sciendo

Artificial Intelligence in Security: Balancing Institutional Efficiency and the Protection of Individual Rights

Narciz BĂLĂȘOIU
Bucharest University of Economic Studies, Bucharest, Romania
narciz.balasoiu@rei.ase.ro

Abstract. *Artificial Intelligence has become the light-motif of the much-cited technological revolution, which will change society as we know it at a paradigmatic level. The Security and Law Enforcement Sector is one of the preferred fields for integrating these disruptive technologies. In this study, I will highlight the main levels of integration of artificial intelligence (AI) into the security work toolkit, whether we are talking about detection, control, response or decision-making processes. Equally, the galloping dynamics of these technologies raise a series of challenges from the perspective of the fundamental rights and freedoms of the individual. Thus, advantages such as reducing human errors or increasing predictive capabilities are countered by challenges/risks such as algorithmic bias or excessive surveillance. The levels of autonomy specific to AI-based systems are also considered, in order to increase the efficiency of the collaborative human-machine approach. Finally, the article highlights the need to build a regulatory framework that balances the efforts to innovate and optimize security systems with the imperative need to respect fundamental individual rights and freedoms. As AI becomes increasingly present in the security field, careful regulation and oversight of its use are essential to prevent abuses and guarantee responsible and transparent use.*

Keywords: Artificial Intelligence, Security, Automation, Autonomy, Emerging technologies, Data protection, AI ethics.

Introduction

It has become increasingly difficult to use new artificial intelligence (AI) technology because marketing data, technical needs, and commercial goods are all changing so quickly. In this context, it is both relevant and complicated to try to get a research-based knowledge of how AI is applied in security technology. When deciding whether or not to use AI-based security solutions, security professionals' problems with AI integration and decision-making require ongoing adaption. Because of this, it is very important to stress how AI is used in security technologies across all types, domains, and ranges of AI paradigms, as well as across the entire technical lifecycle of operation.

Professionals are trying to figure out the pros and cons of using or not employing artificial intelligence as technology moves forward quickly. As a summary of the main themes and issues in the literature on the subject, the current use of AI in security technologies, the chances of more widespread usage, the possible problems, and the list of risk factors are not thoroughly covered. In the same respect, the information here can help one think critically about artificial intelligence and how it can be used in the real world. It can also prove relevant in weighing the pros and cons of using AI to improve security technologies and make sure that these risks are noted in the organization's risk registers.

Many current technologies, such security systems and devices, use artificial intelligence algorithms to make them better. An algorithm is a set of instructions or a single instruction that tells a computer or system how to accomplish something. Security technologies can do a lot of things with AI algorithms, such as finding patterns and signals (like the sound of gunshots),

DOI: 10.2478/picbe-2025-0133

finding strange behaviour patterns (like behavioural analysis in surveillance systems), classifying and comparing images (like computer vision to reflect the difference between a person and an animal), and finding and identifying images or materials (such as contraband or compounds in X-ray scanners). Using AI in security technology can greatly improve operational security by making it easier to find things and faster to do so, reducing operator fatigue and stress, and allowing security professionals to focus on the areas that need it most.

AI could help management save costs, better allocate resources, make decisions, and even provide them the chance to intervene early to reduce insider challenges. Many security technologies use basic AI to do a specific job. AI-powered technologies that can do multiple specialized tasks at once may seem to be smarter in some scenarios. However, this doesn't always mean that the AI is smarter. Intelligence usually is alleviated when decisions are harder and more integrated, not when a system or device can do more specialized tasks.

Literature review

In the context of the accelerating evolution of artificial intelligence (AI) and its integration into the security domain, the literature reflects a complex landscape marked by conceptual dilemmas, technical challenges and ethical debates. Existing studies outline an emerging paradigm in which AI is redefining traditional methods of surveillance, threat analysis and response, fundamentally transforming institutional security mechanisms.

Theoretical approaches to AI applied to security fall into two categories: those that analyze the impact of functionality on institutional effectiveness and those that problematize the ethical and legal implications of its use. In the first case, studies highlight AI's ability to optimize anomaly detection processes (Chen et al., 2020), improve the accuracy of biometric identification systems (Nguyen & Zhao, 2021), and reduce reaction time in critical situations (Smith et al., 2019). On the other hand, the literature emphasizes associated risks including algorithmic errors (Brennan, 2022), excessive surveillance tendencies (Zuboff, 2018) and impact on the fundamental rights of the individual (Crawford, 2021).

