World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Review Article)

The intersection of Artificial Intelligence and cybersecurity: Challenges and opportunities

Adewale Daniel Sontan [1,]* and Segun Victor Samuel [2]

[1] *Independent Researcher, Newark, New Jersey, 07103, USA.*
[2] *Independent Researcher, Johannesburg, South Africa.*

World Journal of Advanced Research and Reviews, 2024, 21(02), 1720–1736

Publication history: Received on 13 January 2024; revised on 24 February 2024; accepted on 26 February 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.21.2.0607

Abstract

The fusion of artificial intelligence (AI) with cybersecurity represents a paradigm shift in our efforts to safeguard digital assets against a dynamic threat landscape. This manuscript comprehensively analyses AI's transformative role in cybersecurity, covering foundational principles, advanced methodologies, and ethical considerations. This article begins with exploring fundamental AI techniques such as machine learning and natural language processing. The manuscript delineates their applications in bolstering threat detection, vulnerability analysis, and incident response. Traditional approaches to vulnerability analysis are juxtaposed with AI-driven methodologies, highlighting the efficacy of automated scanning, threat prioritization, and adaptive risk assessment. Moreover, the manuscript delves into the pivotal role of AI-driven automation in expediting incident response, minimizing human error, and fortifying overall security postures. Ethical and privacy concerns surrounding AI deployment in cybersecurity are carefully examined, emphasizing the importance of responsible decision-making, privacy protection, and transparency. Looking ahead, emerging trends such as adversarial machine learning and zero trust security present promising avenues for further exploration, offering opportunities to enhance digital resilience against evolving threats.

Keywords: Artificial Intelligence; Cyber Security; Threat Detection; Incident Response; Vulnerability Scanning; Machine Learning

1. Introduction

In the contemporary landscape, the proliferation of digital technologies has undeniably revolutionized virtually every aspect of human society. From the seamless connectivity facilitated by cloud computing to the intricate web of interconnected devices comprising the Internet of Things (IoT) [1], our reliance on digital infrastructure has become fundamental to modern existence. These technologies have ushered in unprecedented levels of efficiency, productivity, and connectivity, reshaping how we conduct business, communicate with one another, and interact with the world around us [2].

However, amidst this digital revolution lies a shadowy underbelly characterized by the omnipresent threat of cyber-attacks [2]. With each technological advancement, threat actors -- from individual hackers to sophisticated cybercriminal syndicates and even nation-state adversaries -- exploit vulnerabilities in our digital ecosystem to perpetrate various malicious activities [3]–[5]. These threats manifest in diverse forms, encompassing everything from ransomware attacks that encrypt critical data and demand hefty ransoms to sophisticated phishing schemes that steal sensitive information or infiltrate networks for espionage purposes [6].

* Corresponding author: Sontan Adewale Daniel

The consequences of these cyber-attacks can be severe and far-reaching, extending beyond mere financial losses to encompass reputational damage, operational disruptions, and even threats to national security [7]–[9]. For businesses, the fallout from a successful cyber-attack can result in significant financial losses, loss of customer trust, and legal ramifications [10]. In critical infrastructure sectors such as energy, healthcare, and transportation, the potential impact of cyber-attacks on public safety and essential services cannot be overstated [11].

Moreover, the evolving nature of cyber threats presents a formidable challenge for defenders, as attackers continuously adapt their tactics, techniques, and procedures to circumvent existing security measures [12]. The rapid pace of technological innovation further compounds this challenge, as new vulnerabilities emerge with each new advancement, providing fertile ground for exploitation by malicious actors [13].

As such, the need for robust cybersecurity measures has never been more pressing. Organizations across industries must proactively fortify their defences against cyber threats, employing a multi-layered approach encompassing robust perimeter defences, continuous monitoring and detection capabilities, and rapid incident response protocols. Furthermore, fostering a cybersecurity awareness and resilience culture among employees and stakeholders is essential to mitigating the human factor in cyber risk [14].

The emergence of Artificial Intelligence (AI) marks a pivotal moment in cybersecurity, presenting a paradigm shift in how we defend against and mitigate cyber threats [15]. Once relegated to the realms of science fiction and speculative futurism, AI has now transcended its conceptual confines to become a tangible force with profound implications for cybersecurity.

Unlike traditional cybersecurity approaches that often rely on manual intervention and predefined rulesets, AI introduces a new era of automation and intelligence-driven defence mechanisms. At the heart of this transformation lies techniques such as Machine Learning (ML) and Deep Learning (DL) [15], [16], which endow AI systems with the capacity to analyze vast amounts of data at unprecedented speeds and complexities.

Machine Learning algorithms, for instance, excel at recognizing patterns and correlations within datasets, enabling them to discern subtle anomalies that may evade traditional rule-based detection methods [15]. By continuously learning from historical data and adapting to evolving threat landscapes, ML-powered systems can identify emerging threats with remarkable accuracy, even in the absence of explicit instructions or predefined rules.

Deep learning, a subset of ML inspired by the structure and function of the human brain's neural networks [17], further enhances AI's cybersecurity capabilities. Through sophisticated neural architectures capable of hierarchical feature extraction and representation learning, Deep Learning models can autonomously extract intricate features from raw data [16], enabling them to uncover subtle indicators of compromise that may elude human analysts.

The implications of AI for cybersecurity are profound and far-reaching. By augmenting human capabilities with intelligent automation and data-driven decision-making, AI promises to revolutionize various facets of cybersecurity operations, including threat detection, vulnerability assessment, and incident response.

In the realm of threat detection, AI-powered systems can sift through vast volumes of network traffic, log data, and security event telemetry in real-time, flagging suspicious activities and potential intrusions with unprecedented speed and accuracy [18]. Moreover, AI's ability to contextualize disparate data sources and correlate seemingly unrelated events enables it to discern subtle indicators of compromise indicative of sophisticated cyber-attacks [19].

Similarly, in vulnerability assessment and management, AI-driven approaches streamline the process of identifying and prioritizing security vulnerabilities within complex IT environments [20]. By analyzing code repositories, system configurations, and historical attack data, AI can pinpoint exploitable weaknesses precisely, allowing organizations to allocate resources more effectively and proactively mitigate the most critical risks [21].

In incident response, AI empowers security teams with enhanced capabilities for rapid detection, analysis, and containment of security incidents [22]. Through automated incident triage, AI can prioritize alerts based on their severity and likelihood of impact, enabling responders to focus their efforts where needed most. Furthermore, AI's capacity for predictive analytics and threat intelligence correlation enables it to anticipate emerging threats and recommend proactive countermeasures to bolster defences against future attacks [23].

In essence, the emergence of AI as a game-changer in cybersecurity represents a fundamental shift in the way we approach cyber defence. By leveraging the power of intelligent automation and data-driven decision-making, AI enables defenders to stay one step ahead in the perpetual cat-and-mouse game of cybersecurity, levelling the playing field against increasingly sophisticated adversaries [15], [24]. As organizations continue to embrace AI-driven security solutions, they stand poised to bolster their resilience against evolving cyber threats and safeguard their digital assets with unprecedented efficacy and efficiency.

The primary aim of this review is to provide a comprehensive exploration of the intersection between AI and cybersecurity, focusing on the challenges and opportunities that arise from their integration. In essence, we seek to dissect the intricate relationship between AI and cybersecurity, unravelling the complexities that emerge when these two domains converge.

Central to our examination is a critical analysis of the role played by AI across various dimensions of cybersecurity, including but not limited to threat detection, vulnerability analysis, and incident response. By scrutinizing the applications of AI in these key areas, we aim to elucidate how AI-driven approaches can enhance our defensive capabilities against the ever-evolving landscape of cyber threats.

