systems behave like decision-support tools, not fully autonomous actors that just run off on their own. Critical national security decisions need human verification, ethical review, and multi-level authorization, before anything gets implemented. This method strengthens accountability and reduces too much reliance on automated intelligence systems, which is a bad idea in high stakes settings.

C. Anti-Manipulation Security Layer

To improve cybersecurity resilience, the framework includes protection mechanisms against adversarial attacks, manipulated surveillance data, deepfake intelligence, and AI poisoning attacks [12], [24]. Blockchain-based audit logs and tamper detection mechanisms are also used, so integrity and traceability of intelligence information stay intact.

D. Intellectual Property Protection

The framework suggests intellectual property rights (IPR) protection for military AI models and defense algorithms, via patent security, controlled licensing, restricted access, and cybersecurity safeguards [16], [17]. These steps aim to stop unauthorized duplication, theft, and misuse of sensitive AI technologies, keeping valuable parts from walking out the door.

E. Ethical and Legal Governance

The proposed framework stresses ethical AI deployment through privacy protection policies, accountability standards, and transparent surveillance regulations [4], [15]. Overall, it backs responsible use of AI technologies while trying to

balance national security goals with citizen rights, even though that tradeoff can feel awkward at times.

VIII. CASE STUDIES AND DISCUSSION

A. AI-Based Surveillance in India

India has increased its deployment of AI-enabled surveillance and facial recognition systems for public security and intelligence monitoring. While these technologies could enhance the efficiency of monitoring and threat detection, concerns about privacy violations, a lack of transparency and the misuse of citizens' data are increasing.

B. Deepfake and Misinformation Threats

Deepfakes and AI-manipulated digital content are serious threats to intelligence analysis and national security operations [10], [24]. Artificial media and synthetic data can shape public opinion, spread misinformation, and build false narratives of security.

C. Cybersecurity Threats to AI Systems

Surveillance data was manipulated through adversarial attacks. Intelligence systems and defense infrastructure can be compromised through AI poisoning attacks [12]. Such threats underscore the importance of cybersecurity safeguards and robust governance mechanisms for AI-enabled national security systems.

D. Discussion

The above case studies demonstrate that while AI improves intelligence analysis and surveillance effectiveness, unregulated use of AI poses ethical, legal and security issues. Accordingly, secure governance, explainable decision-making, human oversight, and intellectual property protection are essential to the responsible use of AI in national security operations.

IX. CONCLUSION

Artificial Intelligence (AI) has improved national security and intelligence analysis, through things like surveillance, cyber defense, predictive threat detection, and automated monitoring systems, but you know it is not only smooth sailing. When dependence on AI increases it also brings serious problems tied to privacy violations, biased decisions, little to no transparency, adversarial attacks, and even unauthorized access to sensitive defense technologies.

In this paper we propose a Secure and Explainable AI Governance Framework that kind of ties together explainable AI, human oversight, anti-manipulation security mechanisms, ethical governance, and intellectual property protection for AI-based national-security systems. The idea is that this framework helps enable transparent, accountable, and secure deployment of AI technologies, while at the same time lowering risks that come from uncontrolled or unethical AI usage in defense and intelligence operations.

REFERENCES

- [1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Hoboken, NJ, USA: Pearson, 2021.
- [2] N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*. Oxford, U.K.: Oxford University Press, 2014.
- [3] S. Wachter, B. Mittelstadt, and C. Russell, "Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI," *Computer Law & Security Review*, vol. 41, 2021.
- [4] European Commission, "Ethics Guidelines for Trustworthy AI," 2019. [European Commission AI Ethics Guidelines](#)
- [5] IEEE, "Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems," 2019. [IEEE Ethically Aligned Design](#)
- [6] OpenAI, "GPT-4 System Card," 2023. [OpenAI GPT-4 System Card](#)
- [7] A. Doshi-Velez and B. Kim, "Towards a Rigorous Science of Interpretable Machine Learning," arXiv preprint arXiv:1702.08608, 2017.
- [8] D. Gunning, "Explainable Artificial Intelligence (XAI)," *Defense Advanced Research Projects Agency (DARPA)*, 2017. [DARPA XAI Program](#)
- [9] B. Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. New York, NY, USA: W. W. Norton & Company, 2015.
- [10] P. W. Singer and A. Brooking, *LikeWar: The Weaponization of Social Media*. Boston, MA, USA: Houghton Mifflin Harcourt, 2018.
- [11] S. Zuboff, *The Age of Surveillance Capitalism*. New York, NY, USA: PublicAffairs, 2019.
- [12] M. Rouse, "Adversarial Attack," *TechTarget*, 2023. [TechTarget Adversarial Attack Definition](#)
- [13] IBM, "What is Explainable AI?" 2024. [IBM Explainable AI](#)
- [14] Microsoft, "Responsible AI Principles," 2024. [Microsoft Responsible AI](#)
- [15] UNESCO, "Recommendation on the Ethics of Artificial Intelligence," 2021. [UNESCO AI Ethics Recommendation](#)

[16] World Intellectual Property Organization (WIPO), “WIPO Technology Trends 2025: Artificial Intelligence,” 2025.

WIPO AI Technology Trends

[17] A. R. Mattioli, “Military Artificial Intelligence and Intellectual Property Challenges,” *Journal of National Security Law*, vol. 14, no. 2, pp. 45–58, 2023.

[18] National Institute of Standards and Technology (NIST), “AI Risk Management Framework,” 2023.
[NIST AI Risk Management Framework](#)

[19] A. Turing, “Computing Machinery and Intelligence,” *Mind*, vol. 59, no. 236, pp. 433–460, 1950.

[20] R. Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems*, 3rd ed. Indianapolis, IN, USA: Wiley, 2020.

[21] J. Kaplan, *Artificial Intelligence: What Everyone Needs to Know*. Oxford, U.K.: Oxford University Press, 2016.

[22] NATO, “Artificial Intelligence Strategy,” 2021.
[NATO AI Strategy](#)

[23] MIT Sloan, “The Risks of AI Surveillance,” 2023.
MIT Sloan AI Surveillance Risks

[24] World Economic Forum, “Global Risks Report: AI and Cybersecurity,” 2024.
World Economic Forum AI Risks

[25] B. Goodman and S. Flaxman, “European Union Regulations on Algorithmic Decision-Making and a Right to Explanation,” *AI Magazine*, vol. 38, no. 3, pp. 50–57, 2017.