Advantages and challenges of AI systems in security

Moftah Ismail
PhD

DOI: 10.29322/IJSRP.16.01.2026.p16917
https://dx.doi.org/10.29322/IJSRP.16.01.2026.p16917

Paper Received Date:15th December 2025
Paper Acceptance Date: 12th January 2026
Paper Publication Date: 20th January 2026

Abstract

In recent years, artificial intelligence (AI) has become a crucial component in optimizing and maintaining IT infrastructures. The implementation of predictive models based on historical data (logs, metric data) enables proactive issue detection, thereby reducing the number of unplanned downtimes and optimizing response times to technical issues. Simultaneously, AI models for energy optimization use consumption data to efficiently manage resources and reduce operational costs of the infrastructure. The use of advanced AI techniques, such as reinforcement learning algorithms, allows for dynamic adjustment of infrastructure operations to achieve energy savings

Key words: Artificial, implementation, detection, intelligence.

Introduction

In today's digital world, IT infrastructure is the foundation upon which almost every organization rests, from small businesses to global corporations. Given the increasing dependence on technology, the efficiency, security and reliability of IT systems are becoming key factors in maintaining a competitive advantage. IT infrastructure includes all technical components that enable the operation of computer systems, including servers, networks, databases, applications and cloud services.

As organizations face constant changes in business and technology requirements, the complexity and scalability of IT infrastructure becomes a challenge. Given the rapid development of technology, infrastructure is becoming increasingly dynamic, which implies frequent changes in capacity and performance needs.

Managing the IT infrastructure in this environment is not a simple task. Traditional approaches that rely on manual configuration, intervention and reaction to emerging problems are no longer sufficient to monitor performance and quickly resolve technical challenges. The problem of scaling, optimizing resources, as well as maintaining real-time performance are becoming more and more challenging for IT managers and organizations. In this sense, the need for predictive tools that can anticipate potential failures, detect anomalies in system operation and allocate resources in a timely manner become crucial for ensuring stability and business continuity.

AI Security Risk	Description	Key Defense Strategies
Data Poisoning	Malicious data injected into training sets	Secure data pipelines Data auditing Federated learning
Model Inversion	Extraction of sensitive data from model outputs	Differential privacy RBAC Query monitoring
Prompt Injection	Malicious prompts manipulating model behavior	Input validation Context separation Zero-trust approach
Model Theft	Replication of model via API probing	Watermarking Rate limiting Anomaly detection
Evasion Attacks	Slight input changes cause incorrect outputs	Adversarial training Anomaly detection Robust architectures
Lack of Transparency	Inability to understand model decisions	Explainable AI tools Hybrid models
Supply Chain Vulnerabilities	Insecure third-party components	Dependency scanning Patching Source validation
Insecure APIs	Poorly secured model endpoints	API gateways RBAC Request monitoring
Deepfakes	AI-generated media used for fraud or misinformation	Detection tools MFA Digital watermarking
Poor Governance	No policies guiding AI use	Formal governance Internal training Model tracking

Picture 1.The top 10 AI security risks and key strategies to protect systems [1]

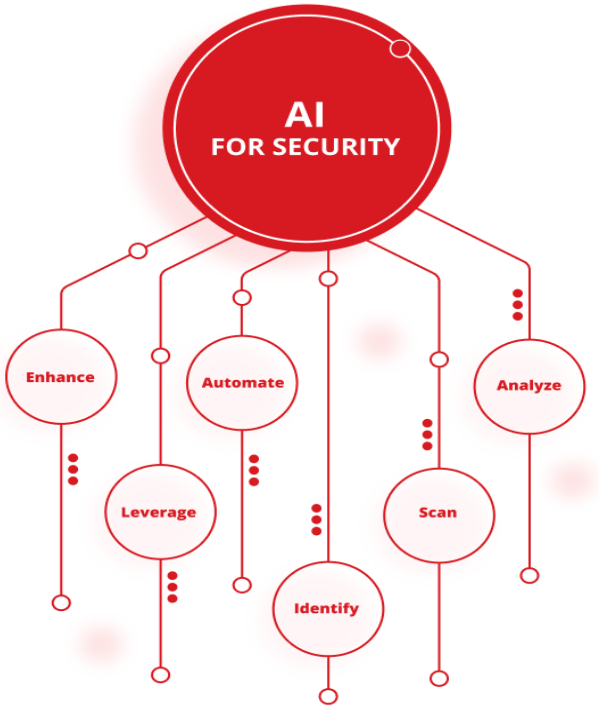
In today's digital age, the security of information systems is becoming one of the most critical issues for organizations, governments and individuals. Artificial intelligence (AI) is increasingly recognized as a key tool in the fight against cybercrime. Thanks to its ability to process large amounts of data, identify behavioral patterns and automate incident responses, AI is transforming the approach to cybersecurity.

