Walisongo Law Review (Walrev), Vol 7 No. 2 (2025)
DOI: 10.21580/walrev.2024.6.2.27851
P-ISSN: 2715-3347 E-ISSN: 2722-0400

ARTIFICIAL INTELLIGENCE IN CYBERSECURITY LEGAL AND ETHICAL CHALLENGES IN REGULATING AUTONOMOUS DEFENSE SYSTEMS

Grahadi Purna Putra
Universitas Khairun, Indonesia

*Correspondence: *grahadipurna@gmail.com*

Citation

Purna Putra, Grahadi. 2025. "Artificial Intelligence in Cybersecurity Legal and Ethical Challenges in Regulating Autonomous Defense Systems". *Walisongo Law Review (Walrev)* 7 (2):179-94. https://doi.org/10.21580/walrev.2025.7.2.27851.

Abstract: The emergence of Artificial Intelligence (AI) in autonomous cybersecurity defense systems has created a profound regulatory and ethical. Current doctrines of responsibility, which center on the role of humans, are insufficient to regulate AI systems that act independently, particularly in the context of cross-border cyber incidents. This study uses a normative and comparative legal approach, examining regulatory frameworks such as the EU AI Act and Russian cyber law to assess the consequences of legal fragmentation and weak international harmonization. It shows that outdated laws do not adequately cover all aspects of fault, transparency, and fairness in AI-supported online security. The paper suggests granting legal status to computer programs and establishing ways for people to engage in the process, in addition to examining how these algorithms affect us, as a means of ensuring that they are accountable. This two-pronged approach integrates fairness, transparency, and proportionality into AI governance, while strengthening ethical oversight. Further study recommends a proportional division of responsibility among developers, distributors, and regulators, as well as international harmonization to prevent regulatory arbitrage. By integrating conceptual and practical insights, this research advances anticipatory and ethics-based governance to ensure that AI-based cybersecurity systems operate effectively while upholding human rights and global security.

Kemunculan kecerdasan buatan (AI) dalam sistem pertahanan keamanan siber otonom menciptakan kekosongan regulasi dan etika. Doktrin tanggung jawab yang berpusat pada peran manusia, tidak memadai untuk mengatur sistem AI yang bertindak secara independen, terutama dalam konteks insiden siber lintas batas. Studi ini menggunakan pendekatan hukum normatif dan komparatif, mengkaji kerangka regulasi seperti undang-undang AI Uni Eropa dan hukum siber Rusia untuk menilai konsekuensi fragmentasi hukum dan harmonisasi internasional yang lemah. Studi ini menunjukkan bahwa undang-undang yang sudah ketinggalan zaman tidak mencakup semua aspek kesalahan, transparansi, dan keadilan dalam keamanan daring yang didukung AI secara memadai. Penelitian ini menyarankan

pemberian status hukum kepada program komputer dan menetapkan cara bagi orang-orang untuk terlibat dalam prosesnya, selain mengkaji bagaimana algoritma ini memengaruhi kita, sebagai cara untuk memastikan akuntabilitasnya. Pendekatan bercabang dua ini mengintegrasikan keadilan, transparansi, dan proporsional ke dalam tata kelola AI, sekaligus memperkuat pengawasan etika. Studi lebih lanjut merekomendasikan pembagian tanggung jawab yang proporsional antara pengembang, distributor, dan regulator, serta harmonisasi internasional untuk mencegah arbitrase regulasi. Dengan mengintegrasikan wawasan konseptual dan praktis, penelitian ini memajukan tata kelola antisipatif dan berbasis etika untuk memastikan bahwa sistem keamanan siber berbasis AI beroperasi secara efektif sambil menjunjung tinggi hak asasi manusia dan keamanan global.

Keywords: AI governance; legal liability; cybersecurity ethics.

INTRODUCTION

Artificial intelligence (AI) driven autonomous defense systems (ADS) are changing the face of cybersecurity. Designed to identify, analyze, and eliminate threats without human involvement, these systems offer improvements in operational efficiency and response times. However, the automation raises immediate operational, legal, and ethical concerns. Autonomous systems eroding the decision-making roles of human actors poses the question of who is accountable. The traditional legal frameworks which would assign fault and responsibility are unlikely to apply. The "black box" problem in AI also raises the question of how to make accountability calculable. The 2021 Colonial Pipeline ransomware attack illustrates these challenges. Commanded by the intrusive AI, the systems automation cybersecurity measures initiated a full pipeline shutdown, a response which balanced the economic costs of statewide fuel shortages with the operational costs of extending the cyber-attack. The economic losses are estimated to be in the billions of dollars. (Perlroth, 2021) This example illustrates the ambiguity of AI-based cybersecurity: how swiftly threats can be neutralized, and how quickly "thinking machines" can be left to operate on their own and cause immense collateral damage. It creates the accountability conundrum: who is liable when autonomous defense systems decide and act systems that impact national security and the economy the most - the system, the developer, the consumer, or the AI system? This profound misalignment between the pace of innovation and law creates immediate gaps that must be filled. This is why legal digital defense scholars and practitioners must go back to the basics and rethink the core principles of law. (Tarun Kumar Vashishth, 2025)