Ultimately, we aim to offer readers a comprehensive understanding of the opportunities and challenges posed by the intersection of AI and cybersecurity. By shedding light on both the transformative potential and the ethical considerations inherent in this integration, we hope to pave the way for a more informed and responsible approach to harnessing the power of AI in cybersecurity, thereby bolstering our collective defences against emerging cyber threats.

2. Opportunities of AI in Cybersecurity

2.1. Threat Detection

2.1.1. Traditional methods of Threat Detection

Cybersecurity has long relied on traditional methods like signature-based and rule-based detection to identify threats [25]. These methods have served us well, but as attackers become more sophisticated and threats evolve ever-increasingly, these approaches reveal their limitations. Signature-based detection struggles with novel attacks, while rule-based methods often generate false positives and miss complex threats [26]. A brief description of the operations, effectiveness and limitations are highlighted in Table 1 [26], [27].

Table 1 Traditional methods of threat detection in cybersecurity

Method	Description	Operation	Effectiveness	Limitations
Signature-Based Detection	Relies on predefined patterns of known threats. These signatures are typically derived from indicators of compromise (IOCs), such as file hashes, network traffic patterns, or behavioural characteristics associated with known malware or attack techniques.	Security systems, such as antivirus software or intrusion detection systems (IDS), Compare data/traffic against known signatures. If a match is found, the system triggers an alert or takes predefined action to block or mitigate the threat.	Effective against known threats.	It may struggle to detect novel or previously unseen attacks, as attackers can easily evade detection by modifying their tactics, techniques, and procedures (TTPs) or using polymorphic malware variants.
Rule-Based Detection	Uses predefined rules or heuristics to identify suspicious activities or behaviours indicative of a security threat. These rules are typically based on known attack patterns, vulnerability exploits, or abnormal behaviour thresholds.	Security systems analyze incoming data or network traffic in real-time, applying predefined rules to identify potentially malicious activity. If a match is found, the system triggers an alert or takes appropriate action, such as blocking suspicious activity or	Effective for detecting known attack patterns or common security vulnerabilities.	It may generate false positives or miss sophisticated attacks that do not conform to predefined rules, making it less effective against novel or advanced threats.

		notifying security personnel for further investigation.		

While signature-based and rule-based detection methods have been foundational in cybersecurity for many years, they have limitations when it comes to detecting unknown or zero-day threats. As cyber threats become increasingly complex and dynamic, organizations are turning to more advanced and adaptive approaches, such as AI-driven threat detection, to augment traditional methods and enhance their cybersecurity defences.

2.1.2. AI-based Threat Detection

AI possesses a unique capability to analyze vast amounts of data at an unprecedented speed, far surpassing the capabilities of human analysts [28]. This proficiency is rooted in the advanced algorithms and computational power that underpin AI systems, enabling them to sift through massive datasets comprising diverse sources of security telemetry with remarkable efficiency. Here's a brief explanation of how AI can analyze data to identify anomalies and potential threats faster than humans:

- Advanced Algorithms: AI systems leverage sophisticated algorithms, including but not limited to ML and DL, to process and analyze large volumes of data. These algorithms are designed to learn from data patterns, adapt to new information, and make predictions or decisions based on the insights derived from the data [29].
- Parallel Processing: AI systems are capable of parallel processing, allowing them to analyze multiple data streams simultaneously. Unlike human analysts who are limited by cognitive capacity and attention span, AI can handle massive datasets in parallel, significantly reducing the time required for analysis [30].
- Real-time Analysis: AI systems can perform real-time analysis of streaming data, such as network traffic logs and system event logs, as it is generated [31]. This enables organizations to detect and respond to security threats in real-time, minimizing the impact of potential cyber-attacks.
- Pattern Recognition: AI excels in identifying patterns, anomalies, and deviations within data. By analyzing historical data and learning from past incidents, AI can recognize abnormal behaviours or activities that may indicate a potential security threat [32]. This pattern recognition capability enables AI to flag suspicious events quickly and accurately.
- Automation: AI-driven automation streamlines the data analysis process, eliminating the need for manual intervention at every step. Once trained, AI systems can autonomously analyze data, identify anomalies, and trigger alerts or responses based on predefined criteria [15]. This automation significantly accelerates the pace of threat detection and response, enabling organizations to react swiftly to emerging cyber threats.
- Scalability: AI systems are highly scalable and capable of handling increasingly large and complex datasets without sacrificing performance [33]. AI can scale horizontally as data volumes grow by distributing computation across multiple processors or nodes, ensuring consistent and efficient analysis even as the data load increases.

2.1.3. Machine Learning and Deep Learning

It is noteworthy to state that ML and DL are pivotal AI techniques that have revolutionized threat detection in cybersecurity. Below is a detailed discussion of these techniques and their applications in threat detection:

Machine Learning (ML)

There are different machine learning techniques, as described below

- Supervised Learning: In supervised learning, AI systems are trained on labelled datasets, where each data point is associated with a specific class or category (e.g., malicious or benign). Through iterative training iterations, the system learns to map input data to the correct output labels, enabling it to classify new, unseen data accurately [34].
- Unsupervised Learning: Unsupervised learning involves training AI systems on unlabeled datasets, where the system aims to identify patterns or structures within the data without explicit guidance [35]. This technique is particularly useful for anomaly detection, as it can uncover deviations from normal behaviour without knowing what constitutes an anomaly.
- Reinforcement Learning: Reinforcement learning is a trial-and-error-based learning paradigm where AI systems learn to make decisions through interaction with an environment [36]. In the context of threat

detection, reinforcement learning can be employed to optimize decision-making processes, such as adaptive response strategies to evolving cyber threats.

Deep Learning (DL)

DL is a subset of ML that employs artificial neural networks with multiple layers of interconnected nodes known as neurons. These deep neural networks (DNNs) can automatically learn hierarchical representations of data, enabling them to extract intricate patterns and features from raw input data. The different types of DL are described thus:

- Convolutional Neural Networks (CNNs): CNNs are a deep neural network specifically designed to process structured grid-like data, such as images or time-series data [37]. In threat detection, CNNs can be applied to analyze network traffic patterns or identify malicious patterns in malware binaries.
- Recurrent Neural Networks (RNNs): RNNs are well-suited for processing sequential data, making them ideal for tasks such as natural language processing (NLP) and time-series analysis [38], [39]. In cybersecurity, RNNs can be used to analyze log data or detect anomalies in system behaviour over time.
- Generative Adversarial Networks (GANs): GANs consist of two neural networks – a generator and a discriminator – that are trained simultaneously in a competitive manner [40]. GANs have applications in generating synthetic data for augmenting training datasets or generating adversarial examples for testing the robustness of threat detection systems.
- By leveraging these AI techniques, threat detection systems can continuously adapt to evolving threats and identify previously unseen attack vectors with high accuracy. Moreover, ML and DL algorithms' inherent flexibility and adaptability enable threat detection systems to detect complex and sophisticated threats that may evade traditional signature-based approaches. A summary of the different types of ML and DL Techniques are presented in Table 2.

Table 2 Summaries of types of Machine and Deep Learning

Technique	**Description**	**Application in Threat Detection**
Supervised Learning (ML)	Trains on labelled data (malicious/benign) to classify new data accurately.	Identifying known threats and malware.
Unsupervised Learning (ML)	Identifies patterns in unlabeled data to detect anomalies.	Finding unknown threats and unusual behaviour.
Reinforcement Learning (ML)	Learns through trial-and-error, optimizing decisions.	Adapting response strategies to evolving threats.
Convolutional Neural Networks (DL)	Analyzes structured data like images and time-series.	Identifying malicious patterns in network traffic or malware binaries.
Recurrent Neural Networks (DL)	Processes sequential data like log data or system behaviour.	Detecting anomalies in system behaviour over time.
Generative Adversarial Networks (DL)	Generates synthetic data or adversarial examples for testing.	Augmenting training datasets and testing system robustness.

