IJSRA

International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

Int. J. Sci. Res. Arch.
IJSRA
International Journal of Science and Research Archive
Research Journal Archive, INDIA

(Review Article)

Check for updates

Cyber threats in the age of artificial intelligence: Exploiting advanced technologies and strengthening cybersecurity

Driss ABBADI * and Abdelkader Lachkar

Department of Legal, Political and Economic Research, Interdisciplinary Faculty of Taza, University of Sidi Mohamed Ben Abdellah, Fez, Morocco.

International Journal of Science and Research Archive, 2024, 13(01), 2576–2588

Publication history: Received on 06 September 2024; revised on 12 October 2024; accepted on 15 October 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.13.1.1961

Abstract

The rapid advancement in artificial intelligence (AI) technologies has led to the emergence of new and complex cyber threats. These threats rely on AI capabilities to target digital systems and infrastructures with greater precision and efficiency, making them more difficult to detect and counter. AI is being used to develop automated cyber-attacks, such as malware that can adapt and learn from its digital environment. Conversely, AI is also employed to enhance cyber security by enabling early threat detection, analyzing suspicious behaviors, and providing rapid responses to hacking incidents.

Based on the above, the research idea we intend to present at this article on how to achieve a balance between leveraging AI technologies to protect digital systems and the growing risks associated with using these technologies to develop more sophisticated cyber-attacks.

Research Methodologies: The study requires employing both the descriptive-analytical and inferential methodologies.

Keywords: Cyber Threats; Artificial Intelligence; Advanced Technologies; Cybersecurity

1. Introduction

Artificial Intelligence (AI) in today's digital age is one of the most prominent technological innovations that has brought about a transformative shift in many fields, including cyberspace. AI enables systems to perform tasks beyond traditional human capabilities, such as analyzing big data, making automated decisions, and self-learning. These capabilities have made AI a vital tool in enhancing cybersecurity, as it is used to detect, predict, and prevent attacks. However, despite its immense benefits, this technology is also being exploited maliciously by cyber attackers.

The advanced technologies provided by AI represent a double-edged sword. On the one hand, they strengthen the ability of systems to withstand cyber threats by enabling early detection of suspicious activities and analyzing abnormal behaviors. On the other hand, they are being exploited by malicious actors to develop complex cyberattacks that are difficult to detect, such as AI-powered malware and intelligent phishing attacks.

The issue lies in the fact that AI-based cyberattacks are not limited to the traditional threat to digital security, but pose a new threat that requires innovative and novel cybersecurity strategies. In this context, important questions arise about how to positively harness AI to enhance cybersecurity, and at the same time, how to confront the advanced cyber threats that rely on AI.

* Corresponding author: Driss ABBADI

1.1. The Importance of the Topic

The subject under study is one of the vital and important issues. Cyber threats are gaining increasing importance in the age of artificial intelligence due to the development of advanced technologies that can be used in dual ways, thereby heightening the risks of cyberattacks. Attackers are able to exploit AI to develop new and more complex methods for breaching security systems, posing a real challenge to cybersecurity. In this context, enhancing cybersecurity becomes essential to protect sensitive data and information, and to ensure the safety of information systems. This necessitates the development of AI technologies to improve the ability to detect and respond to attacks, making this topic a key focus in global security discussions.

1.2. Research Problem

Amid the rapid development of AI technologies, humanity faces new and growing challenges in the field of cybersecurity. Cyber threats have become more complex and diverse, with attackers exploiting advanced AI tools such as machine learning and big data analysis to execute precise and sophisticated attacks that are difficult to detect or counter using traditional means. On the other hand, these same technologies can play a crucial role in enhancing cybersecurity defenses by improving the capabilities of threat detection systems and enabling faster and more efficient responses. However, this duality in the use of AI presents a complex issue: how to balance leveraging this technology to enhance global cybersecurity while preventing its harmful use against individuals and institutions.

1.3. Research Questions

- What are the advanced cyber threats resulting from the exploitation of AI?
- What are some real-world cases of advanced cyber threats resulting from AI exploitation?
- How is the technical aspect manifested in enhancing cybersecurity using AI and advanced technologies?
- How is the legislative and political aspect manifested in enhancing cybersecurity?

1.4. Study Objectives

This study aims to explore the mutual effects between the development of AI and the increasing cyber threats. It analyzes the methods used by cyber attackers to exploit technologies such as machine learning, neural networks, and predictive analytics to carry out complex and precise attacks targeting digital infrastructure and critical institutions. The study also addresses how AI can be employed in developing advanced cyber defense systems capable of effectively detecting and countering unconventional threats. Additionally, the study does not overlook the legislative and political aspects that contribute to achieving cybersecurity.

1.5. Study Significance

The significance of this study lies in the growing challenges posed by the rapid technological advancements to digital security. With the development of AI, cyber attackers have become more capable of carrying out complex and unexpected attacks, threatening individuals, institutions, and critical infrastructure. By studying this topic, the importance of developing innovative strategies that leverage AI technologies to enhance cyber defense systems and enable them to respond to advanced threats quickly and efficiently is highlighted.

Moreover, the study contributes to guiding discussions on the formulation of legislative and political efforts that align with this evolving digital era, helping to strengthen protection against cyberattacks and create a safer digital environment.

2. Research Methodology

The study will adopt a descriptive analytical method to provide a detailed description of cyber threats in the age of AI and explain the importance of developing innovative strategies that utilize AI technologies to enhance cyber defense systems and enable them to respond to advanced threats with speed and efficiency.

2.1. Research Axes

The study addresses the topic through two main axes: the first focuses on advanced cyber threats resulting from AI exploitation, while the second deals with how advanced AI technologies, legal and legislative efforts, and international collaborations can be used to enhance cybersecurity.