Focusing on ethics, autonomous systems in cybersecurity not only diminish human supervision, but also pose questions regarding equitable decision-making and affect individual rights to privacy. When there is no human verification, AI actions, however autonomous,

raises questions regarding their moral ethical validity. Such systems can respond to cyber threats far quicker than any human, but this quick decision-making inherently lacks explainability and the qualitative assessment that human reasoning provides. Biased algorithms, particularly those informed by biased datasets, can promote unjust decision-making. (Kulothungan, 2024) In terms of defense, these biases can lead to the unmerited targeting of people or systems. Concern arises from automated digital countermeasures lacking proportional AI moral reasoning. (Satory, 2024) These technological dilemmas require not only business philosophy adjustments, but also regulatory changes. (Pasupuleti, 2024)

The present legal systems do not fully appreciate the complexities artificial intelligence brings to self-governing cybersecurity. Many current laws do not adequately cover the legal implications of self-operating algorithms. (Mingo, 2024) For instance, the law is silent on who 'owns' the decisions made by an AI, or the questions of responsibility arising from an AI's actions. (Li, 2024) There are also potential liabilities concerning the copying of other technologies by AI models during the process of threat mitigation. (Mustafa, 2024) Such an intricate legal scenario supplements the argument for the modernized regulatory focus, which considers the legal personhood of self-governing systems and their operators, (Sundar Tiwari, 2020) Worldwide implications of cybersecurity also pose challenges to legal certainty concerning extraterritorial jurisdiction and global legal alignment.

The interaction of cross-border data flows, global threat actors, and competing jurisdictions brings legal challenges to the principles of sovereignty and territoriality. (Yakubova, 2024) While attempts like the EU's GDPR and AI Act propose solutions to these issues, their effects are minimal for jurisdictions outside the EU. (Pasunuru, 2025) The lack of consistency between countries' laws makes it even harder to respond to coordinated global cyber challenges, especially when multiple AI systems operate within different jurisdictions. (Rawol, 2024) Thus, the need for collaboration at the international level and the need harmonize frameworks for the governance of AI become imperative to avoid llegal fragmentation and avoid gaps in enforcement. (Gutsalyuk, 2024)

One of the primary ethical concerns regarding AI in the domain of cybersecurity is the lack of transparency and the interpretability of algorithms. It impedes the legal assessment and undermines the trust of the public. (Bhatti, 2023) If a cybersecurity system is incapable of explaining its decision, it cannot be legally accountable to its doctrines. The advancement of Explainable AI (XAI) is, therefore, necessary to resolve this issue. (Stoilova, 2024) Furthermore, auditing for compliance of laws such as data protection, civil liberties, anti-discrimination and discrimination is only possible with transparency. Otherwise, ADS may violate fundamental rights cloaked in the pretext of cybersecurity. (Kulothungan, Securing the AI Frontier, 2024) Thus, explainability is not merely a technical requirement, it is also a legal requirement and ethical requirement.

Recent academic work emphasizes the opportunities and risks associated with using AI in defense-focused cybersecurity. AI performs anomaly detection, fraud prevention, and

intrusion response, among other tasks. (Vashishth, 2024) Yet AI creates new risks such as the potential for adversarial attacks and other forms of manipulation. (Pasupuleti, Legal and Regulatory Frameworks for AI in Cybersecurity, 2024) There are warnings from several scholars regarding the AI arms race, in which bad actors apply AI to develop more advanced and nuanced attacks. (Li, 2024) The inherent dual-use nature of AI technologies calls for governance to differentiate legitimate defensive uses from abusive exploitation. (Mustafa, 2024) Therefore, all ethical safeguards should be integrated within each phase of the AI system life cycle, from design and deployment to post-implementation monitoring. (Satory, 2024)

When there is malfunctioning of AI systems and there are unintentional repercussions, the possible encroachment of the rights of an individual, and the malfunctioning AI systems are employed, the ethical problems that arise are tangled with legal responsibility. (Mingo, 2024) Can automated systems in cyberspace claim self-defense? How should the law understand algorithmic actions when it comes to the concept of intent? (Rawol, 2024) These questions indicate the gaps in the current tort and criminal liability frameworks ignoring the context of AI. Moreover, reputational and compliance risks muster for the organizations that implement ADS and do not observe the requisite due diligence. (Kulothungan, Securing the AI Frontier: Urgent Ethical and Regulatory Imperatives for AI-Driven Cybersecurity, 2024) To alleviate these risks, some legal scholars suggest the advancement of legal personhood constructed on "algorithmic agents" or "digital trustees" which would assume a portion of the liability. (Gutsalyuk, 2024)

This article identifies and analyzes the legal and ethical issues within the regulation of autonomous AI-based defense systems. Through a doctrinal approach, comparative law, and literature review, this research identifies how the legal systems are responding or, indeed, not responding to threats posed by AI. This research identifies the need for regulation, drawing on the principles of transparency, accountability, and equity to propose governance frameworks for AI in cybersecurity. This research attempts to balance the need for governance frameworks that empower innovation and the need for the protection of rights. The practical outcomes of this research are intended for lawmakers, technologists, and ethicists aimed at building a safe and equitable digital future. Accordingly, this research aims to answer the central question of how legal gaps and fragmentation in regulation shape the governance of autonomous cyber defense systems, and the issues of algorithm accountability and the ethics of oversight in AI cyber defense systems.