As for identifying the main technological usage models, the literature delineates three AI paradigms in security: symbolic, statistical and sub-symbolic (Russell & Norvig, 2020). The symbolic paradigm, based on rule-based systems and formal logic, is used in expert systems and decision analysis (Laird et al., 2021). The statistical paradigm, based on machine learning techniques, dominates in predictive analytics and behavioral pattern recognition (Goodfellow et al., 2016). Finally, the sub-symbolic paradigm, based on neural networks and deep learning, is increasingly present in areas such as facial recognition and security image analysis (LeCun et al., 2015).

The literature identifies four major directions in which AI is transforming security mechanisms: observation technologies, detection systems, access control, and response mechanisms (Schneier, 2021). Empirical studies show that AI-based surveillance systems are capable of analyzing massive volumes of data in real time, detecting anomalies in the behavior of individuals (Ferguson, 2020). However, this increased efficiency raises questions about data privacy and the risk of algorithmic discrimination (O'Neil, 2016).

In terms of detection systems, AI-based technologies have proven useful in analyzing hazardous materials and cyber threats (Gorwa et al., 2022), but the literature emphasizes the need for constant calibration to avoid false alarms and misinterpretation (Koopman et al., 2019).

Another area of interest is access control, where AI has revolutionized identity verification through biometric recognition and behavioral authentication techniques (Wayman, 2021). However, research highlights potential vulnerabilities related to the security of these

systems and the risk of compromised personal data (Raji & Buolamwini, 2019). Finally, in terms of response mechanisms, AI is being used to support decision support in crisis situations by providing response scenarios based on probabilistic threat analysis (Bryson, 2020). However, ethical debates are intense, especially when using AI in autonomous lethal decision-making systems (Asaro, 2019).

Recent studies outline optimistic outlooks on the evolution of AI in security, but also highlight emerging challenges (Bostrom, 2014). On the one hand, the literature suggests that AI will continue to improve decision-making processes and significantly reduce human error (Rahwan et al., 2019). On the other hand, there are concerns about the increasing autonomy of security systems, the possibility of using AI for repressive purposes, and the difficulty of regulating this field within a coherent international framework (Calo, 2020).

One can conclude that the literature portrays an ambivalent landscape of the role of AI in security: although the technology promises increased efficiency and advanced detection capabilities, the implications for individual rights and ethical dilemmas cannot be neglected. In this context, the development of AI in security needs to be accompanied by a robust regulatory and ethical framework capable of maintaining a balance between innovation and the protection of fundamental societal values.

Methodology

The methodology used in this article is based on a qualitative approach, with the analysis focusing more on the examination of the specialized literature, the main objective being the identification of the main areas of interest in the intersection of artificial intelligence and security. In this context, we used a combination of analytical and documentary means, including the comparative analysis of the most relevant scientific research efforts in the field.

The research began with a synthesis, a review of the existing specialized literature, which included academic studies, official reports, as well as reference works by renowned authors in the fields of artificial intelligence and security. This stage allowed the outline of a well-systematized conceptual framework, highlighting the main technological paradigms and their implications for institutional efficiency in line with individual rights protection standards.

The analysis then focuses on the critical analysis of the main trends and challenges associated with the use of artificial intelligence in the field of security. We have taken a comparative approach, calibrating the technological, legal and ethical advantages and risks, based on case studies and positions developed in the specialized literature. The methodology used did not involve the extraction of empirical data or the application of quantitative techniques, given the exploratory nature of the research, as well as the high degree of classification and confidentiality of this type of information. Instead, the study focused on a deductive approach, in which theoretical concepts and available data were analyzed to draw relevant conclusions regarding the impact of artificial intelligence on security. The methodology used generated the optimal research framework for balancing academic rigor with practical aspects, thus facilitating a better understanding of how artificial intelligence impacts operational tools in the security sphere, without compromising the standards of protection of fundamental individual rights.