The different examples of successful AI-driven threat detection implementations are described in Table 3.

Table 3 Examples of successful AI-driven threat detection

Use Case	Example	Description
Anomaly Detection [41]	Darktrace	Uses AI to detect deviations from normal network traffic, user behaviour, and system logs. Identifies potential threats like data exfiltration, insider threats, and malware infections in real-time.
Behavioural Analysis [42], [43]	Splunk User Behavior Analytics (UBA)	Utilizes AI to analyze user activities across IT systems and applications. Detects suspicious activities like insider threats, credential misuse, and unauthorized access attempts.
Predictive Analytics [44], [45]	FireEye Threat Analytics Platform (TAP)	Leverages AI to analyze historical attack data and predict future threats. Identifies emerging attack vectors and helps organizations proactively defend against them.

These examples illustrate the effectiveness of AI-driven threat detection implementations across various use cases, including anomaly detection, behavioural analysis, and predictive analytics. By harnessing the power of AI algorithms, organizations can enhance their cybersecurity posture, detect threats in real-time, and respond swiftly to mitigate the impact of cyber-attacks.

2.2. Vulnerability Analysis

2.2.1. Traditional Approaches to Vulnerability Analysis:

Traditional vulnerability analysis methods involve manual inspection and assessment of various components within an organization's IT infrastructure, including software code, system configurations, and network architecture. Security professionals typically rely on their expertise and experience to identify potential security weaknesses that malicious actors could exploit. This process often includes:

- Manual Code Review: Security experts manually review software code to identify vulnerabilities, such as buffer overflows, injection flaws, or insecure authentication mechanisms [48]. This involves examining the code line by line, looking for coding errors or design flaws that could lead to security breaches.
- System Configuration Review: Security professionals inspect system configurations, including operating system settings, network configurations, and access controls, to identify misconfigurations or weaknesses that could expose the organization to cyber threats [49]. This involves comparing the current configuration against security best practices and industry standards to identify areas of concern.
- Network Infrastructure Assessment: Security teams analyze the organization's network infrastructure, including routers, switches, firewalls, and other network devices, to identify potential vulnerabilities or misconfigurations that could be exploited by attackers [50]. This may involve conducting network scans, reviewing firewall rules, and analyzing network traffic patterns to identify anomalous behaviour.

While effective in identifying known vulnerabilities and common security weaknesses, these traditional methods have limitations [50], [51]:

- Labour-Intensive: Manual vulnerability analysis is labour-intensive and time-consuming, requiring significant human effort and expertise to review large volumes of code, configurations, and network data.
- Limited Scalability: Traditional methods may struggle to scale to large and complex IT environments, particularly in organizations with extensive software development projects or distributed network infrastructure.
- Limited Coverage: Traditional approaches may overlook novel or emerging threats that do not match predefined signatures or patterns, leaving organizations vulnerable to zero-day exploits or advanced persistent threats.
- Human Error: Manual vulnerability analysis is prone to human error, with security professionals potentially overlooking critical vulnerabilities or misconfigurations due to fatigue, oversight, or lack of expertise.
- Difficulty in Keeping Pace with Threat Landscape: The rapid evolution of cyber threats and attack techniques makes it challenging for traditional vulnerability analysis methods to keep pace with the changing threat

landscape. New vulnerabilities are discovered regularly, requiring constant vigilance and proactive measures to address emerging security risks.

While these traditional methods serve us well, the evolving cyber landscape demands more. As organizations face increasingly sophisticated and dynamic cyber threats, there is a growing need for more advanced and automated approaches to vulnerability analysis, such as those enabled by AI and machine learning algorithms.

The next section will explore how AI revolutionizes vulnerability analysis, offering a more automated, scalable, and proactive approach to securing our digital castles.

2.2.2. AI-based Vulnerability Scanning:

AI can revolutionize vulnerability scanning and prioritize critical vulnerabilities by applying machine learning algorithms to analyze extensive datasets. AI-powered vulnerability scanners automate the process of identifying potential security weaknesses across software code, system configurations, and network assets. These scanners utilize sophisticated algorithms to detect patterns, anomalies, and historical attack data, enabling them to uncover vulnerabilities that may have previously gone unnoticed.

The automation of vulnerability scanning using AI has several significant impacts:

- Efficiency: By automating vulnerability scanning, AI-driven systems significantly reduce the time and effort required to identify and remediate security weaknesses. This efficiency allows organizations to conduct more frequent scans and assessments, ensuring continuous security posture monitoring [50], [52].
- Prioritization: AI-powered vulnerability scanners can prioritize vulnerabilities based on their severity, exploitability, and potential impact on the organization's security posture. By assigning risk scores or ratings to vulnerabilities, these systems enable security teams to focus their remediation efforts on addressing the most critical threats first, thereby maximizing the effectiveness of their security measures [53].
- Resource Allocation: By prioritizing critical vulnerabilities, AI-driven systems help organizations allocate their resources more effectively [54]. Security teams can direct their attention and resources towards addressing high-risk vulnerabilities that pose the greatest threat to the organization's assets and data, thereby reducing the overall risk of a successful cyber-attack.
- Accuracy: AI algorithms excel at identifying subtle patterns and anomalies within complex datasets, enabling them to accurately identify vulnerabilities that may be overlooked by traditional methods [54], [55]. This enhanced accuracy reduces the likelihood of false positives and false negatives, ensuring that security teams can trust the results generated by AI-driven vulnerability scanners [55].
- Adaptability: AI-powered vulnerability scanners can continuously learn and adapt to new threats and attack techniques, ensuring that organizations remain protected against evolving cyber threats [52], [53]. By analyzing historical attack data and security trends, these systems can update their models and algorithms to detect previously unseen vulnerabilities and attack vectors, thereby enhancing the organization's resilience to emerging security risks [55].

2.2.3. Tools and Platforms Leveraging AI for Vulnerability Analysis

Several tools and platforms leverage AI for vulnerability analysis, automating the detection and prioritization of security weaknesses across an organization's IT infrastructure.

In the fight against cyber threats, organizations are increasingly turning to AI to bolster their defences. AI-powered vulnerability analysis tools like IBM Security QRadar, Qualys AI-driven Threat Detection, and Tenable.sc are automating the detection and prioritization of security weaknesses across entire IT infrastructures. Let's take a closer look at these leading examples:

- IBM Security QRadar: This comprehensive platform acts as a security intelligence hub, leveraging AI and machine learning to scan for vulnerabilities and identify potential threats in real-time [56]. By analyzing network traffic, log data, and security events, QRadar can pinpoint critical vulnerabilities based on their severity, exploitability, and potential impact [57]. This allows organizations to focus their remediation efforts on the most pressing issues, maximizing their security posture.
- Qualys AI-driven Threat Detection: Operating in the cloud, this platform continuously monitors network assets, applications, and endpoints for vulnerabilities [58]. Its AI algorithms automate vulnerability scanning and prioritize critical issues, helping organizations stay ahead of potential threats. By leveraging machine learning

for threat detection and classification, Qualys provides actionable insights and recommendations for remediation, enabling proactive responses to security risks.

- Tenable.sc: Formerly known as SecurityCenter, this platform incorporates AI and machine learning to automate vulnerability scanning and threat detection [59]. Tenable.sc conducts comprehensive assessments of network assets, identifying vulnerabilities and misconfigurations across diverse IT environments. It prioritizes critical vulnerabilities based on their severity and potential impact, allowing organizations to prioritize remediation efforts and reduce their exposure to cyber threats.