RESEARCH METHOD

This research adopts a normative juridical legal method and statutory, conceptual, and comparative approaches to measure the sufficiency of existing laws for the legal regulation of AI and cybersecurity, particularly self-defense systems. Identifying and evaluating legal norms and the ethical dimensions of algorithmic accountability, human rights, and outer space cybersecurity defense is critical for the normative approach. (Imelda Mardayanti, 2024)

Legal materials consist of legislation, legal writings, and judiciary decisions on autonomous technologies in the public and military domains. (Niţă, 2024)

In addition to the doctrinal approach, the current research incorporates a comparative legal viewpoint by studying the AI regulation frameworks in the European Union, Russia, the United States, and Uzbekistan. (Lipinsky, 2020) The comparative method serves the purpose of identifying legal gaps and areas of harmonization in systems with disparate legal traditions. (Eresko, 2024) It also demonstrates the cross-border nature of cyber threats and the need for integrated governance of cybersecurity. (Zubaedah, 2024)

For this research, primary sources such as national and international legislation, soft law, and pertinent court and administrative decisions, were supplemented with secondary sources which include peer-reviewed articles, reports, and legal commentary as well as digital governance reports available through SCOPUS, DOAJ, SINTA, and Web of Science. As a consequence of the country case studies, the works were taken from legal databases, academic repositories, and government repositories to support. This was in keeping with best practices for research reliability and the respective sources research integrity. (Matyuk, 2022)

There were three sets of data sources: (1) Texts of regulations such as the GDPR, the AI Act, and national laws pertaining to cybersecurity; (2) Doctrinal writings of legal scholars pertaining to AI liability as well as data ethics, and (3) Comparative studies of how various states manage the governance of AI. (Zhaltyrbayeva, 2023) This type of triangulation enables the researcher to address both the normative and the empirical aspects of the legal regulation of AI. (Zhaltyrbayeva, 2023)

The data were analyzed through the qualitative doctrinal approach, using the grammatical, systematic, and teleological methods. (Niţă, 2024) The researcher also carried out a conceptual synthesis of the keywords “autonomy,” “explainability,” “accountability,” and “cyber sovereignty.” (Dmitry A. Lipinsky, 2020) A critical comparison of these concepts across legal jurisdictions sought to determine whether they were functionally integrated and defined in law or were marked by contextual divergence. (Imelda Mardayanti, 2024)

Thematic selection helped focus data reduction. Documents were screened using three main criteria: (1) relevance pertaining to legal or ethical regulation of AI in defense cybersecurity; (2) credibility of publication; and (3) jurisdictional relevance. (Zubaedah, 2024) To minimize analytical imprecision, sources that were vague, or relied on outdated legal or ethical frameworks were eliminated. (Eresko, 2024) This established methodological framework provides a rigorous legal study on the extent to which the complex demands of autonomous AI systems in cybersecurity align with (or misalign) in practice and in theory with current regulatory instruments. This has proposed insights on the need for legal reform in practice and theory, focusing on the legal position, ethical, and technological relevance. (Yakubova, The Legal Challenges of Regulating AI in Cybersecurity: A Comparative Analysis of Uzbekistan and Global Approaches, 2024)

RESULT AND DISCUSSION

Legal Gaps and Regulatory Fragmentation in Governing Autonomous Cyber Defense Systems

In most places, laws meant to govern artificial intelligence (AI)-based autonomous cyber defense technology are lagging behind the technology. Many areas are largely regulated by the outdated human-centered laws and regulations, and this creates a huge legal vacuum. Legal accountability for systems with autonomous AI decision-making remains uncertain. For example, the GDPR and the EU AI Act focus on data privacy and AI legislation, respectively, but there are no agreed norms on legal accountability for transnational cyber incidents that involve AI. (Kacimi, 2022) Furthermore, there is little to no legislation that distinguishes between traditional automation and current AI systems that can learn, adjust, and modify their activity real-time. (Wong, 2020) The fundamental challenges for the private sector and states in these instances involve unregulated denial of responsibility and unintended consequences of uncooperative international arrangements.

Variations in country and sector regulations governing AI lead to imbalances in oversight and enforcement. While the EU is pioneering frameworks such as the Digital Services Act and the AI Act, there is minimal sectorial AI cybersecurity legislation in Asean and Central Asian countries. (Yakubova, The Legal Challenges of Regulating AI in Cybersecurity: A Comparative Analysis of Uzbekistan and Global Approaches, 2024) In Uzbekistan and similar countries, legislation on AI cybersecurity is vague, abstract, and generalized and does not respond to the specific demands of algorithmic systems. (Mustafa, 2024) The imbalances pose challenges to international legal collaboration and compromise joint cybersecurity initiatives, leading in practice to AI systems operating on the periphery of agreed ethical and legal principles, which diminishes trust and interoperability globally. The imbalances may lead to legal gaps and regulatory arbitrage, where entities gamble with jurisdictions that permit the clandestine implementation of dangerous AI systems.

Another critical concern pertains to the lack of proper attribution and accountability frameworks around the autonomous systems employed in cyber defense. Unlike systems where a human directly oversees every action, the traditional legal frameworks of intent, negligence, and command responsibility do not easily fit. (Vashishth, 2024) For example, an AI firewall misidentifies and neutralizes traffic from a foreign server. Under current international law, there is no clear answer regarding whether the liability lies with the developer, the deployer, or the algorithm itself. (Pasupuleti, Legal and Regulatory Frameworks for AI in Cybersecurity, 2024) This is problematic in terms of restitution and dispute and diplomatic negotiations. This is especially pronounced in the case of multiple countries. Without an updated legal framework, cross-border cyber-AI incidents will continue to escalate, with no clear legal frameworks, adjudication, or enforceable normative limits.