Results and discussions

There are six main ways to think about and solve problems in AI security tech, including vision, reasoning, cognition, planning, and communication. These cardinal paths are called macro-approaches. Furthermore, the six paradigms can be grouped into three major groups: symbolic,

statistical, and sub-symbolic (Russell & Norvig, 2020). They are logic, cognition, probabilistic methods, machine learning, embodied intelligence, and search optimization. Symbolic paradigms are the several AI methodologies that employ symbols as the basis for computation. Moreover, these symbols can be expressed in rules, which can be logic or knowledge.

An AI process like this takes in data and, based on whether it meets a set of rules, gives a specific output. Logic-based AI methods use logical statements to show what an agent knows about its environment, its goals, and its current state. Using inference or deduction with these logical truths leads to the right computational decision to reach certain goals. Logic-based techniques cover a wide range of topics, such as knowledge representation, different types of reasoning (nonmonotonic, abductive, and inductive), and computational logic (Laird, Lebiere, & Rosenbloom, 2021). There are two basic parts to knowledge-based approaches: a knowledge-based part and an inference engine that functions as a control part to come up with new information or judgments. The knowledge base has facts about the world as it is now. These facts can be shown in numerous ways, such as declarative, procedural, heuristic, structural, or meta-knowledge, which includes ontologies and large databases.

The second part, the inference engine, uses methods like rule-based, pattern-based, and case-based reasoning to come up with new knowledge or decisions. Statistical paradigms use a set of mathematical operators on their inputs to get a specific output. For instance, a typical AI in computer vision takes the pixels of an image as input and utilizes a sequence of operations based on the color and location of the pixels to group together pixels that belong to different objects. We can tell what objects are what by looking at the measurements from these groupings. Most of the time, statistical paradigms use probabilistic and machine learning methods (Goodfellow, Bengio, & Courville, 2016). Probabilistic techniques use probabilistic representations, which are graphical models that show how uncertain relationships and knowledge about the world might be. The graphical models show how the data is spread out, and you may use statistical inference methods to make decisions. Machine learning methods automatically create models from input data that can be used to make choices or predictions.

There are three main types of machine learning: unsupervised learning, supervised learning, and reinforcement learning. Some people have called semi-supervised learning a fourth type of machine learning. Unsupervised learning methods look for patterns in input data without needing to classify them. Reinforcement learning and supervised learning both need input data to be labelled. Classification methods are used in supervised learning to figure out which category something belongs to. Numerical regression creates a function that first shows how inputs and outputs are related, and then uses this information to predict how the outputs will change as the inputs change. The purpose of reinforcement learning is to reward a learning agent for giving excellent responses and punish it for giving bad replies. This will eventually help the agent learn how to solve its challenge.

Subsymbolic paradigms are like neurons in the brain since they can learn on their own thanks to the training and growth of the neural network architecture. Renown authors state that the sub-symbolic paradigm is one in which "no specific knowledge representation should be provided ex-ante," which means that knowledge representations are not given "before the event." So, the ideas of affective computing, autonomous systems, distributed artificial intelligence, environmental computing, and evolutionary algorithms all fall within the sub-symbolic paradigm. When looking at the architectural paradigms of AI from a systems-based point of view, only the distributed intelligence paradigm and some uses of evolutionary algorithms qualify. The other paradigms are mostly just implementations of AI approaches from the symbolic or statistical

paradigms. When designing intelligent behaviour for integrated and localized agents, embedded intelligence approaches take into account situatedness and embeddedness. Localization, or situatedness, is the connection between an agent and its surroundings. Embodiment, on the other hand, is the limits that an agent's body, senses, and motor systems put on it.

The early development of bio-inspired computational intelligence approaches in robotics is linked to the study of embodied intelligence. These techniques focus on morphological computation and sensory-motor coordination in robotic models (Sheridan, 2016). Formal attempts to define levels of autonomy have changed from focusing on what computers can do to how well people and machines work together to get things done. There are differences between current ideas and definitions of integration, automation, and levels of autonomy, as well as between how well humans and AI do their jobs at each level or categorization. So, the best way to characterize these ideas is through a system engineering method, which looks at how much human control or intervention there is at each stage to tell the difference between integration, automation, and autonomy: Putting components together or adding them to make a whole.

Integrated security means using a mix of security technologies, functions, and devices, or just combining diverse security services that can talk to each other to do more complicated tasks with as little automation as possible - the approach, technique, or system of running or regulating a process with as little human input as possible, using automated tools such electronic equipment.