ML models can be trained to recognize normal user or system behavior, and based on that, detect anomalies that may indicate compromise. Such systems are particularly effective in detecting the so-called zero-day attacks, which cannot be identified using conventional methods because they are not yet known to the security community.[2]

Tools like Splunk, Darktrace, and IBM QRadar use AI to scan network flows, logs, and user behaviors, identifying threats that would be difficult or impossible for a human to recognize. A key advantage is the ability of these systems to process large amounts of data in real time and to adapt to new types of attacks.[3]

[1] https://mindgard.ai/blog/top-ai-security-risks
[2] Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In 2010 IEEE Symposium on Security and Privacy

Another important component of AI in cyber security is the automation of response to threats. AI-based systems can not only identify an attack, but also take automatic steps to neutralize it. This is most often carried out through the so-called Security Orchestration, Automation and Response (SOAR) tools.



Picture 2. What AI using includes [4]

When an AI system detects a compromised user account, it can automatically deactivate the account, isolate the affected node in the network, and initiate an incident response protocol. This approach reduces the reaction time from several hours to just a few seconds, which minimizes damage.[5]

AI systems in threat defence

AI systems can be used to automatically classify threats by risk level, allowing security teams to prioritize the most dangerous incidents. Automation reduces the number of false positives, relieving security teams of having to manually analyze each warning signal.

The benefits of introducing AI in cyber security are numerous: higher detection accuracy, faster response to incidents, adaptability to new types of threats and the possibility of continuous learning. However, there are also challenges. AI systems can become the target of attacks through the so-called adversarial attacks, where attackers manipulate input data to confuse algorithms.[6]

Another challenge is the need for a large amount of quality data for model training, as well as the possibility of bias in the analyses. Also, full reliance on AI can lead to a reduction in critical oversight and human control, which in the cyber security domain can be risky. AI has the potential to revolutionize cyber security through advanced methods of threat detection and response automation. By using machine learning and SOAR systems, organizations can make the performance of security teams

[3] IBM. (2022). Threat Intelligence Index. IBM Security.
[4] https://www.trendmicro.com/content/dam/trendmicro/global/en/what-is/ai/ai-security/ai-security-illustration.png
[5] G. Garcia, et al. (2020). SOAR platforms: Improving response time in security incidents. Journal of Cybersecurity
[6] Biggio, B., & Roli, F. (2018). Wild patterns: Ten years after the rise of adversarial machine learning. Pattern Recognition, 84

more efficient, faster and more resilient to modern threats. However, for the successful implementation of AI systems, it is necessary to ensure transparency, oversight and proper training of security teams.[7]

Aplications of AI in security

The application of artificial intelligence (AI) in the field of security represents a revolutionary step in improving the protection of information systems, infrastructure and public safety. From video surveillance, through the detection of cyber attacks, to predictive analytics in police and military operations, AI systems offer capabilities that exceed the capacities of traditional approaches. However, such implementation is not without its challenges – especially when it comes to ethical dilemmas, technical limitations and potential errors in threat detection.

AI systems, especially those based on machine learning and deep learning, are capable of analyzing huge amounts of data in real time and detecting threats much faster than human personnel could. This ability is particularly important in situations where the speed of reaction is crucial - such as preventing attacks, detecting unauthorized access or automatically blocking suspicious activity in information systems.