Jurisdictional disputes deepen the issue of regulatory fragmentation. The minimalist architecture of cyber defense systems may lead to distributed servers, transnational networks, and decentralized AI infrastructures. This may cause overlapping claims of jurisdiction and

accountability. (Romanova, 2024) Take the case of a Singapore-headquartered system that is hosted on a cloud platform located in Europe and, at the same time, deflects threats from a hostile actor in Russia. Which court will have jurisdiction over disputes that may arise? (Satory, 2024) The Budapest Convention and other existing treaties provide little in the way of guidance on AI-specific enforcement. The problem of dual-use technologies also deepens the complexity of the interplay among military, civilian, and commercial applications. It makes the issue of inapplicability of international humanitarian law and domestic criminal law that much more pronounced.

Regulatory enforcement challenges also stem from the lack of specific legislation on AI technologies. National cyber laws continue to categorize "software agents" or "digital systems" without recognizing the differences between basic software programs versus dynamic AI technologies that evolve relative to the context. (Pasunuru, 2025) As a result, enforcement authorities struggle to prosecute cases involving the misuse of AI. Additionally, the attribution of intent requirements for legal action, whether against a person or a corporation, creates a legal enforcement deadlock when AI systems are the "actors." (Pasunuru, 2025) In both the U.S. and Russia, cyber laws are also missing essential "AI" clauses, illustrating the compartments of cyber legislation and the lack of legal definitions for "autonomous systems" and "algorithmic accountability." AI systems, particularly those used in cyber defenses, that lack regulatory oversight may misclassify threats on the basis of race, language, or geopolitical region (Pasunuru, 2025) Such systems institutionalize harmful biases, particularly when laws lack provisions on the "transparency and fairness" that must be required in audits of biased systems. (Wong, 2020) Current legal instruments do not provide the algorithmic explainability required to detect and mitigate legal and political risk. Therefore, biased enforcement and discrimination at the global level, violation of international treaties, and negative international relations may result from AI decisions that lack legal oversight.

Compliance with domestic legal requirements is further complicated by transnational deployments of AI. AI tools deployed by some multinational cybersecurity suppliers operate across borders in multiple countries, which raises questions about the laws of which country are relevant. (Yakubova, The Legal Challenges of Regulating AI in Cybersecurity: A Comparative Analysis of Uzbekistan and Global Approaches, 2024) Several complicated legal issues remain. These include the position of the AI on autonomous monitoring systems, the AI logic, and the AI evolution in the cloud. (Kacimi, 2022) Legal inconsistencies are the result of cross-border data flows, disparate standards on data privacy, and localized security expectations. Legal inconsistencies contribute to a loss of cybersecurity trust and fragmented compliance. These observations highlight the slow and uncoordinated attempts to bring together AI governance in the world and the absence of an aligned policy approach to AI governance. This includes the minimal number of enforceable internationally binding agreements available, like the ethical AI principles published by the OECD and the Council of Europe, coupled with the absence of cohesive regulatory frameworks. (Pasupuleti, Legal and Regulatory Frameworks for AI in Cybersecurity, 2024) Uncoordinated AI policies

contribute to duplication, indecision, and strategic divergence between countries. This is a growing concern in the context of cybersecurity, which faces rapidly evolving threats, whereby the regulatory frameworks are, at best, stagnant and underdeveloped, leaving significant legal gaps in the most advanced systems.

To help close these gaps, legal scholars and policymakers propose the establishment of AI-targeted global minimum standards, regulatory bodies, and frameworks for the sharing of liability. Suggested innovations include mandates for algorithmic traceability, sovereign AI audit mechanisms, and the approximation of tri-partite jurisdiction over AI under international cyber law. (Romanova, 2024) Such innovations will require political will and technical, as well as legal, adaptiveness. (Kulothungan, Securing the AI Frontier: Urgent Ethical and Regulatory Imperatives for AI-Driven Cybersecurity, 2024) To close, the governance of AI in autonomous cyber defense calls for the reconfiguration of the still largely reactive and fragmented law to one that is anticipatory, harmonized, and, most importantly, designed to ethical standards.

Cross-jurisdictional differences in understanding and regulating AI-enabled cybersecurity are stark. The AI Act of European Union employs risk-based regulation and identifies high-risk AI applications including autonomous cyber defense systems and imposes stringent requirements as including transparency, human oversight, and accountability. In comparison, Russian cyber law, irrespective of recognizing AI as a national priority, lacks provisions of any kind specifically aimed at the cybersecurity sector and remains vague and generalized. (European Commission, 2021) There are two lessons to be learnt from this kind of divergence. First, the trust of users and legal certainty are enhanced by the comprehensive frameworks of regulation, as is the case with the European Union. Secondly, generalized frameworks may be flexible, but may provide no points of accountability and weak enforcement. From a governance perspective, the lack of comprehensive legal frameworks in multiple jurisdictions highlighted the pressing need for the unification of global standards to combat regulatory arbitrage and enhance joint cyber-resilience. (Eresko, Legal Support of the Security of the Information Space of the Russian Federation in the Field of Artificial Intelligence, 2024)