In automation, actions are carried out using human logic and decisions that have already been made, all within a known or assumed frame of reference. There are no decisions made during operations. Technical results that include the machine making decisions while it is running in reaction to outside, unplanned occurrences, yet a human is involved in some way and has some direct control over this process. Semi-autonomous systems can do things on their own and move quickly, unlike automated systems. Autonomy is the state of being autonomous, which is being able to control oneself or having the right to do so. An autonomous system is one in which decisions are made inside the system and do not need or involve human decision making. A fully autonomous system can handle new and unexpected situations on its own, without any preprograming, scripting, or help from people. An automated system doesn't make decisions on its own; it follows a scenario, which may be very complex, in which all possible actions have already been decided. If the system runs into a problem that wasn't intended for, it stops and waits for somebody to help (for example, "phone home").

So, for an automated system, the decisions have either already been made and written down, or they need to be made by someone else. An autonomous system, on the other hand, makes its own choices. It attempts to reach the objectives on its own, without any help from people, even when things don't go as planned or when dynamics are unclear. An intelligent autonomous system uses more advanced methods to make decisions than other systems. These methods are typically similar to those that people use. In the end, the quality of the choices an autonomous system makes is what determines how smart it is.

Current uses of artificial intelligence in security technologies

Experts underline that security technologies can be divided into four main groups: observation, detection, control, and response. Different types and categories of security technologies may utilize AI in different ways. To better understand how AI is used in security technologies, it can be truly revealing to look at how AI is used in the areas afford mentioned. One could state that the main jobs of observational technologies are to find a threat that is coming, describe the threat, help plan a reaction, and aid with investigations after something has happened. Network surveillance

systems are likely the most common type of observation technology. Detection technologies are a big part of a more complete intrusion detection system (IDS) and physical protection system (PPS). They are mostly used to find out if an unauthorized action is happening or has already happened by finding the stimulus of the action and sending an alarm to a control center for evaluation. Detection technologies can comprise a variety of systems and devices for detecting intrusions, such as alarm systems, illumination systems, x-ray technology, trace detection systems, acoustic threat detection sensors and systems, RADAR and SONAR sensors and systems.

Access and egress control are two common uses of the control function in asset protection. ASIS International defines access control as any method used to limit or manage access to a person, place, facility, complex, system, or asset. Most of the time, control technologies are "access control" systems that use a lot of different sensors and credentials. Response is the endeavour to stop, limit, or lessen an event. It could also involve an evaluation that the occurrence doesn't need immediate action.

When it comes to protecting assets, responding usually means getting human security forces involved and letting people make decisions. But "response" as a security function has changed a lot since the days when people were the only ones who could respond. Now, it might incorporate a variety of technologies that help people respond. Communications, disposable barriers, vehicle barriers, and weapons systems are some of the technologies that are used to respond to security threats today. There are several parts or aspects that make up each of these technologies, and they can have varying technological outcomes and levels of control.

Benefits and opportunities for artificial intelligence in security technologies

AI will keep getting better, and there are many real benefits to using AI in security solutions. A lot of these have to do with economic gains, such as higher productivity and lower costs from better narrow AI jobs. The biggest benefit to society will be using response technology like drones and robots to get people out of danger. There are several ways that AI can be used to improve security technologies. These chances are likely to come up in areas including surveillance, detection, control, reaction, and integrated security technologies. Here are some of them:

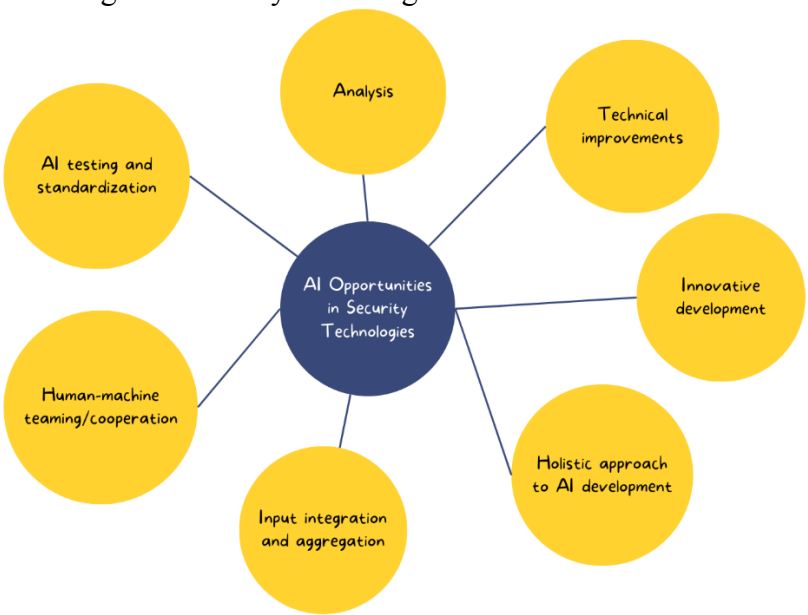

Figure 1. AI Opportunities in Security Technologies

Source: Authors' own research.

Standardization opportunities include issuing worldwide standards for AI and making common networking protocols that work on all platforms, equipment, and devices. Standardizing protocols will make it easier to connect systems and demand more and different types of inputs to enhance superior data interpretation. The creation and standardization of testing methods will help improve accuracy and dependability, giving security managers peace of mind that systems meet the levels of proficiency promised by manufacturers and suppliers in a range of operational situations. It is important to improve safety, quality, resolution, processing power, and signal noise reduction, as well as the accuracy and reliability of systems, especially in real-world situations and changing settings.

New sensing devices, the creation of source data sets, sensors and devices that can move about more easily, and better scaling and processing capabilities are all other ways to improve technology. Observation technologies, in particular, need better ways to recognize and classify objects and manage the huge amounts of data they create. More automation and greater integration of internal and external systems (such building and security systems) would help security technology. Bringing together data from several sensors and using analytics could help make better security decisions and share information across systems, facilities, and security teams to improve security response.

To do this, though, integrated systems need to change from a simple input/output architecture to a more complex one that allows for coordinated decision making across devices and at all levels of automation and administration. Better integration can also make it easier to use these technologies. For example, multi-platform control can make it easier to access and control systems or devices like drones and robots while they are still in use.

Security technologies will have a lot of chances to grow thanks to analytics and predictive analytics that help find patterns and anomalies. To help with decision making, give more context in changing situations, and give changing assessments for distributed information and response systems, new analytical methods will need to be created. Tools for making judgments, including "prioritized options" for reaction, can help security professionals make decisions that are in line with the law and have the best chance of success.

AI-generated response alternatives, along with human-machine teamwork, could make it much easier for human operators to adopt important decisions. Human-machine teaming (HMT) looks at and encourages collaborative development by bringing people and machines together to reach certain goals and abilities. HMT's goal is to help people work together, not to give computers tasks or change the way people do things. HMT can work with various kinds of security technologies, but it's especially useful with response technologies like drones and robots, where trust and safety concerns make it better for people to control systems and equipment.

So, HMT might make a big difference for security operators in integrated security management systems without putting people's safety and security at risk by making bad decisions. Management, industry, and international reforms must support technical advancement. This means that AI must be developed in a way that is holistic and coordinated across all sectors and fields that are related to safety, such as safety, facility management, engineering, research, and both academic and commercial development.

To lay the groundwork for progress in all areas of technology, we need to comprehensively investigate social ideas like trust, safety, responsibility, accountability, and integrity as qualitative drivers of AI. We also need to make rules and guidelines for how to create and use AI-based security solutions in ways that are acceptable to society. AI adoption should be

clearly linked to risks at the level of holistic management, especially when it comes to what happens if it fails.

In short, IA is easier to adopt when the consequences of failure are less severe, and adoption should be lessened as the repercussions of failure become more severe. In a practical sense, this means encouraging the use of AI in more logical situations, like showing user credentials (like an access card) or where the chances of detection are high because the materials or signatures are stable (like X-ray detection or trace detection of targeted compounds). There may also be more chances in places where many levels of defense-in-depth are working at the same time. This is because the chances of being caught by another technological or procedural measure are lower when one fails. So, the defense-in-depth strategy should be used when putting AI techniques into action, and the best security results should be achieved by carefully placing people and machines. There are many ways that the security industry can shape the social evolution of AI technology through holistic approaches to AI development. These rules and frameworks can help not just with business, law, society, and politics, but also with the security industry as an example of how AI might be used.