These are just a few examples of the many AI-powered vulnerability analysis tools available today. By automating the process of identifying and prioritizing security weaknesses, these solutions are helping organizations strengthen their defenses and mitigate the risk of cyberattacks. As the digital landscape becomes increasingly complex and threats evolve, AI promises to play a vital role in safeguarding our digital world.

2.3. Incident Response

2.3.1. Traditional Incident Response Methods in Cybersecurity

Traditional incident response methods in cybersecurity involve a series of manual steps designed to identify, contain, and mitigate security incidents. These methods typically adhere to established frameworks such as the NIST Cybersecurity Framework or the SANS Incident Response Plan [61]. Traditional incident response encompasses several key activities:

- Identification and Triage of Security Alerts: Security teams monitor various sources of security alerts, including intrusion detection systems, security information and event management (SIEM) platforms, and endpoint detection and response (EDR) tools [62]. Analysts investigate its severity and relevance when an alert is triggered to determine if it warrants further investigation.
- Containment of the Incident: Once a security incident is confirmed, efforts are made to contain it to prevent further damage [63]. This may involve isolating affected systems or segments of the network, shutting down compromised services, or blocking malicious traffic.
- Investigation and Analysis: Analysts conduct a detailed investigation into the incident to determine its scope, impact, and root cause. This may involve analyzing log files, network traffic, system configurations, and other forensic evidence to understand how the incident occurred and what data or systems may have been compromised [64].
- Remediation and Recovery: Once the incident is contained and analyzed, remediation efforts begin to restore affected systems and data to a secure state [63]. This may involve patching vulnerabilities, restoring from backups, or deploying additional security controls to prevent similar incidents in the future.
- Post-Incident Review and Lessons Learned: After the incident has been resolved, a post-incident review is conducted to assess the effectiveness of the response efforts and identify areas for improvement [65]. Lessons learned from the incident are documented and used to update incident response plans, improve security controls, and enhance staff training.

While traditional incident response methods have been the cornerstone of cybersecurity operations for many years, they have several limitations:

- Time-Consuming and Resource-Intensive: Manual incident response processes can be time-consuming and resource-intensive, particularly in large and complex environments. Analysts may struggle to keep pace with the volume of security alerts and the speed at which modern cyber-attacks unfold.
- Risk of Human Error: Manual processes are prone to human error, leading to delays in detection and response or overlooking critical indicators of compromise. Analysts may also misinterpret data or make incorrect decisions under pressure, increasing the organization's exposure to cyber threats.
- Limited Scalability: Traditional incident response methods may struggle to scale to meet the demands of modern cyber threats, particularly in organizations with limited resources or expertise. As cyber-attacks become more sophisticated and widespread, manual response efforts may be insufficient to adequately protect the organization's assets and data.

Overall, while traditional incident response methods are effective in addressing security incidents, they have limitations in terms of scalability, efficiency, and effectiveness. To address these challenges, organizations are increasingly turning to automation and AI-driven solutions to augment and enhance their incident response capabilities.

2.4. AI-driven Automation into Incident Response

AI offers significant potential in automating repetitive tasks during incident response, thereby expediting remediation efforts and enhancing the overall efficiency of security operations. AI-powered incident response platforms leverage advanced algorithms, including machine learning and natural language processing, to analyze security alerts, prioritize incidents, and execute predefined response actions autonomously. These platforms are capable of handling large volumes of security alerts in real-time and can make data-driven decisions to address security incidents swiftly and effectively.

The integration of AI-driven automation into incident response processes yields several notable impacts [66]–[68]:

- Acceleration of Incident Response Times: By automating repetitive tasks such as triaging security alerts, categorizing incidents, and initiating predefined response actions, AI significantly reduces the time required to detect and respond to security incidents. This swift response helps minimize the window of opportunity for attackers, mitigating the potential impact of security breaches on the organization.
- Reduction of Human Error: AI-driven automation minimizes the risk of human error inherent in manual incident response processes. By consistently applying predefined response actions based on predefined rules and algorithms, AI ensures a consistent and reliable approach to incident handling, reducing the likelihood of oversight or misinterpretation of security data.
- Enhancement of Overall Security Posture: By freeing up security personnel from mundane and repetitive tasks, AI enables them to focus on more strategic and high-value activities, such as threat hunting, analysis, and proactive security measures. This shift towards proactive security measures strengthens the organization's overall security posture, making it more resilient against evolving cyber threats.
- Scalability and Flexibility: AI-powered incident response platforms are highly scalable and adaptable to changing threat landscapes and organizational requirements. These platforms can handle large volumes of security alerts and incidents without sacrificing performance, ensuring that organizations can effectively respond to varying complexity and scale security incidents.
- Continuous Learning and Improvement: AI algorithms can learn from historical incident data and feedback from security analysts, continuously improving their performance and accuracy over time. By leveraging machine learning techniques, AI-driven incident response platforms can adapt to new threats, trends, and attack techniques, enhancing their effectiveness in detecting and mitigating security incidents.

Utilizing AI for efficient incident response and threat prediction yields several other significant impacts [64], [69], [70]:

Prediction and Prevention of Future Attacks: By analyzing historical incident data and identifying emerging threat patterns, AI can predict and preemptively mitigate future attacks before they occur. This proactive approach to threat detection and prevention helps organizations stay one step ahead of cyber adversaries and minimize the impact of potential security breaches.

Improvement of Response Strategies: AI-driven analysis of incident data provides valuable insights into common attack patterns, tactics, and techniques used by threat actors. This information enables organizations to develop more effective response strategies tailored to specific threat scenarios, such as targeted phishing campaigns, ransomware attacks, or insider threats. By understanding how attackers operate, organizations can implement proactive security measures and countermeasures to mitigate the risk of successful cyber-attacks.

Enhancement of Situational Awareness and Threat Visibility: AI-powered threat intelligence platforms offer enhanced situational awareness and threat visibility across the organization's IT infrastructure. By continuously monitoring and analyzing security data in real-time, AI can identify anomalous behaviour, suspicious patterns, and potential indicators of compromise that may indicate an ongoing security incident or emerging threat. This increased visibility enables security teams to detect and respond to security incidents more rapidly, minimizing the time to detection and containment.

Reduction of Response Times and Resource Requirements: By automating the analysis of incident data and identifying actionable insights, AI streamlines incident response processes and reduces the time and resources required to respond to security incidents. AI-driven threat intelligence platforms can prioritize alerts, incidents, and vulnerabilities based on their severity, impact, and relevance, enabling security teams to focus their efforts on the most critical threats and security risks.

In summary, AI-driven analysis of incident data enables organizations to predict and prevent future attacks, improve response strategies, enhance situational awareness, and reduce response times and resource requirements. By leveraging AI for efficient incident response, organizations can strengthen their overall cybersecurity posture and mitigate the risk of cyber threats and security breaches.