The Oversight of Algorithms and Ethics in AI-Driven Cybersecurity Operations

The use of Artificial Intelligence (AI) in the field of cybersecurity has resulted in phenomenal progress, and at the same time, poses questions concerning ethics, particularly in the areas of accountability, fairness, and transparency. While AI-enabled cybersecurity systems can learn and adapt without the need for rules, traditional systems, and adapt without the need for rules, traditional systems and adapt without the need for rules, traditional systems do not learn, and adjust, and erase accountability under the "black box" of AI. (Chellappan, 2024) The growing delegation of AI systems to make independent decisions on detecting and mitigating threats, questions of accountability and ethics go beyond what law and technology can presently offer. The AI field urgently needs accountability and transparency initiatives like Explainable AI (XAI) and other models such as SHAP and LIME

that aim to provide AI systems with the necessary interpretability to provide transparency in core decisions, remaining the key stakeholders in AI systems. (Lakshmi, 2024)

Legally and ethically, designers, deployers, and setters accountable rules need to willingly commit to beyond just a technical fix for algorithmic accountability. In AI-driven cybersecurity, where an automated decision can lock the system, limit access to the data, or report a customer for suspicious activities, accountability gaps can paralyze operations or violate human rights. The distributed accountability problem here is whether the system's algorithm developer, the cybersecurity vendor, or the organization implementing the system is liable to the accountability gaps. (Vashishth, 2024) The absence of a clear accountability paradigm to counter this ambiguity is the patchy AI governance across jurisdictions. Active and institutionalized algorithmic impact assessments and human-in-the-loop mechanisms are recommended by scholars to meet the ethical standards required in the cyber defense. (Pappachan, 2024)

In cybersecurity AI, one of the most alarming ethical challenges is the issue of biased algorithms, particularly when learning systems identify, flag, or ignore particular user behaviors. The bias embedded in threat scoring systems may lead to racial, gender, or behavioral profiling, especially in systems that work with surveillance and biometric-checked systems. (Vish Iyer, 2025) The primary goal of Explainable AI (XAI) is to remove bias by providing transparency on feature influence on the outcome of the model, but technical complexity and performance-interpretability trade-offs are challenges that remain. Auditing and bias fairness checking must be built in the AI systems development lifecycle maintenance to be able to shift the bias effectively. (Goel, 2024)

The lack of transparency in AI, particularly in deep learning, becomes a direct threat to governance in cybersecurity. When decisions made by AI systems cannot be readily explained or undone, they create "black box" scenarios which hinder audits, and compliance checks, and in turn, prevent organizations from assessing the legally and ethically defensible boundaries of their AI systems. (Muthusubramanian, 2024) S The absence of adequate explainability in high-risk cybersecurity challenges illustrates the necessity of scholars treating XAI as a means to satisfy legal procedural requirements as well as fulfill the demand for technological transparency. As per the regulations and guidelines, the explainability frameworks ought to establish minimum requirements that are commensurate with the autonomy and consequences of the systems in question. (Lunn, 2024)

The possibility of cybersecurity AI systems causing harm and danger without escalatory intent requires ethical oversight. AI systems might by themselves initiate indefinite countermeasures or isolate critical components of an infrastructure, and unreasonable damage might occur based on erroneous data. This makes the need for ethics of oversight proportionality, necessity, and oversight. (Sindiramutty, 2024) Cross-disciplinary ethics committees with legal scholars, cyber-security experts, ethicists, and stakeholders can provide institutional checks on AI release and use. Moreover, AI incident review boards can function

for the purpose of dissecting failures and recommending safeguards similarly to how aviation boards function post-accidents. (Ramya, 2023)

The probable use of AI for unreasonable expected gain or unforeseen exploit must also be addressed for accountability. Attackers might exploit certain weaknesses in machine learning algorithms, like data poisoning, adversarial inputs causing false negatives, false positives, and other methods. (Raj, 2024) This abuse cross the line of unreasonable use and requires the necessary technical counter-measures to be accompanied by legal accountability. Legal doctrines such as strict liability, negligence, and duty of care must be adjusted to these novel threats. Decision logs, response protocols, and model versioning provide the audit trail and are also important for these legal doctrines to be carried. (Gyandeep, 2024)

The extent to which the public trusts a cybersecurity system depends on the ethics of how the system is managed and on how these issues are communicated. Stakeholders and users expect not only technical proficiency but also legal and social congruity. Organizations need to embrace positive ethics through open documentation, algorithmic transparency, and participatory design methods that engage end users. (Engelmann, 2023) Furthermore, balanced multi-stakeholder governance including public policy regarding AI technology development will alleviate the unequal distribution of power and facilitate the system serving the common good, rather than the private interests of the developers. (Hagen-Zanker, 2023) Algorithmic accountability and the ethics of cybersecurity, in conclusion, depend on the intersection of law, technology, and administrative order. (Elgar, 2024) The potential provided by explainable AI tools will only be realized if organizations adopt a human rights-focused framework and active oversight to explain these tools. (Lunn, 2024)