2.2. Section one: Advanced Cyber Threats in the Age of Artificial Intelligence

With the rapid advancement in artificial intelligence (AI) technology, new applications have emerged that allow attackers to exploit this technology to carry out "advanced" cyberattacks. These attacks utilize techniques that adapt to targeted systems and have the ability to hide or learn from the digital environment, making them more precise and harder to detect. To address this dilemma, this section focuses on analyzing the different types of advanced cyber threats that have emerged as a result of AI exploitation, with reference to some real-life incidents.

2.3. First: Cyber Threats Resulting from AI Exploitation

This section examines how advanced AI technologies are being leveraged to execute complex cyberattacks. The main types of these threats can be addressed as follows:

2.3.1. Adversarial AI Attacks:

Adversarial attacks exploit vulnerabilities in AI models to manipulate the system into making incorrect decisions, disrupting security systems. These attacks are among the most significant challenges facing modern machine learning and AI systems. They involve subtly modifying input data in a way that remains undetectable to humans but causes AI systems to make incorrect decisions. For instance, by adding slight "noise" to an image, a human may still recognize its content without issue, but an AI model might interpret the image entirely incorrectly [1].

This type of attack is not limited to images but can also affect textual, audio, or video data. One well-known example of adversarial attacks is in facial recognition systems, where minor alterations to a photo can prevent the system from accurately identifying individuals, posing a significant cybersecurity threat [2]. Additionally, these attacks present a considerable risk to AI applications in sensitive sectors, such as altering traffic signals in a way that seems normal to human observers but is misinterpreted by the system, potentially leading to catastrophic decisions [3]. Furthermore, these attacks could extend beyond security and transportation to healthcare systems. Nowadays, many healthcare systems rely on AI to analyze medical data and make diagnostic decisions. However, if these systems are subjected to adversarial attacks, diagnostic results could be manipulated, leading to inappropriate treatments for patients [4].

2.3.2. AI in Malware:

AI is increasingly used to develop advanced malware that is more effective at evading detection and adapting to targeted systems. AI has drastically changed the landscape of cybersecurity, particularly in the area of malware [5]. AI-powered malware can bypass traditional defense mechanisms in unprecedented ways, making detection and mitigation more complex. These malware programs rely on advanced techniques such as machine learning (ML) and deep learning (DL) to enhance their ability to analyze the behavior of targeted systems and identify exploitable vulnerabilities. Additionally, AI allows malware to dynamically modify itself based on the targeted system's environment, making it harder for traditional antivirus programs and security systems to detect them.

This development enables malware to remain hidden for extended periods, allowing it to execute attacks in stages and with greater precision. AI-powered malware also uses big data analytics to assess behavior patterns and systems it targets, quickly processing large volumes of data to determine the most effective attack strategies [6].

On the other hand, cybersecurity researchers are working on developing AI-based defensive techniques to combat these threats, such as using machine learning to predict malware behavior and analyze attack patterns before they occur, allowing defensive systems to take proactive measures [7]. Nevertheless, the ongoing race between AI-driven malware and smart defense technologies continues, with both sides striving for superiority in this fast-evolving field.

2.3.3. Exploitation of AI in Ransomware and Phishing Attacks:

AI enables the analysis of vast amounts of data at high speed, allowing attackers to design sophisticated phishing attacks based on accurate data collected about victims. The use of AI in ransomware, phishing, and social engineering has become an increasing challenge in cybersecurity, as AI technology has made these attacks more complex and effective. Cybercriminals rely on AI to analyze users' personal data and customize phishing messages, increasing the likelihood of victims falling into the trap [8].

AI-enhanced tools enable attackers to automate and scale up their operations, making attacks faster, larger, and more targeted. These AI-driven attacks craft convincing messages tailored to specific recipients, increasing the success rate of deception and data theft.

In addition to traditional phishing tactics, attackers increasingly use AI-backed voice and video cloning techniques to impersonate trusted individuals such as family members, coworkers, or business partners. By manipulating and creating highly realistic audio and video content, these attackers deceive victims into revealing sensitive information or approving fraudulent transactions [9]. AI-powered phishing is becoming more targeted and personalized, using data collected from multiple sources to determine optimal times and methods for attacks [10].

Regarding ransomware attacks, AI is opening new possibilities for cybercriminals. Previously, some limiting factors in ransomware attacks involved the expertise and effort required to execute a successful attack. While ransomware-as-a-service (RaaS) exists, allowing attackers to outsource part or all of the ransomware operation, this requires trust in the ransomware service provider. With AI, many of these limitations are removed. Attackers can automate time-consuming tasks and improve their operations, making it easier for them to scale attacks, improve their effectiveness, and increase the number of potential victims [11].

Examples include:

- Using machine learning to blend with normal activities, such as hiding data breaches within regular traffic, making it difficult for organizations to detect and stop attacks.
- Automating the search for targets on social media by aggregating information from multiple sites, enabling faster and more effective phishing attacks.
- Crafting more effective phishing emails, even in non-native languages, using AI to draft convincing and precise messages.
- Analyzing vulnerabilities to determine how to bypass security measures before they are patched.
- Writing and refining code using natural language commands, making the attack process much easier and faster [12].

2.3.4. Control of Malicious Botnets through Artificial Intelligence:

The use of artificial intelligence in cybersecurity faces many challenges, including the risk of counter-attacks, where malicious botnets are designed to exploit vulnerabilities in AI algorithms [13]. A botnet is a network of compromised internet-connected computers controlled remotely by attackers for malicious and illegal purposes. The term comes from the "bot" software, short for "robot," due to its automated behavior. A bot is a highly advanced piece of malware that incorporates elements from viruses, worms, spyware, and other malicious programs. The person who controls a botnet is known as a "botmaster" or "bot herder," and they go to great lengths to keep their identity hidden.