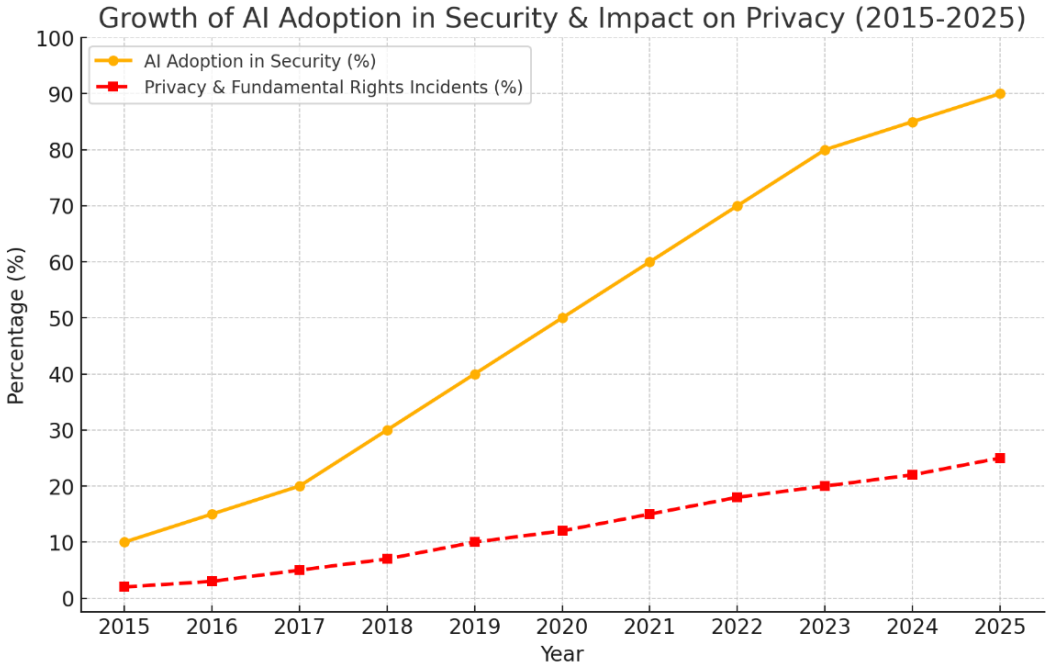

Figure 2. Growth of AI Adoption in Security & Impact on Privacy (2015-2025)
Source: McKinsey & Company's 2024 survey, AI adoption has significantly increased across various sectors.

Opportunities exist for innovative development in the realm security technologies. There are new chances to come up with new ways to use drone swarms and robotics in response technologies. There may also be chances to improve geospatial AI in response technologies. For instance, geofencing might be used to set virtual boundaries (such latitude, longitude, altitude, etc.) within which a drone or robot can move. There may be more opportunities to use smart deployment features with different types of less-lethal weapons. The creation of new types of smart autonomous weapons may also be useful in some situations or applications. Integrated security technology can make new uses possible, such as automated crowd control that uses data from various sensors to better understand the situation and find threats. Some progress has been achieved in several areas, although these technologies are still in the early stages of development.

Conclusion

There is a vast array of hazards involved with developing artificial intelligence for security systems, and we don't completely understand what those risks are yet. At the global level, the quest for technical advancement may lead to political divides, upset power balances, increase the exploitation of poor countries, and encourage the violation of people's rights and privacy. The creation of military and security response systems that can employ or release force on their own could one day be able to decide who lives and who dies or hurt or kill people.

Commercially accessible security systems can't yet reach this level of smart autonomy, but the desire for military superiority and the fact that military-to-military trade is always changing will probably make the use of force by itself a reality. Because these technologies could cause harm, the conversation on intelligent autonomy needs to include legal, moral, ethical, and human rights issues that are brought up through relevant international governance platforms. Because of this, it's important to point out how artificial intelligence is being used right now in all stages of the technical lifecycle of security solutions, as well as in all sorts, areas, and ranges of AI paradigms.

Technology is moving at full gallop pace, and so are the efforts to figure out what the future holds for AI and what the risks and opportunities are. There is not a full discussion of how AI is already used in security technologies, the chances for further use, the hazards that could come with it, and a list of risk factors. Instead, this is a summary of the main directions and concerns in the literature.