Examples of AI-Powered Incident Response Solutions are highlighted in Table 4

Table 4 Examples of AI-Powered Incident Response Solutions

Name	Description	Impact
Darktrace [71]	Darktrace's Autonomous Response technology leverages AI algorithms to detect and respond to cyber threats in real-time. By continuously monitoring network traffic and user behaviour, Darktrace's AI-powered platform can identify anomalies and deviations from normal patterns indicative of potential security incidents. When a threat is detected, the system autonomously takes action to contain and mitigate the threat, preventing further damage to the organization's systems and data.	Darktrace's Autonomous Response technology enables organizations to respond to cyber threats swiftly and effectively, reducing the time to detection and containment. By automating incident response actions, Darktrace helps organizations minimize the impact of security incidents and strengthen their overall security posture.
Cynet [72]	Cynet's Security Platform combines AI-driven detection with automated response capabilities to protect organizations from advanced cyber threats. The platform continuously monitors endpoint, network, and user activity, leveraging machine learning algorithms to detect and analyze suspicious behaviour indicative of potential security incidents. When a threat is detected, Cynet's platform automatically initiates response actions, such as isolating affected endpoints, blocking malicious IP addresses, or quarantining suspicious files to contain and mitigate the threat.	Cynet's AI-driven detection and automated response capabilities enable organizations to rapidly detect and contain security incidents, reducing the risk of data breaches and other cyber threats. By automating response actions, Cynet helps organizations streamline their incident response processes and minimize the impact of security incidents on their business operations.
Palo Alto Networks Cortex XSOAR [73]	Palo Alto Networks Cortex XSOAR is an AI-powered security orchestration, automation, and response platform that streamlines incident response workflows, automates repetitive tasks, and coordinates response efforts across security teams and tools. Cortex XSOAR integrates with a wide range of security products and data sources, allowing organizations to centralize incident data and automate response actions based on predefined playbooks and workflows.	Cortex XSOAR enables organizations to orchestrate and automate incident response processes, reducing manual effort and accelerating response times. By integrating with existing security tools and data sources, Cortex XSOAR helps organizations improve collaboration between security teams and streamline incident response efforts, ultimately enhancing the organization's overall security posture.

Overall, AI-powered incident response solutions such as Darktrace, Cynet, and Palo Alto Networks Cortex XSOAR leverage AI algorithms to detect and respond to cyber threats in real-time, automate response actions, and streamline incident response workflows. These solutions help organizations enhance their incident response capabilities, reduce the time to detection and containment, and strengthen their overall security posture in the face of evolving cyber threats.

3. Challenges of AI in Cybersecurity

3.1. Data Privacy Concerns

With the proliferation of AI in cybersecurity, there's a growing concern about the potential misuse of personal data. AI algorithms often require access to large datasets for training, some of which may contain sensitive information about individuals. If not handled properly, there's a risk that AI algorithms could inadvertently expose or misuse this personal data, leading to privacy violations and breaches of confidentiality.

To address data privacy concerns, it's crucial to prioritize data anonymization techniques during AI development. By anonymizing personal data before feeding it into AI algorithms, organizations can minimize the risk of privacy breaches while still benefiting from the insights derived from the data. Furthermore, ethical considerations should guide the development and deployment of AI systems, ensuring that they adhere to principles of fairness, transparency, and respect for individual privacy rights.

Various regulations and frameworks govern the use of personal data in AI cybersecurity applications. For example, the General Data Protection Regulation (GDPR) in the European Union mandates strict rules for the processing and protection of personal data, including requirements for data minimization, purpose limitation, and user consent [15], [74]. Additionally, frameworks such as Privacy by Design and Privacy Enhancing Technologies (PETs) provide guidelines for integrating privacy protections into AI systems from the outset, ensuring that data privacy concerns are addressed proactively.

3.2. Explainability and Transparency

The "black box" problem refers to the opacity of complex AI models, where it's challenging to understand how they arrive at their decisions or predictions. This lack of transparency can undermine trust in AI systems, especially in critical applications like cybersecurity, where decision-making must be justified and understandable [47], [75]. Complex neural networks and deep learning algorithms often exacerbate this issue, as they operate in ways that humans do not interpret easily.

Explainable AI (XAI) is essential for security professionals to trust and validate the outputs of AI-driven cybersecurity systems. By providing insights into how AI algorithms reach their conclusions, XAI enables security analysts to understand the rationale behind AI-generated alerts, recommendations, or decisions [19]. This transparency fosters trust and confidence in AI systems, empowering security professionals to make informed judgments and take appropriate action in response to security incidents.

Significant advancements have been made in the field of explainable AI, with researchers developing techniques and methodologies to enhance the transparency of AI algorithms. Techniques such as feature importance analysis, decision tree visualization, and model-agnostic interpretability allow stakeholders to gain insights into AI decision-making processes [47]. Furthermore, organizations and industry groups actively promote transparency and accountability in AI development, advocating for adopting XAI principles and standards to ensure that AI systems are trustworthy and accountable.

3.3. Bias and Fairness

AI models trained on biased or unrepresentative datasets may inherit and amplify existing biases, leading to unfair or discriminatory outcomes. Biases in training data can stem from historical disparities, societal prejudices, or sampling biases, which may result in AI systems making biased predictions or decisions that disproportionately impact certain groups or individuals [75].

Ensuring unbiased data and model development is critical for promoting fair and ethical AI in cybersecurity. Organizations must strive to identify and mitigate training data biases through data preprocessing, bias detection, and algorithmic fairness testing. By promoting diversity and inclusivity in dataset collection and model training, organizations can develop AI systems that are more equitable and unbiased in their decision-making processes.

Industry efforts and best practices for mitigating bias in AI cybersecurity systems include developing bias detection and mitigation tools, establishing diverse and representative datasets, and promoting transparency and accountability in AI development processes [76]. Additionally, organizations can implement fairness-aware algorithms and techniques to mitigate bias in AI models, ensuring equitable outcomes across different demographic groups and contexts. By adopting these measures, organizations can build AI cybersecurity systems that are fair, transparent, and accountable, thereby promoting trust and confidence in their deployment and use.

4. Ethical and Privacy Considerations

4.1. Potential Ethical Implications of AI in Cybersecurity

The potential ethical implications of AI in cybersecurity are described below [76], [77]:

- Automated Decision-Making: The use of AI in cybersecurity raises concerns about automated decision-making, where AI algorithms autonomously detect and respond to security threats. Ethical considerations arise regarding the accountability and transparency of these decisions, especially when they impact individuals' rights and freedoms.
- Weaponization of AI: There are ethical concerns surrounding the potential weaponization of AI in cyber warfare and offensive cyber operations. The development and deployment of AI-powered cyber weapons raise questions about the morality of using AI to conduct attacks and the potential for unintended consequences or collateral damage.
- Surveillance and Privacy: AI-driven cybersecurity measures, such as intrusion detection systems and network monitoring tools, may inadvertently infringe on individuals' privacy rights by capturing and analyzing their digital activities without their consent. The ethical implications of mass surveillance and data collection must be carefully considered to balance security needs with individual privacy rights.

4.2. Privacy Concerns Associated with AI-Driven Security Measures

The privacy concerns associated with AI-Driven Security measures include [57], [78]:

- Data Collection and Usage: AI-powered security measures rely on extensive data collection and analysis, raising concerns about individuals' personal information privacy. Unauthorized access to sensitive data or misuse of personal information by AI algorithms could lead to privacy violations and breaches of confidentiality.
- Algorithmic Bias: AI algorithms used in cybersecurity may exhibit biases that disproportionately impact certain individuals or groups. Biased algorithms could result in discriminatory outcomes, such as false positives or unfair targeting, leading to privacy concerns and potential harm to affected individuals.
- Third-Party Risks: Outsourcing AI-driven security measures to third-party vendors or service providers introduces privacy risks, as organizations may have limited control over how their data is collected, processed, and protected. Ensuring the privacy and security of data shared with third parties is essential to mitigate the risk of data breaches or unauthorized access.