Unlike older malware such as viruses and worms, the primary motivation behind running a botnet is financial gain. Botnets can be incredibly lucrative, earning operators hundreds of dollars a day. A botmaster can either rent out the processing power of the botnet to others or make direct profits by sending spam, distributing spyware to aid in identity theft, and even extorting money from companies by threatening a Distributed Denial of Service (DDoS) attack.

It is no surprise that many network security researchers consider botnets to be one of the most dangerous security threats on the internet today. Once a single bot program is installed on a victim's computer, the possibilities are almost endless. For instance, the attacker can obtain your online passwords, drain your bank accounts, use your computer as a remotely controlled "zombie" to search for other victims, send spam emails, and even launch DDoS attacks. Bots and botnets are the latest trend in the evolution of internet malware. Their black-hat developers have built upon decades of experience with viruses, worms, trojans, and other malware to create highly sophisticated programs that are difficult to detect and remove.

Typical botnets contain several hundred to several thousand members, although some have been discovered with more than 5.9 million members. As of January 2011, industry analysts estimated that up to 590 million computers (approximately 36% of all internet hosts) were infected with bot software.

Before the rise of botnets, the primary motivation for online attacks was fame and notoriety. By design, these attacks were loud and easy to detect. Notable examples include the Melissa email virus (1999), the ILOVEYOU virus (2000), the Code Red virus (2001), the Slammer virus (2003), and the Sasser virus (2004). While the impact of these viruses and worms was severe, the damage was relatively short-lived, mainly limited to the cost of downtime and the work hours required for cleanup. Once the infected files were removed and vulnerabilities patched, the attackers no longer had control [14].

In contrast, botnets are built on the premise of extending the attacker's control over their victims. To achieve long-term control, a bot must remain hidden throughout every phase of its lifecycle, unlike its predecessors. As a result, most bots have a relatively small network footprint and do not generate much traffic during regular operation. Once installed, the only required network activity consists of incoming commands and outgoing responses [15].

2.4. Second: Cyberattacks Born from Artificial Intelligence (A Study of Real Incidents)

In recent years, the world has witnessed a significant development in the use of artificial intelligence (AI) across various fields, including cyberattacks. While AI provides powerful tools for protecting systems and networks, it has also become an advanced tool exploited to execute complex and effective cyberattacks. This article explores the most notable cyberattacks that have utilized AI, focusing on real-life incidents supported by academic references.

2.4.1. Deepfake Attack on British Engineering Giant "Arup":

Deepfake technology, which leverages advanced AI and machine learning techniques to create highly realistic but fake media, has emerged as a significant cybersecurity challenge by manipulating audio and visual content to produce misleading and convincing simulations. This technology—deepfakes—has the potential to undermine trust in digital media and create a range of security risks.

It can be used to impersonate individuals in phishing attacks, manipulate public opinion through false information, and disrupt organizational operations through misleading communications [16]. One real-life example of this attack involves the British engineering giant "Arup," which revealed in January 2024 that it was defrauded through a deepfake scam amounting to $25 million. Arup, a multinational design and engineering company behind globally renowned buildings such as the Sydney Opera House, confirmed that one of its employees in Hong Kong was deceived into transferring $25 million to fraudsters. The company reported the fraud to Hong Kong police, confirming that fake voices and images were used in the process. In February, Hong Kong police revealed that the complex scam tricked the employee into joining a video call with individuals they believed to be the company's CFO and other staff members, who turned out to be deepfake versions. The police did not disclose the names of the company or the parties involved. Rob Greig, Arup's global head of IT, commented: "Like many other companies around the world, our operations are regularly targeted by attacks, including invoice fraud, phishing scams, WhatsApp voice impersonations, and deepfakes. We've seen the number and complexity of these attacks rise sharply in recent months" [17].

2.4.2. Deepfake Attacks in Public Politics:

In 2024, shortly before the New Hampshire primary elections, an AI-generated robocall mimicking President Biden's voice was created, urging voters not to participate in the primary, falsely claiming they should "save" their votes for the general election in November 2024. An ordinary voter hearing this message might reasonably believe that Biden had actually recorded the message and that they should comply with the request—effectively restricting their right to vote. Looking ahead, it's easy to imagine other fake messages from trusted voices being used to dissuade citizens from voting, weaken their ability to vote, or create false alarms about emergencies such as fires or attacks, prompting voters to stay home on election day. Furthermore, there is a real risk that AI could intensify the disproportionate targeting of disinformation against Black and Brown voters, who already face many barriers to equal participation in the democratic process. AI also creates new opportunities for bad actors to undermine election administration or sow unjustified doubts about election outcomes. AI technologies can easily be used to fake images or false evidence of wrongdoing, such as tampering with or shredding ballots. This would not only erode public confidence in election results but could also trigger more public threats of violence against election officials. In recent years, nonpartisan election workers have faced unprecedented levels of threats and harassment while attempting to ensure a smooth and fair democratic process, and AI technology could make their situation worse. Even after all votes are cast, AI could be used to fabricate audio recordings of a candidate falsely claiming they manipulated results or to generate other misinformation that might convince supporters of a losing campaign to disrupt vote counting and certification processes, which have already become increasingly politicized and were the basis for the attempt to sabotage the 2020 presidential election. AI-generated fake media has already been used to influence major elections in Argentina and Slovakia [18].

2.4.3. Emotet Malware Attacks:

In 2021, a malware strain called "Emotet" emerged, utilizing AI to analyze victims' online behavior. This malware was capable of generating personalized phishing emails, increasing the chances of deceiving users. Additionally, Emotet adapted to cybersecurity defenses and improved the effectiveness of attacks by learning from previous attempts [19].