Arguments around the intersection of AI and cybersecurity may delve into thorny issues of established theories and the law. In liability theory, one of the main theories of international law, the 'owned' agency proposed by autonomous systems raises the question of the legal implications of human agency and non-human agency. (Plakokefalos, 2014) Another theory is agency theory: when AI acts as a 'digital agent', is liability 'sure' to remain with the principal (the deploying entity) and/or partially with the agent? (Meckling, 1976) In addition to the above, one may consider the moral AI and technology ethics frameworks that will require the designers and deployers of AI to balance the principles of proportionality, fairness, and explainability, and to provide 'embedded' moral autonomy. (Taddeo, 2016) All of these theories materially explain the 'negative' law gap and assist in the mental construction of new liability theories in relation to accountable AI-powered defense systems. (Burri, 2017)

For example, imagine the scenario when the AI-powered intrusion detection system (IDS) incorrectly identifies a legal data transfer as a hacking attack and auto-denies network access. (Chaudhary, 2024) In such cases when the auto-shutdown affects the critical hospital database server, the consequences could lead to loss of life due to delayed or obstructed medical services. (Vish Iyer, A Value-Based Approach to AI Ethics: Accountability, Transparency, Explainability, and Usability, 2025) In this hypothetical example, several legal and ethical issues can arise: Do the hospital authorities using the IDS tool, the AI tool

developers, or the regulatory authorities share the responsibility for the loss of human life due to the auto-shutdown feature of the IDS tool? This hypothetical scenario illustrates the hazards of real AI implementation when the "black box" effect affects the AI system disproportionately. This creates a pressing imperative for enforceable AI cybersecurity regulations for both explainability as well as accountability. (Goel, Ethical Considerations in Implementing Artificial Intelligence in Cybersecurity," in Advances in Information Security, Privacy, and Ethics, 2024)

CONCLUSION

In fact, this project finds that the existing law has a very important gap that cannot currently be covered because the law has failed to regulate AI-based self-governing defense systems. The project has importance because it introduces the solution of the very limited recognition of algorithmic agents and human-in-the-loop techniques/algorism impact assessment approaches as the innovative accountabilities. This will not only protect the previous studies because the previous studies were only concerned with the technical solutions but failed to consider the liability resolution. This provides important implications because first, the liability concerned with the AI-based cybersecurity should be allocated fairly among the producers and the regulators. This indicates that the current lack of international regulations has serious implications because states face the risk of arbitrage as the regulations are not standardized at the international level. Thus, the existing lack of regulations at the international level has made states vulnerable because the current lack of standardization has led to the risk of lack of cybersecurity at the international level. [*W*]

BIBLIOGRAPHY

Akhtar, Zarif Bin, and Ahmed Tajbiul Rawol. "Harnessing Artificial Intelligence (AI) for Cybersecurity." Computing and Artificial Intelligence 2, no. 2 (2024). https://doi.org/10.59400/cai.v2i2.1485

Andre Nollkaemper and Ilias Plakokefalos, *Principles of Shared Responsibility in International Law* (Cambridge: Cambridge University Press, 2014), https://doi.org/10.1017/CBO9781139924528.

Ashraf, Z., and Nadeem Mustafa. "AI and Cyber Laws." Dalam Advances in Healthcare Information Systems and Administration, disunting oleh M. Usman dan G. M. Bhatti, 250–268. Hershey, PA: IGI Global, 2024. https://doi.org/10.4018/979-8-3693-7051-3.ch015.

Bhatti, David Samuel, et al. "AI's Challenge to Ethics and Law." Dalam 2023 International Conference on Information Management & Intelligent Computing (ICIMIC), 2023. https://doi.org/10.1109/INMIC60434.2023.10466146.

Chaudhary, Gyandeep. "Unveiling the Black Box: Bringing Algorithmic Transparency to AI." Masaryk University Journal of Law and Technology 18, no. 1 (2024): 57–74. https://doi.org/10.5817/MUJLT2024-1-4.

Chellappan, Ramesh Babu. "From Algorithms to Accountability: The Societal and Ethical Need for Explainable AI." Preprint, Research Square, 2024. https://doi.org/10.21203/rs.3.rs-5277731/v1.

Duca, Daniela, and Alex Hagen-Zanker. "Building Ethical AI Systems in the Public Sector: Governance, Transparency and Trust." AI and Ethics 3 (2023): 415–430. https://doi.org/10.1007/s43681-022-00241-w.

Engelmann, Alana. "Algorithmic Transparency as a Fundamental Right in the Democratic Rule of Law." Brazilian Journal of Law, Technology and Innovation 1, no. 2 (2023): 169–188. https://doi.org/10.59224/bjlti.v1i2.169-188.

Eresko, P. V. "Legal Support of the Security of the Information Space of the Russian Federation in the Field of Artificial Intelligence." Vestnik Universiteta imeni O. E. Kutafina 122, no. 10 (2024): 69–76. https://doi.org/10.17803/2311-5998.2024.122.10.069-076.

European Commission. *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*. COM (2021) 206 final, Brussels, 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206.

Goel, Pawan Kumar. "Ethical Considerations in Implementing Artificial Intelligence in Cybersecurity." Dalam Advances in Information Security, Privacy, and Ethics, 77–95. Hershey, PA: IGI Global, 2024. https://doi.org/10.4018/979-8-3693-6517-5.ch005.