For example, in modern surveillance systems for video analytics, AI algorithms are trained to recognize behavior patterns, analyze movements and even facial emotions, thus enabling the detection of potentially dangerous situations such as theft, violence or access to protected areas. When the system detects an anomaly – such as a person staying too long in an unusual location, or moving against the direction of crowd movement – it can automatically alert security services, trigger door locks, or activate higher-level cameras.

A similar principle is used in cyber security, where AI tools continuously monitor network traffic and user behavior. Using techniques such as anomaly detection or predictive analytics, these systems can recognize unusual patterns that indicate potential attacks (eg phishing, DDoS, ransomware) before the damage actually occurs.

In addition to speed, a key advantage of AI in these contexts is its ability to learn from previous incidents – meaning the system becomes more effective over time. In this way, AI not only accelerates the response, but also enables proactive security, which represents a fundamental shift in the protection paradigm.[8]

Intelligent monitoring and analysis in the UK

In the UK, several cities have introduced advanced AI technologies into their video surveillance systems to improve public safety and the effectiveness of police interventions. One of the most famous examples is London, where intelligent systems for video analytics based on artificial intelligence are used.

London is one of the most informed cities when it comes to the application of smart surveillance technologies. Millions of CCTV cameras are installed across the city, and AI systems analyze the footage in real time. These systems use algorithms for facial recognition, suspicious behavior detection and Automatic Number Plate Recognition (ANPR).

For example, AI can recognize a person suspected of criminal acts even if he is among a large group of people. When the system detects such a person, it automatically notifies the police officers, which accelerates the search and intervention.

In addition, in cases of mass gatherings or protests, AI analytics help detect potential incidents such as fights or crowds that can escalate into dangerous situations. This reduces reaction time and the risk of escalation.

[7] Darktrace. (2021). Autonomous Response Technology in Cyber Defense. Technical White Paper
[8] Bou-Harb et al. (2020). *Cyber Threat Intelligence: Towards an AI-Driven Security Strategy*

Cities like Manchester and Glasgow are also using AI technologies to improve safety. A system has been developed in Manchester that uses AI to detect unauthorized access in industrial zones and sensitive facilities. In Glasgow, AI has been implemented into traffic control systems, where it monitors and identifies unsafe driving patterns that can lead to traffic accidents.

The implementation of AI in the security systems of these cities resulted in a reduction in the crime rate in certain locations, as well as an increase in the efficiency of police patrols and interventions. Also, the costs associated with manual monitoring and inspections have been reduced.

Conclusion

AI systems significantly contribute to the standardization and automation of processes within security environments, which enables the minimization of errors caused by the human factor. Human operators often work under stressful conditions, under pressure and with limited time resources, which can lead to fatigue, inattention or reduced concentration. All of this increases the risk of omissions and errors that can have serious security consequences.

Automated AI tools, on the other hand, do not suffer from these limitations. They continuously and consistently monitor systems and processes, analyzing data according to predefined models and algorithms. This is particularly important in areas such as malware detection, where timely and accurate identification of malicious software is critical to preventing harmful attacks.

In addition, AI systems make it possible to reduce the number of false alarms that would otherwise burden human operators, because through continuous learning and adaptation, they improve their models and reduce classification errors. In this way, security personnel can focus on real threats and make better decisions, while routine and burdensome tasks are taken over by automation. In short, the integration of AI into security systems not only increases efficiency, but also the quality of protection by reducing the risk of human errors and improving detection accuracy.

Refferences:

- Biggio, B., & Roli, F. (2018). Wild patterns: Ten years after the rise of adversarial machine learning. Pattern Recognition, 84
- Bou-Harb et al. (2020). *Cyber Threat Intelligence: Towards an AI-Driven Security Strategy*
- Darktrace. (2021). Autonomous Response Technology in Cyber Defense. Technical White Paper
- G. Garcia, et al. (2020). SOAR platforms: Improving response time in security incidents. Journal of Cybersecurity
- https://mindgard.ai/blog/top-ai-security-risks
- https://www.trendmicro.com/content/dam/trendmicro/global/en/what-is/ai/ai-security/ai-security-illustration.png
- IBM. (2022). Threat Intelligence Index. IBM Security.
- Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In 2010 IEEE Symposium on Security and Privacy