Emotet infected over 1.6 million computers worldwide, causing damage amounting to hundreds of millions of dollars globally. Acting Deputy Attorney General John Carlin stated, "Emotet malware and its network infected hundreds of thousands of computers across the United States, including our critical infrastructure, causing millions of dollars in damage to victims worldwide." U.S. Attorney Matt Martin from the Northern District of Carolina added, "Cybercrime transcends geographic and political borders and costs American citizens and businesses billions of dollars annually. This was certainly true with Emotet" [20].

2.4.4. Attacks on Autonomous Vehicles:

Security breaches targeting sensor systems in autonomous vehicles (AVs) are a major concern due to the potential risk they pose to vehicle safety and passenger well-being. AVs rely on a complex array of sensors, essential for understanding the surrounding environment and facilitating critical decision-making. Compromising these systems could lead to distorted environmental perception, increasing the risk of accidents and safety hazards. Historically, there have been numerous demonstrations highlighting the potential for such security attacks on AV sensor systems. These include sophisticated spoofing attacks capable of tricking sensors into recognizing non-existent obstacles, as well as jamming attacks that disrupt the sensors' ability to detect objects accurately. Alarmingly, the equipment needed to carry out such attacks is often simple and easily accessible, like laser pointers or wireless transmitters, raising the possibility of more attackers [21].

For example, experts demonstrated that it is possible to manipulate CMOS camera sensors to prevent driver-assistance programs in Tesla and Baidu's Apollo vehicles from reading traffic signs. This experiment represents a deadly threat to drivers if it falls into the hands of hackers. Instead of directly targeting driver-assistance programs, the attack focuses on CMOS camera sensors installed in Tesla vehicles, as well as Baidu Apollo's self-driving taxis. Specifically, the GhostStripe technology, used to read traffic signals, relies on virtual markings invisible to humans to read traffic signals. GhostStripe marks are typically visible in specific wavelengths, such as infrared, and self-driving vehicle cameras are equipped with filters to detect these specific wavelengths. When a camera captures a scene, the image processing algorithm searches for these invisible lines. If traffic signals are equipped with these markings, they can be easily detected and recognized, even in low-light, fog, or rain conditions [22].

2.4.5. Banking Algorithm Manipulation Attacks:

In the first half of 2021, ransomware attacks increased by 1318%, with banks being disproportionately affected. BEC (Business Email Compromise) attacks may have increased by 4% due to the COVID-19 pandemic. Bank hacks are on the rise. Since banks are interconnected, a cyberattack on one bank could threaten another. State-sponsored cyberattacks pose a threat to U.S. banks. As internet and mobile banking use increases, cybercrime is also on the rise. Cybercrimes include credit card fraud, spam, ATM thefts, and identity theft. Valuable banking data is at risk. Hackers can exploit financial and banking data in various ways. Banks' digital footprints increase the attack surface. Cyberattacks can disrupt power, military equipment, and sensitive information. They can steal medical records, disrupt phone and computer networks, and destroy data and systems. This makes the banking sector highly vulnerable to attack. Hackers can exploit financial and banking data in multiple ways [23].

2.4.6. DDoS Attacks on AWS Using AI:

In 2020, Amazon Web Services (AWS) was subjected to a Distributed Denial of Service (DDoS) attack. The attackers used AI to organize and direct traffic in a complex and targeted manner. AI enhanced the effectiveness of the attack, causing a significant service disruption. This attack, reported on AWS, relied on a DDoS attack using the CLDAP protocol (Connectionless Lightweight Directory Access Protocol), which was combined with amplification attacks—a technique common in large-scale attacks. Reflection and amplification attacks remain preferred tools, alongside CLDAP and other popular attacks such as exposed UDP ports, DNS, NTP, SSDP, and SNMP services that rely on UDP. What makes these attacks more dangerous is their two main characteristics: first, the amplification of the attacker's payload can generate traffic five, ten, or even a hundred times larger than their requests. Second, they can spoof to hide the attacker’s traces while directing the payload toward a specific target of their choice [24].

These incidents highlight the increasing challenges facing cybersecurity as AI becomes more prevalent. As this technology continues to evolve, cyberattacks are expected to increase in complexity and effectiveness, requiring organizations to take proactive steps to develop defensive technologies, also leveraging AI, to counter these persistent threats.

2.5. Section two: Enhancing Cybersecurity through AI Technologies and Legislative and Policy Efforts

In the midst of the rapid technological advancements taking place globally, cybersecurity has become a critical concern for institutions, governments, and individuals alike. As reliance on digital technology increases across various aspects of daily life, new and complex cyber threats have emerged that jeopardize the informational infrastructure and sensitive data. This growing threat necessitates the development of more advanced protective measures that can adapt to the constantly evolving risks. In this context, artificial intelligence (AI) plays a pivotal role in enhancing cybersecurity by offering innovative solutions that facilitate early detection of cyberattacks, analyzing large data sets to identify suspicious patterns, and countering cyber threats at their early stages. Additionally, AI technologies help automate numerous cybersecurity processes, such as network monitoring and user behavior analysis, enabling faster and more accurate detection of attacks.

Moreover, it is essential to recognize the importance of legal and policy efforts, which must evolve alongside technological advancements to achieve comprehensive cybersecurity.

In this section, we will explore how AI and advanced technologies are being applied in the field of cybersecurity, as well as the legislative and policy efforts that must be considered to achieve a holistic approach to cybersecurity.

2.6. First: The Technical Dimension in Enhancing Cybersecurity through the Use of Artificial Intelligence

Cybersecurity is one of the major challenges facing organizations in the era of advanced information technology, as cyber threats continue to grow and become increasingly complex. In this context, the technical dimension emerges as a key factor in strengthening cybersecurity, particularly through the use of artificial intelligence (AI) technologies.