Gutsalyuk, M. AI in Cybersecurity: An Instrument for Defense and Attack. 2024. https://doi.org/10.69635/978-1-0690482-1-9-ch23.

Gyandeep Chaudhary, *"Unveiling the Black Box: Bringing Algorithmic Transparency to AI," Masaryk University Journal of Law and Technology* 18, no. 1 (2024): 57–74, https://doi.org/10.5817/MUJLT2024-1-4.

Hooper, Kerrie, and Stephanie J. Lunn. "A Scoping Review of Transparency and Explainability in AI Ethics Guidelines." Dalam *Proceedings of the 37th International Florida Artificial Intelligence Research Society Conference (FLAIRS-37)*, disunting oleh Kerrie Hooper dan Stephanie J. Lunn, 46–52. Miami, FL: Florida International University, 2024. https://doi.org/10.32473/flairs.37.1.135594.

Ilieva, Roumiana, and Gloria Stoilova. "Challenges of AI-Driven Cybersecurity." Dalam 2024 57th International Scientific Conference on Information, Communication and

Energy Systems and Technologies (ICEST), 2024. https://doi.org/10.1109/ET63133.2024.10721572.

Iyer, Vish, Muhanad S. Manshad, and Daniel Brannon. "A Value-Based Approach to AI Ethics: Accountability, Transparency, Explainability, and Usability." Mercados y Negocios 54 (2025): 123–140. https://doi.org/10.32870/myn.vi54.7815.

Kacimi, Sahra. "Legal Vacuum to Regulate AI." Dalam Artificial Intelligence and International Economic Law, disunting oleh Shin-yi Peng, Thomas Streinz, and Ching-Fu Lin, 123–142. Singapore: Springer, 2022. https://doi.org/10.1007/978-981-19-1496-6_7.

Keeling, Amanda, and Edward Elgar. "Regulating AI Through Ethical Governance: The Role of International Coordination." Dalam AI Regulation in the Global South, disunting oleh L. Banda dan C. Maina, 144–160. Cheltenham: Edward Elgar, 2024. https://doi.org/10.4337/9781035318965.00018.

Kulothungan, Vikram. "Securing the AI Frontier: Urgent Ethical and Regulatory Imperatives for AI-Driven Cybersecurity." arXiv, 2025. https://doi.org/10.48550/arXiv.2501.10467.

Kulothungan, Vikram. "Securing the AI Frontier: Urgent Ethical and Regulatory Imperatives for AI-Driven Cybersecurity." Dalam Proceedings of the 2024 IEEE International Conference on Big Data, 2211–2220. New York: IEEE, 2024. https://doi.org/10.1109/BigData62323.2024.10826010.

Li, Fangshu. "Application and Challenges of Artificial Intelligence in Cybersecurity." Dalam Applied and Computational Engineering: Proceedings of the 2024 International Conference on Applied and Computational Engineering, 2024. https://doi.org/10.54254/2755-2721/47/20241480.

Lipinsky, Dmitry A., Roman A. Romashov, Aleksandra A. Musatkina, Svetlana G. Golenok, and Elena A. Bryleva. "The Problems of Legal Regulation of AI: A Rather-Legal Research." Dalam Digital Transformation of the Economy: Challenges, Trends and New Opportunities, disunting oleh Irina S. Ibragimova, 470–481. Cham: Springer, 2020. https://doi.org/10.1007/978-3-030-39319-9_47.

Luciano Floridi and Mariarosaria Taddeo, "What is Data Ethics?," *Philosophy & Technology* 29, no. 4 (2016): 327–329, https://doi.org/10.1007/s13347-016-0218-7.

Mardayanti, Imelda, Yenni Arfah, and Dedy Dwi Arseto. "Pemanfaatan Artificial Intelligence (AI) dalam Pembentukan Peraturan Perundang-undangan serta Implikasinya terhadap Etika dan Keamanan." Cakrawala Sosial Politik 3, no. 1 (2024): 11–22. https://doi.org/10.70021/csp.v3i1.136.

Matyuk, Yulia S. "Legal Regulation of Artificial Intelligence: Foreign Practices." Russian Journal of Legal Studies 9, no. 1 (2022): 1–16. https://doi.org/10.17816/RJLS91009.

Michael C. Jensen and William H. Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics* 3, no. 4 (1976): 305–360, https://doi.org/10.1016/0304-405X(76)90026-X.

Mingo, Horace C. "The Emerging Cybersecurity Challenges with Artificial Intelligence." Dalam Cybersecurity Issues and Challenges in the Drone Industry, 2024. https://doi.org/10.4018/979-8-3693-3226-9.ch010.

Mohitkar, Chhavi, and D. Lakshmi. "Explainable AI for Transparent Cyber-Risk Assessment and Decision-Making." Dalam Advances in Computational Intelligence and Robotics, 99–112. Hershey, PA: IGI Global, 2024. https://doi.org/10.4018/979-8-3693-7540-2.ch010.

Muthusubramanian, Muthukrishnan, et al. "Demystifying Explainable AI: Understanding, Transparency, and Trust." International Journal for Multidisciplinary Research 6, no. 2 (2024): 1–15. https://doi.org/10.36948/IJFMR.2024.V06I02.14597.

Niţă, Gabriel. "Artificial Intelligence Regulation: Approaches and Implications." Adjuris International Academic Publisher 1 (2024): 145–158. https://doi.org/10.62768/adjuris/2024/1/10.