So, this material is a summary that offers a framework to think about what AI is, what it can really do, and the pros and cons of using AI-enhanced security technologies, while making sure that these risks are included in the organization's risk registers. There are a lot of challenges involved with making security technologies smarter, and we don't know what the full effects will be yet. At the global level, the desire for technical growth can lead to political divides, upset power balances, increase the exploitation of poor countries, and encourage the violation of privacy and individual rights.

Military and security response systems that may employ or release force on their own may one day have the capacity to decide who lives and dies or hurt people. Commercially accessible security technology can't yet reach this level of cognitive autonomy, but the drive for military superiority and the fact that military-to-military trade is always changing will probably make the deployment of force without human intervention a reality. Because these technologies could cause harm, the conversation about intelligent autonomy must include legal, moral, ethical, and human rights issues that are communicated through appropriate international governance platforms.

References

Asaro, P. (2019). The labor of surveillance and bureaucratized killing: New subjectivities of military drone operators. *Social Semiotics, 29*(1), 73-91. https://doi.org/10.1080/10350330.2018.1504731

Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. Oxford University Press.

Brennan, M. (2022). Artificial intelligence and algorithmic bias: Addressing fairness in security technologies. *AI & Society, 37*(3), 435-450. https://doi.org/10.1007/s00146-021-01189-2

Bryson, J. J. (2020). The artificial intelligence of ethics and the ethics of artificial intelligence. *Ethics and Information Technology, 22*(1), 1-10. https://doi.org/10.1007/s10676-019-09550-5

Calo, R. (2020). Artificial intelligence policy: A primer and roadmap. *Stanford Law Review, 72*(5), 766-810.

Chen, X., Liu, Y., & Wang, Z. (2020). Machine learning for cybersecurity: Challenges and opportunities. *Computers & Security, 94*, 101862. https://doi.org/10.1016/j.cose.2020.101862

Crawford, K. (2021). Atlas of AI: Power, politics, and the planetary costs of artificial intelligence. Yale University Press.

Ferguson, A. G. (2020). The rise of big data policing: Surveillance, race, and the future of law enforcement. NYU Press.

Gorwa, R., Binns, R., & Katzenbach, C. (2022). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society, 9*(1), 1-15. https://doi.org/10.1177/20539517221086422

Koopman, P., & Wagner, M. (2019). Autonomous vehicle safety: An interdisciplinary challenge. *IEEE Security & Privacy, 17*(3), 41-48. https://doi.org/10.1109/MSEC.2019.2893731

Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2021). A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *AI Magazine, 42*(1), 27-42. https://doi.org/10.1609/aimag.v42i1.5294

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444. https://doi.org/10.1038/nature14539

Nguyen, T., & Zhao, R. (2021). Biometric authentication systems and artificial intelligence: Opportunities and risks. *Journal of Information Security and Applications, 60*, 102888. https://doi.org/10.1016/j.jisa.2021.102888

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

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C., ... & Wellman, M. (2019). Machine behaviour. *Nature, 568*(7753), 477-486. https://doi.org/10.1038/s41586-019-1138-y

Raji, I. D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of public scrutiny on large-scale face recognition systems. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 429-435. https://doi.org/10.1145/3306618.3314244

Russell, S. J., & Norvig, P. (2020). Artificial intelligence: A modern approach. Pearson Education.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks, 61*, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003

Schneier, B. (2021). Artificial intelligence and the future of security: Risks, ethics, and governance. *Harvard Kennedy School Journal of AI & Security, 3*(1), 57-72.

Sheridan, T. B. (2016). Human-robot interaction: Status and challenges. *Human Factors, 58*(4), 525-532. https://doi.org/10.1177/0018720816644364

Smith, H., Jones, M., & Patel, A. (2019). AI-based security systems: A critical analysis. *Cybersecurity Journal, 5*(2), 89-105.

Wayman, J. (2021). Biometric systems: Design, performance, and applications. Springer.

Zuboff, S. (2018). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.

We would like to thank the European Commission and the Government of Romania for funding this research within the Project "Accountable Governance and Responsible Innovation in Artificial Intelligence", grant number 760047/23.05.2023, financed through the National Plan of Recovery and Resilience PNRR-III-C9-2022–I8.