4.3. Strategies for Mitigating Ethical and Privacy Risks

4.3.1. Here are the strategies for mitigating ethical and privacy risks

- Ethical Frameworks and Guidelines: Organizations should adopt ethical frameworks and guidelines for the responsible development and use of AI in cybersecurity. These frameworks should emphasize principles such as transparency, accountability, fairness, and respect for individual rights to guide ethical decision-making and behaviour.
- Privacy-Enhancing Technologies: Implementing privacy-enhancing technologies, such as encryption, anonymization, and differential privacy, can help mitigate privacy risks associated with AI-driven security measures. Organizations can reduce the likelihood of privacy breaches and violations by protecting sensitive data and minimizing the collection of personally identifiable information.
- Algorithmic Transparency and Explainability: Enhancing the transparency and explainability of AI algorithms used in cybersecurity can help address ethical concerns related to automated decision-making. Providing clear explanations of how AI algorithms reach their conclusions and allowing for human oversight and intervention can increase trust and confidence in AI-driven security measures.

5. Future Directions and Opportunities

As AI revolutionizes cybersecurity, exciting new trends and opportunities are emerging. This section explores promising avenues for the future, highlighting key areas of focus and potential challenges.

5.1. A. Emerging Trends in AI-Driven Cybersecurity

Adversarial Machine Learning: Adversarial machine learning techniques are emerging as a critical area of research in AI-driven cybersecurity, focusing on developing robust defences against adversarial attacks that target AI algorithms.

Zero Trust Security: The adoption of zero trust security architectures is gaining momentum, driven by the need to protect against sophisticated cyber threats and insider attacks. AI technologies are crucial in implementing and operationalizing zero-trust principles by continuously monitoring and verifying user and device behaviours.

5.2. Opportunities for Further Research and Development:

Explainable AI: Advancing research in explainable AI is essential for improving the transparency and interpretability of AI-driven cybersecurity systems. Researchers can enhance trust and confidence in AI algorithms among security professionals and stakeholders by developing techniques to explain AI decision-making processes in understandable terms.

Privacy-Preserving AI: Research in privacy-preserving AI aims to develop techniques and methodologies that enable AI algorithms to perform complex computations while preserving the privacy of sensitive data. Researchers can address privacy concerns and compliance requirements without sacrificing performance or accuracy by integrating privacy-preserving techniques into AI-driven security measures.

5.3. Predictions for the Future of AI and Cybersecurity Integration:

Increased Automation: Integrating AI into cybersecurity is expected to lead to increased automation of security processes, enabling organizations to detect, respond to, and mitigate cyber threats more efficiently and effectively.

AI-Powered Threat Intelligence: AI-driven threat intelligence platforms will provide organizations with actionable insights into emerging cyber threats, enabling proactive defence strategies and threat-hunting activities.

In conclusion, addressing ethical and privacy considerations, exploring emerging trends, and leveraging opportunities for further research and development are essential for realizing the full potential of AI in cybersecurity and ensuring a safer digital future.

6. Conclusion

In conclusion, the fusion of artificial intelligence (AI) with cybersecurity heralds a new era in our ongoing battle against digital threats. This study elucidated the foundational concepts of AI and its applications in cybersecurity, illustrating how machine learning, natural language processing, and other AI techniques are revolutionizing threat detection, incident response, and vulnerability analysis.

This study further delved into the intricate landscape of vulnerability analysis, juxtaposing traditional approaches with AI-driven methodologies. By automating vulnerability scanning, prioritizing critical threats, and continuously learning from historical data, AI-powered solutions address longstanding challenges of scalability, accuracy, and adaptability in cybersecurity. Furthermore, the discourse on incident response underscored the pivotal role of AI-driven automation in expediting threat mitigation and enhancing overall security posture. Through accelerated incident response times, reduced human error, and continuous learning capabilities, AI-equipped platforms are reshaping the frontline defences against cyber adversaries.

However, the ethical and privacy considerations underscore the imperative of responsible AI deployment. As AI becomes increasingly intertwined with cybersecurity operations, organizations must navigate the ethical complexities of automated decision-making, guard against the weaponization of AI, and uphold privacy rights in the face of pervasive surveillance and data collection.

Looking forward, the future of AI in cybersecurity holds immense promise. Emerging trends such as adversarial machine learning and zero trust security present fertile ground for further research and development, offering opportunities to fortify our digital resilience against evolving threats. Moreover, ensuring robust ethical frameworks and privacy safeguards will be paramount as AI continues to evolve. Transparency, accountability, and fairness must guide the development and deployment of AI-driven cybersecurity solutions to maintain trust and confidence in these technologies.

Ultimately, the journey towards a safer digital landscape requires a concerted effort to balance technological innovation with ethical principles and privacy rights. By fostering collaboration among stakeholders, promoting transparency, and adhering to ethical guidelines, we can harness the full potential of AI to build a more secure, resilient, and trustworthy digital ecosystem for generations to come.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] E. Lacka and T. C. Wong, "Examining the Impact of Digital Technologies on Students' Higher Education Outcomes: The Case of the Virtual Learning Environment and Social Media," *Stud. High. Educ.*, 2019, doi: 10.1080/03075079.2019.1698533.

[2] A. Urbinati, D. Chiaroni, V. Chiesa, and F. Frattini, "The Role of Digital Technologies in Open Innovation Processes: An Exploratory Multiple Case Study Analysis," *R D Manag.*, 2018, doi: 10.1111/radm.12313.

[3] J. M. Couretas, "Cyber Analysis and Targeting," *An Introd. to Cyber Anal. Target.*, pp. 1–12, 2022, doi: 10.1007/978-3-030-88559-5_1.

[4] C. C. Zou, W. Gong, and D. Towsley, "Code red worm propagation modeling and analysis," *Proc. ACM Conf. Comput. Commun. Secur.*, pp. 138–147, 2002, doi: 10.1145/586110.586130.

[5] T. Mahjabin, Y. Xiao, G. Sun, and W. Jiang, "A survey of distributed denial-of-service attack, prevention, and mitigation techniques," *Int. J. Distrib. Sens. Networks*, vol. 13, no. 12, Dec. 2017, doi: 10.1177/1550147717741463/ASSET/IMAGES/LARGE/10.1177_1550147717741463-FIG13.JPEG.

[6] T. Palys and C. Atchison, "Qualitative Research in the Digital Era: Obstacles and Opportunities," *Int. J. Qual. Methods*, 2012, doi: 10.1177/160940691201100404.

[7] R. Collier, "NHS ransomware attack spreads worldwide," *C. Can. Med. Assoc. J.*, vol. 189, no. 22, p. E786, Jun. 2017, doi: 10.1503/CMAJ.1095434.

[8] Andy Greenberg, "The Untold Story of NotPetya, the Most Devastating Cyberattack in History," *Wired*, 2018. https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/ (accessed Jan. 22, 2024).

[9] S. Ghafur, S. Kristensen, K. Honeyford, G. Martin, A. Darzi, and P. Aylin, "A retrospective impact analysis of the WannaCry cyberattack on the NHS," *npj Digit. Med. 2019 21*, vol. 2, no. 1, pp. 1–7, Oct. 2019, doi: 10.1038/s41746-019-0161-6.

[10] L. Agostini and A. Nosella, "Industry 4.0 and Business Models: A Bibliometric Literature Review," *Bus. Process Manag. J.*, 2021, doi: 10.1108/bpmj-03-2021-0133.

[11] I. Iyamu *et al.*, "Defining Digital Public Health and the Role of Digitization, Digitalization, and Digital Transformation: Scoping Review," *Jmir Public Heal. Surveill.*, 2021, doi: 10.2196/30399.

[12] Y. Commandré, C. Macombe, and S. Mignon, "Implications for Agricultural Producers of Using Blockchain for Food Transparency, Study of 4 Food Chains by Cumulative Approach," *Sustainability*, 2021, doi: 10.3390/su13179843.