In this section, we will discuss several technical aspects that play a vital role in enhancing the protection of systems and information. We will explore how AI is utilized for early threat detection, data analysis, advanced encryption and blockchain technologies, as well as automated response systems to counter cyberattacks.

2.6.1. Machine Learning in Threat Detection:

Machine learning enables the detection of cyberattacks by analyzing anomalous patterns in data traffic and suspicious behaviors automatically and efficiently. This advanced technology plays a vital role in cybersecurity, allowing systems to continuously learn and improve in detecting cyber threats. Unlike traditional systems that rely on fixed rules to identify attacks, machine learning leverages big data analysis to detect unusual patterns that may signal threats [25]. This adaptability and ability to learn from past data make cybersecurity systems more flexible and effective in countering evolving attacks. For example, deep learning algorithms are used to analyze user behavior and network traffic patterns to detect any abnormal activity that might indicate a breach attempt [26].

Additionally, machine learning plays a pivotal role in email classification, detecting phishing attempts, and spam messages. By analyzing the specific structures of these emails, algorithms can distinguish harmful messages from regular ones with high accuracy [27]. Furthermore, machine learning improves defensive systems through predictive analysis, allowing the prediction of future attacks based on historical patterns, enabling organizations to take preventive measures before an attack occurs [28].

Moreover, the integration of machine learning with big data analytics enhances the effectiveness of cybersecurity systems in detecting complex threats and targeted attacks. These tools improve system responses to zero-day attacks and advanced persistent threats [29]. Hence, machine learning is an indispensable part of any comprehensive cybersecurity strategy, providing tools that can adapt to rapidly changing threats and significantly improve the systems' ability to detect and counter attacks in real-time.

2.6.2. Deep Learning and Big Data Analytics:

Deep learning is one of the most prominent fields of artificial intelligence that has significantly impacted big data analytics. It is characterized by its ability to process vast amounts of complex data and is built on artificial neural networks that mimic human neural networks. Thanks to recent advancements in computing hardware, training these networks on large datasets has become possible, improving prediction accuracy in various applications, including computer vision, natural language processing, and speech recognition [30]. Big data, a crucial part of the modern information revolution, is generated daily across the internet and smart devices. This massive data holds valuable insights for strategic decision-making, but its complexity and volume make traditional analysis methods insufficient. Deep learning offers advanced solutions for extracting hidden patterns and insights from big data [31].

In combating cyber threats, deep learning and big data analytics play a key role by providing advanced techniques for quickly and efficiently detecting and analyzing attacks. While traditional security methods struggle to keep pace with the evolving nature of cyber threats, deep learning, as a branch of machine learning, excels in identifying subtle patterns in large and complex datasets, enabling it to detect unusual activities that may indicate a potential cyberattack. When combined with big data analytics, which allows handling large volumes of data from diverse sources, cybersecurity systems' performance can be improved by offering in-depth insights into threats based on social network analysis, user behavior, and system logs [32].

One of the most notable advantages of deep learning in cybersecurity is its ability to enhance threat detection accuracy. Traditional systems often rely on predefined rules, making them less efficient in identifying unknown vulnerabilities or novel attacks. However, deep learning models exhibit a high level of adaptability and continuous learning, enabling them to detect new and unprecedented attack patterns [33].

Additionally, big data analytics supports deep learning models by providing them with vast datasets to improve the accuracy of learning algorithms. As the amount of available data increases, the system's ability to distinguish between normal and malicious behavior improves. For instance, these systems can detect distributed denial-of-service (DDoS) attacks and phishing attempts in their early stages by analyzing large-scale data, contributing to effectively mitigating the impact of such attacks [34].

The combination of deep learning and big data analytics forms a powerful tool in addressing modern cyber threats. Not only does it enhance the system's ability to detect threats with greater accuracy, but it also provides a rapid response and prediction of potential vulnerabilities, making it an essential part of cybersecurity strategies in the digital era [35].

2.6.3. Advanced Encryption and Blockchain Technologies:

Blockchain and modern encryption techniques provide a secure means of storing data and ensuring its integrity, contributing to the protection of sensitive systems against attacks. Advanced encryption and blockchain technologies are fundamental pillars in the current digital revolution, offering effective solutions in cybersecurity, digital governance, and distributed financial systems. Advanced encryption encompasses a variety of algorithms used to ensure data confidentiality and protect it from unauthorized access. For example, asymmetric encryption and symmetric encryption techniques are used to secure data. Asymmetric encryption relies on a pair of keys—one public key for encryption and a private key for decryption—enhancing the security of digital communications such as e-commerce and public key infrastructure [36].

Blockchain technology, first introduced in the context of digital currencies like Bitcoin, involves distributing data across a decentralized network of nodes, making the data immutable once it is recorded in linked blocks. Data is authenticated through digital signatures based on encryption techniques, ensuring integrity and transparency in digital transactions. These features make blockchain attractive not only in digital currencies but also in other areas like smart contracts, supply chains, and digital identity management [37].

The combination of advanced encryption and blockchain technologies offers an effective way to enhance digital security, with blockchain-based networks utilizing encryption to secure data and verify user identities. This combination makes systems more resistant to cyberattacks and increases trust in digital processes. For instance, these technologies have been used by major banks to enhance the security of financial transactions and improve transparency [38]. Additionally, blockchain's use in health information and digital identity records demonstrates the potential of this technology in enhancing security and privacy in sensitive sectors [39].

Given the increasing reliance on digital systems in everyday life, the development and implementation of advanced encryption and blockchain technologies are essential to ensure the security of financial systems, protect personal data, and enhance user trust in modern technologies. These technologies represent future solutions for addressing global cybersecurity challenges, offering significant levels of security and decentralization while reducing the need for central intermediaries in various processes.