Pappachan, Princy, et al. "Transparency and Accountability." Dalam Advances in Computational Intelligence and Robotics, 87–104. Hershey, PA: IGI Global, 2024. https://doi.org/10.4018/979-8-3693-3860-5.ch006.

P. V. Eresko, "Legal Support of the Security of the Information Space of the Russian Federation in the Field of Artificial Intelligence," *Vestnik Universiteta imeni O. E. Kutafina* 122, no. 10 (2024): 69–76. https://doi.org/10.17803/2311-5998.2024.122.10.069-076.

Pasunuru, Sreekanth. "Cybersecurity in Autonomous Systems: Protecting AI-Driven Applications." Indian Scientific Journal of Research in Engineering and Management 6, no. 1 (2025): 33–47. https://doi.org/10.55041/IJSREM20376.

Pasupuleti, Murali Krishna. Legal and Regulatory Frameworks for AI in Cybersecurity. Singapore: Springer, 2024. https://doi.org/10.62311/nesx/97890.

Pawan Kumar Goel, "*Ethical Considerations in Implementing Artificial Intelligence in Cybersecurity,*" in *Advances in Information Security, Privacy, and Ethics*, ed. M. Usman (Hershey, PA: IGI Global, 2024), 77–95, https://doi.org/10.4018/979-8-3693-6517-5.ch005.

Perlroth, N., Sanger, D. E., & Krauss, C. (2021, May 13). *Cyberattack forces a shutdown of a top U.S. pipeline*. The New York Times.

https://www.nytimes.com/2021/05/08/us/politics/cyberattack-colonial-pipeline.html

Ramya, P., et al. "Advancing Cybersecurity with Explainable Artificial Intelligence: A Review of the Latest Research." Dalam 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), 900–907. New York: IEEE, 2023. https://doi.org/10.1109/ICIRCA57980.2023.10220797.

Romanova, A. N. "Elements of Legislation for Artificial Intelligence Systems." International Journal of Soft Computing and Artificial Intelligence 13, no. 2 (2024): 55–63. https://doi.org/10.5121/IJSCAI.2024.13203.

Satory, Agus, Dimas Nugroho, and Rika Iskandar. "The Legal Challenges of Data Privacy Laws, Cybersecurity Regulations, and AI Accountability." Sahoe Gwahag Nonchong 58, no. 3 (2024): 201–215. https://doi.org/10.59613/zgvwd520.

Sharma, Nikita, and Kanchan Raj. "AI, Accountability, and Cyber Regulation in a Global Context." International Journal of Law and Technology 12, no. 1 (2024): 89–103. https://doi.org/10.55547/IJLT.V12I1.457.

Sindiramutty, Siva Raja, et al. "Ethics and Transparency in Secure Web Model Generation." Dalam Advances in Information Security, Privacy, and Ethics, 210–229. Hershey, PA: IGI Global, 2024. https://doi.org/10.4018/979-8-3693-5415-5.ch012.

Thomas Burri, "The Politics of International Law and International Politics," *German Law Journal* 18, no. 2 (2017): 457–480, https://doi.org/10.1017/S2071832200022113.

Tiwari, Sundar, Vishal Sresth, and Aakash Srivastava. "The Role of Explainable AI in Cybersecurity." International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET) 9, no. 3 (2020). https://doi.org/10.15680/IJIRSET.2020.0903165.

Vashishth, Tarun Kumar, et al. "Ethical and Legal Implications of AI in Cybersecurity." Dalam Advances in Computational Intelligence and Robotics, disunting oleh J. R. Rai dan A. Saxena, 310–329. Hershey, PA: IGI Global, 2025. https://doi.org/10.4018/979-8-3693-7540-2.ch017.

Vish Iyer, Muhanad S. Manshad, and Daniel Brannon, *"A Value-Based Approach to AI Ethics: Accountability, Transparency, Explainability, and Usability," Mercados y Negocios* 54 (2025): 123–140, https://doi.org/10.32870/myn.vi54.7815.

Wong, Anthony. "The Laws and Regulation of AI and Autonomous Systems." Dalam The Future of Digital Governance, disunting oleh M. V. Ramesh dan S. Bhat, 75–89. Cham: Springer, 2020. https://doi.org/10.1007/978-3-030-64246-4_4.

Yakubova, Madinabonu. "The Legal Challenges of Regulating AI in Cybersecurity: A Comparative Analysis of Uzbekistan and Global Approaches." International Journal of Law and Policy 8, no. 2 (2024): 25–38. https://doi.org/10.59022/ijlp.202.

Zhaltyrbayeva, Rauan, et al. "Legal Regulation in the Field of Artificial Intelligence: Assessment and Prospects." Journal of Law and Sustainable Development 11, no. 12 (2023): 149–164. https://doi.org/10.55908/sdgs.v11i12.2049.

Zubaedah, Putri Amalia, Rois Harliyanto, Sahat Maruli Tua Situmeang, Darwin Steven Siagian, and Ema Septaria. "The Legal Implications of Data Privacy Laws, Cybersecurity Regulations, and AI Ethics in a Digital Society." Jurnal Ilmu Hukum dan Teknologi 5, no. 2 (2024): 44–61. https://doi.org/10.59613/29qypw51.