[13] T. Ahmed *et al.*, "Digital Health and Inequalities in Access to Health Services in Bangladesh: Mixed Methods Study (Preprint)," 2019, doi: 10.2196/preprints.16473.

[14] N. Marres, "On Some Uses and Abuses of Topology in the Social Analysis of Technology (Or the Problem With Smart Meters)," *Theory Cult. Soc.*, 2012, doi: 10.1177/0263276412454460.

[15] R. Kaur, D. Gabrijelčič, and T. Klobučar, "Artificial intelligence for cybersecurity: Literature review and future research directions," *Inf. Fusion*, vol. 97, p. 101804, Sep. 2023, doi: 10.1016/J.INFFUS.2023.101804.

[16] A. Singh and B. B. Gupta, "Distributed Denial-of-Service (DDoS) Attacks and Defense Mechanisms in Various Web-Enabled Computing Platforms: Issues, Challenges, and Future Research Directions," *Int. J. Semant. Web Inf. Syst.*, vol. 18, no. 1, 2022, doi: 10.4018/IJSWIS.297143.

[17] A. Borode and P. Olubambi, "Optimization of artificial intelligence models and response surface methodology for predicting viscosity and relative viscosity of GNP-alumina hybrid nanofluid: incorporating the effects of mixing ratio and temperature," *J. Supercomput.*, 2023, doi: 10.1007/s11227-023-05652-y.

[18] F. Pesapane, C. Volonté, M. Codari, and F. Sardanelli, "Artificial Intelligence as a Medical Device in Radiology: Ethical and Regulatory Issues in Europe and the United States," *Insights Imaging*, 2018, doi: 10.1007/s13244-018-0645-y.

[19] N. Capuano, G. Fenza, V. Loia, and C. Stanzione, "Explainable Artificial Intelligence in CyberSecurity: A Survey," *Ieee Access*, 2022, doi: 10.1109/access.2022.3204171.

[20] A. d. Hond *et al.*, "Guidelines and Quality Criteria for Artificial Intelligence-Based Prediction Models in Healthcare: A Scoping Review," *NPJ Digit. Med.*, 2022, doi: 10.1038/s41746-021-00549-7.

[21] G. Nebbione and M. C. Calzarossa, "A Methodological Framework for AI-Assisted Security Assessments of Active Directory Environments," *Ieee Access*, 2023, doi: 10.1109/access.2023.3244490.

[22] H. A. Kholidy, "Multi-Layer Attack Graph Analysis in the 5G Edge Network Using a Dynamic Hexagonal Fuzzy Method," *Sensors*, 2021, doi: 10.3390/s22010009.

[23] L. L. Dhirani, E. Armstrong, and T. Newe, "Industrial IoT, Cyber Threats, and Standards Landscape: Evaluation and Roadmap," *Sensors*, 2021, doi: 10.3390/s21113901.

[24] I. Zografopoulos, J. Ospina, X. Liu, and C. Konstantinou, "Cyber-Physical Energy Systems Security: Threat Modeling, Risk Assessment, Resources, Metrics, and Case Studies," 2021, doi: 10.48550/arxiv.2101.10198.

[25] M. Markevych and M. Dawson, "A Review of Enhancing Intrusion Detection Systems for Cybersecurity Using Artificial Intelligence (AI)," *Int. Conf. KNOWLEDGE-BASED Organ.*, vol. 29, p. 2023, Jul. 2023, doi: 10.2478/kbo-2023-0072.

[26] P. Dini, A. Elhanashi, A. Begni, S. Saponara, Q. Z.-A. Sciences, and undefined 2023, "Overview on Intrusion Detection Systems Design Exploiting Machine Learning for Networking Cybersecurity," *mdpi.comP Dini, A Elhanashi, A Begni, S Sapon. Q Zheng, K GasmiApplied Sci. 2023•mdpi.com*, Accessed: Feb. 20, 2024. [Online]. Available: https://www.mdpi.com/2076-3417/13/13/7507.

[27] S. Moisset, "How Security Analysts Can Use AI in Cybersecurity," 2023. https://www.freecodecamp.org/news/how-to-use-artificial-intelligence-in-cybersecurity/.

[28] N. Jha and S. Ramaprabhu, "Thermal conductivity studies of metal dispersed multiwalled carbon nanotubes in water and ethylene glycol based nanofluids," *J. Appl. Phys.*, vol. 106, no. 8, p. 084317, Oct. 2009, doi: 10.1063/1.3240307.

[29] S. H. Javed, M. Bin Ahmad, M. Asif, S. H. Almotiri, K. Masood, and M. A. Al Ghamdi, "An Intelligent System to Detect Advanced Persistent Threats in Industrial Internet of Things (I-IoT)," *Electron. 2022, Vol. 11, Page 742*, vol. 11, no. 5, p. 742, Feb. 2022, doi: 10.3390/ELECTRONICS11050742.

[30] B. Alaeddine, N. Hmina, and H. Chaoui, "Parallel processing using big data and machine learning techniques for intrusion detection," *IAES Int. J. Artif. Intell.*, vol. 9, p. 553, Sep. 2020, doi: 10.11591/ijai.v9.i3.pp553-560.

[31] A.-T. Costin, D. Zinca, and V. Dobrota, "A Real-Time Streaming System for Customized Network Traffic Capture," *Sensors*, vol. 23, no. 14. 2023, doi: 10.3390/s23146467.

[32] M. Paolanti and E. Frontoni, "Multidisciplinary Pattern Recognition applications: A review," *Comput. Sci. Rev.*, vol. 37, Aug. 2020, doi: 10.1016/J.COSREV.2020.100276.

[33] P. Dini and S. Saponara, "Analysis, Design, and Comparison of Machine-Learning Techniques for Networking Intrusion Detection," *Designs*, vol. 5, no. 1, pp. 1–22, 2021, doi: 10.3390/DESIGNS5010009.

[34] A. Borode, T. Tshephe, P. Olubambi, M. Sharifpur, and J. Meyer, "Experimental study and ANFIS modelling of the thermophysical properties and efficacy of GNP-Al2O3 hybrid nanofluids of different concentrations and temperatures," *SN Appl. Sci.*, vol. 5, no. 12, p. 337, 2023, doi: 10.1007/s42452-023-05574-7.

[35] J. Delua, "Supervised vs. Unsupervised Learning: What's the Difference?," *IBM Blog*, 2021. https://www.ibm.com/blog/supervised-vs-unsupervised-learning/ (accessed Feb. 20, 2024).

[36] A. Srinivasan, "Reinforcement Learning: Advancements, Limitations, and Real-world Applications," *INTERANTIONAL J. Sci. Res. Eng. Manag.*, vol. 07, no. 08, Aug. 2023, doi: 10.55041/IJSREM25118.

[37] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/S13244-018-0639-9/FIGURES/15.

[38] C. Thomas, “Recurrent Neural Networks and Natural Language Processing. | by Christopher Thomas BSc Hons. MIAP | Towards Data Science,” *Towards Data Science*, 2019. https://towardsdatascience.com/recurrent-neural-networks-and-natural-language-processing-73af640c2aa1 (accessed Feb. 20, 2024).

[39] “What are Recurrent Neural Networks? | IBM.” https://www.ibm.com/topics/recurrent-neural-networks (accessed Feb. 20, 2024).

[40] P. Salehi, A. Chalechale, and M. Taghizadeh, *Generative Adversarial Networks (GANs): An Overview of Theoretical Model, Evaluation Metrics, and Recent Developments*. 2020.