2.6.4. Automated Cyberattack Response Systems:

Automated cyberattack response systems involve using artificial intelligence (AI) and machine learning (ML) to proactively eliminate threats before they escalate into breaches. Attackers today leverage automation and AI to launch sophisticated, large-scale cyberattacks, particularly in complex and borderless IT environments like multi-cloud systems. In addition to streamlining cybersecurity operations, automation helps prevent, detect, and respond to cyber threats without human intervention. Organizations can maximize the value of AI by using modern security orchestration techniques to build strong defenses for both present and future threats. Automation eliminates many manual processes and reduces alerts, allowing security operation center (SOC) analysts to complete repetitive security tasks much faster [40].

One of the greatest advantages of automated response systems is their ability to act in real-time without delay or waiting for human intervention. Paquette and Bassett (2018) highlight that these systems possess advanced capabilities for detecting threats and triggering automatic defensive responses, such as shutting down vulnerable systems or isolating suspicious traffic immediately. These automated responses significantly reduce potential damage and minimize the time needed to contain threats, making them effective tools for combating advanced cyberattacks like ransomware and zero-day attacks [41].

Moreover, the ability to quickly adapt to evolving threats is a key feature of these systems, as they adjust to new cyberattacks and use deep learning techniques to predict future attacks. A study published in the "*Journal of Open Source Developments*" indicates that integrating AI with big data analytics helps build systems capable of anticipating cyberattacks before they occur, thus enhancing their ability to protect critical systems such as energy networks and healthcare infrastructures [42].

While AI can bolster defenses against cyberattacks by improving detection and rapid response, there is a need for legislation that ensures the safe and responsible use of these technologies. Policies aim to promote AI innovation while preventing illegal exploitation, thereby balancing efforts to enhance security with legislative and political diligence to protect digital infrastructure, privacy, and data security.

2.7. Second: The Legislative and Political Framework Supporting and Enhancing Cybersecurity

Given the contemporary reality of the widespread use of IT systems and AI that primarily focus on digital alternatives to paper-based systems, and under the immense pressure from the pervasive role of IT in various aspects of life—resulting in cyber threats and legal violations—a multitude of legislative initiatives and political efforts have emerged to address these challenges [43]. Several initiatives have been adopted by states, most notably Russia's first proposal in 1998. The UN General Assembly also established a committee of governmental experts to start discussions on information and communication technologies. Initially, only 15 countries were represented, including the five permanent members of the Security Council, as the focus was on specific areas of expertise. When membership was gradually expanded, the number of countries remained relatively limited (25 countries between 2016 and 2017) [44].

In 2001, the Budapest Convention was enacted as the main legal text for international cooperation in prosecuting and combating cybercrimes [45]. Initially, the Budapest Convention aimed to harmonize cybercrime laws and address the limited capacity for cross-border investigations and prosecutions of online crimes. Recently, its focus has shifted toward addressing the challenges of collecting digital evidence [46]. The success of the convention and its additional protocol is evidenced by the fact that many non-European countries have joined it, and it has gained the status of an international tool with the inclusion of the U.S., Japan, Australia, South Africa, Canada, and others. With the convention's enforcement across all signatory countries, it has become an essential tool for managing cyber risks and threats. Its importance lies in the practical measures that the signatory countries commit to incorporating into their national laws [47].

On February 14, 2017, the Global Commission on the Stability of Cyberspace (GCSC) began its work. This independent commission of 28 international experts from the private sector, academia, the technical community, and NGOs was established by the Dutch Ministry of Foreign Affairs with support from prominent companies and several countries. It was tasked with providing new standards for responsible behavior to maintain security and stability in cyberspace, as outlined in its final report, presented at the Paris Peace Forum in November 2019 [48].

In 2019, Microsoft, Mastercard, and the Hewlett Foundation established the CyberPeace Institute. This independent and neutral NGO, headquartered in Geneva, aims to reduce the frequency, impact, and scale of cyberattacks, hold perpetrators accountable for the harm caused, and assist vulnerable communities. The institute provides evidence-based knowledge, raises awareness of the humanitarian impact of cyberattacks, and gives victims a voice to highlight the damage and consequences of such attacks. Although it lacks official attribution authority, the institute evaluates states' adherence to international law and norms of responsible behavior.

Microsoft's appointment of a representative for cyber matters related to the United Nations in January 2020 further demonstrates the company's desire to actively participate in international discussions [49]. However, the aforementioned initiatives are still limited due to the reluctance of many countries to allow non-state actors to define states' rights, obligations, and responsibilities. The role of non-state actors in international discussions is increasingly recognized, despite the existence of differences between countries.

In addition to these developments, the emergence of cyber threats has led countries and international organizations to intensify efforts to confront the dangers posed by such threats [50]. Among these scholarly efforts is the Tallinn Manual on the International Law Applicable to Cyber Warfare, drafted by a group of leading international law scholars. This manual, along with the principles outlined in the Erice Declaration on Cyber Stability, prepared by the Permanent Monitoring Panel on Information Security of the World Federation of Scientists (WFS), highlights key legal responses to cyber threats.

The Erice Declaration calls for global cyber peace and urges the international community to assess how ICTs support daily life, evaluate cyber threats, analyze the impact of cybercrime and cyber conflicts, and define a roadmap for future action. The declaration, endorsed by the 42nd international seminar on global emergencies held in Erice, Sicily, on August 20, 2009, represents a scientific achievement aimed at fostering cyber peace by ensuring the peaceful and beneficial use of cyberspace. Its goals include preventing or mitigating the effects of cyber conflicts, safeguarding individuals' fundamental rights in democratic societies, and ensuring the free flow of information in cyberspace.

In summary, cybersecurity peace lies in the hands of all actors—companies, citizens, institutions, and public/private initiatives, but especially governments within the framework of international cooperation. The main tools are legal, technical (secured applications), and political (persuading or compelling states to ensure the free flow of information) [51].

3. Conclusion

Artificial Intelligence (AI) in the current digital age is a driving force for change in various fields, including cybersecurity. This study has shown that AI holds immense potential to enhance digital protection through early threat detection and precise analysis of suspicious activities. Conversely, it also poses a complex challenge as cyber attackers exploit it to develop advanced and automated attacks. Therefore, addressing this significant challenge requires a concerted effort in technology, legislation, and politics to achieve an effective balance between strengthening cybersecurity and preventing the misuse of AI in malicious attacks.

The findings of this research include:

- AI enables security systems to improve their ability to counter complex cyber threats, but attackers are also using it to develop advanced cyberattacks.
- The rapid evolution of AI technologies has outpaced advancements in security measures, making it crucial to develop parallel defensive strategies.
- There is a pressing need for international cooperation between governments and institutions to formulate policies and legislation that enhance cybersecurity and keep pace with the growing threats.
- As for the recommendations, the following can be listed:
- Increasing investment in AI systems: It is necessary to boost investments in developing AI-based systems capable of predicting and effectively responding to cyber threats.
- Establishing new legislation: Strengthening cybersecurity requires a clear legal framework to regulate the use of AI in cyberspace and prevent its exploitation in harmful activities.
- Promoting international cooperation: International collaboration between countries and institutions should be enhanced to share information and expertise regarding advanced cyber threats and how AI can be used to counter them.
- Raising awareness and building capacities: It's important to increase awareness and provide training in AI technologies among cybersecurity professionals to ensure the full utilization of its capabilities and mitigate potential threats.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572. 2015.

[2] Kurakin A, Goodfellow I, Bengio S. Adversarial examples in the physical world. arXiv preprint arXiv:1607.02533. 2016.

[3] Szegedy C, Zaremba W, Sutskever I, Bruna J, Erhan D, Goodfellow I, Fergus R. Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199. 2014.

[4] Yuan X, He P, Zhu Q, Li X. Adversarial examples: attacks and defenses for deep learning. IEEE Trans Neural Netw Learn Syst. 2019;30(9):2805-24.

[5] Saxe J, Berlin K. Deep neural network based malware detection using two-dimensional binary program features. In: Proceedings of the 2018 IEEE Security and Privacy Workshops (SPW). 2018;13(2):30-7. Available from: https://doi.org/10.1109/SPW.2018.00017

[6] Anderson HS, Woodbridge J, Filar B. DeepDGA: adversarially-tuned domain generation and detection. In: Proceedings of the 2016 ACM Workshop on Artificial Intelligence and Security; 2016. Available from: https://doi.org/10.48550/arXiv.1610.01969

[7] Goodman B, Flaxman S. European Union regulations on algorithmic decision-making and a "right to explanation." AI Mag. 2017;38(3):50–7.

[8] Begou N, et al. Exploring the dark side of AI: Advanced phishing attack design and deployment using ChatGPT. In: Proceedings of the IEEE Conference on Communications and Network Security (CNS); 2024. Available from: https://bit.ly/3ZBZ7mB.

[9] Begou N, et al. Exploring the dark side of AI: advanced phishing attack design and deployment using ChatGPT. In: Proceedings of the IEEE Conference on Communications and Network Security (CNS); 2024. Available from: https://bit.ly/3ZBZ7mB

[10] Aleroud A, Zhou L. Phishing environments, techniques, and countermeasures: a survey. Comput Secur. 2017;68:160-196. doi:10.1016/j.cose.2017.04.006.

[11] Hurley A. How AI is changing ransomware and how you can adapt to stay protected. Barracuda; 2023 Nov 13. Available from: https://bit.ly/4gFc99h

[12] Hurley A. How AI is changing ransomware and how you can adapt to stay protected. Barracuda; 2023 Nov 13. Available from: https://bit.ly/4gFc99h

[13] Papernot N, McDaniel P, Jha S, Fredrikson M, Celik ZB, Swami A. The limitations of deep learning in adversarial settings. In: 2016 IEEE European Symposium on Security and Privacy (EuroS&P); 2016. p. 372-87. IEEE.

[14] Ramsbrock D, Wang X. The botnet problem. In: Computer and Information Security Handbook. Available from: https://doi.org/10.1016/B978-0-12-394397-2.00012-X

[15] Ramsbrock D, Wang X. The botnet problem. In: Computer and Information Security Handbook. Available from: https://doi.org/10.1016/B978-0-12-394397-2.00012-X

[16] Broklyn P, et al. Deepfakes and cybersecurity: detection and mitigation. 2024 Jul 3. Available from: bit.ly/4di0bPK. Accessed on September 23, 2024.

[17] Magramo K. British engineering giant Arup revealed as $25 million deepfake scam victim. WRALNEWS. 2024 May 17. Available from: bit.ly/3TF08qq. Accessed 2024 Sep 23.

[18] Noti A. How artificial intelligence influences elections, and what we can do about it. Campaign Legal Center. 2024 Feb 28. Available from: bit.ly/4evMrlF. Accessed 2024 Sep 23.

[19] Europol. Emotet dismantled in international cyber operation. Europol Press Release; 2021.

[20] Europol. Emotet dismantled in international cyber operation. Europol Press Release; 2021.

[21] Giannaros A, et al. Autonomous vehicles: Sophisticated attacks, safety issues, challenges, open topics, blockchain, and future directions. J Cybersecur Priv. 2023. Available from: https://bit.ly/3XxS8sp. Accessed 2024 Sep 23.

[22] Loupia M. L'attaque de la technologie GhostStripe empêche les véhicules Tesla et Apollo Baidu de lire les panneaux de signalisation. 2024 May 11. Available from: bit.ly/47F6RpV. Accessed 2024 Sep 23.

[23] Gill MA, et al. Cyber attacks detection through machine learning in banking. Bull Bus Econ. 2023;12(2):34-45. Available from: https://doi.org/10.5281/zenodo.8310116.

[24] Nicholson P. AWS hit by largest reported DDoS attack of 2.3 Tbps. A10 Network. 2020 Jun 24. Available from: bit.ly/4df6Udo. Accessed 2024 Sep 23.

[25] Sommer R, Paxson V. Outside the closed world: On using machine learning for network intrusion detection. In: Proceedings of the IEEE Symposium on Security and Privacy; 2010; Oakland, CA. p. 305-16. Available from: https://doi.org/10.1109/SP.2010.25.

[26] Buczak AL, Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Commun Surv Tutor. 2016;18(2):1153-76. Available from: https://doi.org/10.1109/COMST.2015.2494502.

[27] Chandrasekaran M, Narayanan K, Upadhyaya S. Phishing email detection based on structural properties. In: Proceedings of the New York State Cyber Security Conference; 2006. p. 30-41.

[28] Almseidin M, Alzubi M, Kovacs S, Alkasassbeh M. Evaluation of machine learning algorithms for intrusion detection system. In: 2017 IEEE International Conference on Cyber Security and Cloud Computing (CSCloud); 2017; New York, NY, USA. p. 1-6. doi: 10.1109/SISY.2017.8080566.

[29] Sarker IH, Kayes ASM, Watters P. Cybersecurity data science: An overview from machine learning perspective. J Big Data. 2020;7(1):1-29. doi: 10.1186/s40537-020-00318-5.

[30] Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572. 2015.

[31] Chen CP, Zhang CY. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Inf Sci. 2014;275:314-47. doi: 10.1016/j.ins.2014.01.015.

[32] Chio C, Freeman D. Machine learning and security: Protecting systems with data and algorithms. Sebastopol (CA): O'Reilly Media; 2018.

[33] Shone N, Ngoc TN, Phai VD, Shi Q. A deep learning approach to network intrusion detection. IEEE Trans Emerg Topics Comput Intell. 2018;2(1):41-50.

[34] Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, et al. A state-of-the-art survey on deep learning theory and architectures. Electronics. 2019;8(3):292.

[35] Vinayakumar R, et al. Applying deep learning approaches for network traffic classification and intrusion detection. In: Proceedings of the International Conference on Advances in Computing, Communications and Informatics; 2017. p. 2353-8. Available from: https://doi.org/10.1109/ICACCI.2017.8126198

[36] Rivest RL, Shamir A, Adleman L. A method for obtaining digital signatures and public-key cryptosystems. Scientific Research. 1978;120-6. Available from: https://doi.org/10.1145/359340.359342

[37] Tapscott D, Tapscott A. Blockchain revolution: how the technology behind bitcoin is changing money, business, and the world. New York: Penguin; 2016.

[38] Narayanan A, Bonneau J, Felten E, Miller A, Goldfeder S. Bitcoin and cryptocurrency technologies. Princeton University Press; 2016.

[39] Zyskind G, Nathan O, Pentland A. Decentralizing privacy: Using blockchain to protect personal data. In: 2015 IEEE Security and Privacy Workshops; 2015; San Jose, CA. p. 180-4.

[40] Paloalto. What is Security Automation? [Internet]. [cited 2024 Sep 25]. Available from: https://bit.ly/3ZWFDJV

[41] Bassett S, Paquette M. Improve Security Analytics with the Elastic Stack, Wazuh, and IDS. Elastic Blog [Internet]. [cited 2024 Oct 12]. Available from: https://bit.ly/4gx6rpX

[42] Sharma G, Narayan R. AI-Driven Cybersecurity: Enhancing System Resilience With Advance Security Automation Program (ASAP). J Open Source Develop. 2024;1-19.

[43] Abdel Hamid A. The legal and legislative framework for digitalization and artificial intelligence. Al-Bahith Journal for Legal and Judicial Studies. 2023;50:1.

[44] Douzet F, Gery A. Cyberspace is used, first and foremost, to wage wars: proliferation, security and stability in cyberspace. J Cyber Policy. 2021;6(2):233–52. https://doi.org/10.1080/23738871.2021.1937253.

[45] Martins dos Santos B. Budapest Convention on Cybercrime in Latin America: a brief analysis of adherence and implementation in Argentina, Brazil, Chile, Colombia, and Mexico. Bernabó G, translator. Derechos Digitales América Latina; 2022 [accessed 2023 May 29]. p. 6-7. Available from: http://bitly.ws/FUp7.

[46] Klynge C. Cooperating against cybercrime: 20 years on from the Budapest Convention. [accessed 2023 May 29]. Available from: http://bitly.ws/FVur.

[47] Al-Ashqar Jabour M. Cybersecurity: the concern of the age. Arab Center for Legal and Judicial Research; 2016 [accessed 2023 Apr 14]. Available from: http://bitly.ws/DTBH.

[48] Douzet F, Gery A. Cyberspace is used, first and foremost, to wage wars: proliferation, security and stability in cyberspace. J Cyber Policy. 2021;6(2):233–52. https://doi.org/10.1080/23738871.2021.1937253.

[49] Douzet F, Gery A. Cyberspace is used, first and foremost, to wage wars: proliferation, security and stability in cyberspace. J Cyber Policy. 2021;6(2):233–52. https://doi.org/10.1080/23738871.2021.1937253.

[50] Abdel Aal SM. Legitimate defense against cyber attacks. Egyptian Journal of International Law. 2023;1:79. Available from: https://bitly.ws/3e6IB (Accessed: February 25, 2024).

[51] Ventre D. La cyberpaix: un thème stratégique marginal. Revue internationale et stratégique. 2012;87:83-91. Available from: CAIRN.INFO.