[41] S. Kim, C. Hwang, and T. Lee, “Anomaly Based Unknown Intrusion Detection in Endpoint Environments,” *Electron. 2020, Vol. 9, Page 1022*, vol. 9, no. 6, p. 1022, Jun. 2020, doi: 10.3390/ELECTRONICS9061022.

[42] P. Hakonen, “Detecting Insider Threats Using User and Entity Behavior Analytics,” 2022, Accessed: Feb. 20, 2024. [Online]. Available: http://www.theseus.fi/handle/10024/786079.

[43] L. Universitet and +++ S.------L., “Anomaly detection for automated security log analysis : Comparison of existing techniques and tools,” 2021, Accessed: Feb. 20, 2024. [Online]. Available: https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-177728.

[44] M. Kern, F. Skopik, M. Landauer, and E. Weippl, “Strategic selection of data sources for cyber attack detection in enterprise networks: A survey and approach,” *Proc. ACM Symp. Appl. Comput.*, pp. 1656–1665, Apr. 2022, doi: 10.1145/3477314.3507022.

[45] P. Nanray, *AI-Driven Predictive Analysis in Cybersecurity: Focus on Phishing and Malware Detection*. 2023.

[46] S. Naeem, Aqib Ali, Sania Anam, and Muhammad Munawar Ahmed, “Machine Learning for Intrusion Detection in Cyber Security: Applications, Challenges, and Recommendations,” *Innov. Comput. Rev.*, vol. 2, no. 2, Dec. 2022, doi: 10.32350/ICR.0202.03.

[47] T. Wischmeyer, “Artificial intelligence and transparency: Opening the black box,” *Regul. Artif. Intell.*, pp. 75–101, Jan. 2019, doi: 10.1007/978-3-030-32361-5_4.

[48] M. Di Biase, M. Bruntink, and A. Bacchelli, “A security perspective on code review: The case of chromium,” *Proc. - 2016 IEEE 16th Int. Work. Conf. Source Code Anal. Manip. SCAM 2016*, pp. 21–30, Dec. 2016, doi: 10.1109/SCAM.2016.30.

[49] C. Dietrich, K. Krombholz, K. Borgolte, and T. Fiebig, *Investigating System Operators’ Perspective on Security Misconfigurations*. 2018.

[50] Ö. Aslan, S. S. Aktuğ, M. Ozkan-Okay, A. A. Yilmaz, and E. Akin, “A Comprehensive Review of Cyber Security Vulnerabilities, Threats, Attacks, and Solutions,” *Electron.*, vol. 12, no. 6, Mar. 2023, doi: 10.3390/ELECTRONICS12061333.

[51] “Why Traditional Vulnerability Management isn’t Cutting it Anymore,” *Cybersecurity Insiders*. https://www.cybersecurity-insiders.com/why-traditional-vulnerability-management-isnt-cutting-it-anymore/ (accessed Feb. 21, 2024).

[52] J. Kupsch, B. M. M. I. S. T. (MIST), and undefined 2009, “Manual vs. automated vulnerability assessment: A case study,” *academia.edu*, Accessed: Feb. 20, 2024. [Online]. Available: https://www.academia.edu/download/30662081/MIST2009.pdf#page=87.

[53] S. Khan and S. Parkinson, “Review into State of the Art of Vulnerability Assessment using Artificial Intelligence,” pp. 3–32, 2018, doi: 10.1007/978-3-319-92624-7_1.

[54] X. Tian and D. Tang, “A distributed vulnerability scanning on machine learning,” *Proc. - 2019 6th Int. Conf. Inf. Sci. Control Eng. ICISCE 2019*, pp. 32–35, Dec. 2019, doi: 10.1109/ICISCE48695.2019.00016.

[55] “AI in Cyber Security: Use Cases, Benefits, and Challenges - Eastgate Software.” https://eastgate-software.com/ai-in-cyber-security-use-cases-benefits-and-challenges/ (accessed Feb. 21, 2024).

[56] “Security QRadar | IBM.” https://www.ibm.com/qradar (accessed Feb. 21, 2024).

[57] S. Kumar, U. Gupta, A. K. Singh, and A. K. Singh, “Artificial Intelligence: Revolutionizing Cyber Security in the Digital Era,” *J. Comput. Mech. Manag.*, vol. 2, no. 3, pp. 31–42, Aug. 2023, doi: 10.57159/GADL.JCMM.2.3.23064.

[58] “Qualys Cloud Detection and Response | Qualys.” https://www.qualys.com/apps/cloud-detection-response/ (accessed Feb. 21, 2024).

[59] "Tenable Security Center (Formerly Tenable.sc) | Tenable®." https://www.tenable.com/products/tenable-sc (accessed Feb. 21, 2024).

[60] P. Breda, R. Markova, A. F. Abdin, N. P. Mantı, A. Carlo, and D. Jha, "An extended review on cyber vulnerabilities of AI technologies in space applications: Technological challenges and international governance of AI," *J. Sp. Saf. Eng.*, vol. 10, no. 4, pp. 447–458, Dec. 2023, doi: 10.1016/J.JSSE.2023.08.003.

[61] "Cybersecurity Framework | NIST," Apr. 2018, doi: 10.6028/NIST.CSWP.04162018.

[62] "EDR vs. SIEM - Check Point Software." https://www.checkpoint.com/cyber-hub/threat-prevention/what-is-endpoint-detection-and-response/edr-vs-siem/ (accessed Feb. 21, 2024).

[63] "What is Incident Response? Process, Frameworks, and Tools." https://www.bluevoyant.com/knowledge-center/what-is-incident-response-process-frameworks-and-tools (accessed Feb. 21, 2024).

[64] S. Tatineni, "AI-Infused Threat Detection and Incident Response in Cloud Security," *Int. J. Sci. Res.*, vol. 12, no. 11, pp. 998–1004, Nov. 2023, doi: 10.21275/SR231113063646.

[65] A. ZAMFIROIU and R. C. SHARMA, "Cybersecurity Management for Incident Response," *Rom. Cyber Secur. J.*, vol. 4, no. 1, pp. 69–75, May 2022, doi: 10.54851/V4I1Y202208.

[66] E. Iturbe, E. Rios, A. Rego, and N. Toledo, "Artificial Intelligence for next generation cybersecurity: The AI4CYBER framework," *ACM Int. Conf. Proceeding Ser.*, Aug. 2023, doi: 10.1145/3600160.3605051.

[67] Syed Khurram Hassan and Asif Ibrahim, "The role of Artificial Intelligence in Cyber Security and Incident Response," *Int. J. Electron. Crime Investig.*, vol. 7, no. 2, Jul. 2023, doi: 10.54692/IJECI.2023.0702154.

[68] "AI in Cybersecurity: Incident Response Automation Opportunities." https://www.sisainfosec.com/blogs/ai-in-cybersecurity-incident-response-automation-opportunities/ (accessed Feb. 21, 2024).

[69] R. Sabillon, "Cybersecurity Incident Response and Management," *Res. Anthol. Bus. Asp. Cybersecurity*, pp. 611–620, Sep. 2021, doi: 10.4018/978-1-6684-3698-1.CH028.

[70] D. Schlette, M. Caselli, and G. Pernul, "A Comparative Study on Cyber Threat Intelligence: The Security Incident Response Perspective," *IEEE Commun. Surv. Tutorials*, vol. 23, no. 4, pp. 2525–2556, 2021, doi: 10.1109/COMST.2021.3117338.

[71] "What is Incident Response? Definition & Examples | Darktrace." https://darktrace.com/cyber-ai-glossary/incident-response (accessed Feb. 21, 2024).

[72] "Incident Response Policy: A Quick Guide." https://www.cynet.com/incident-response/incident-response-policy-a-quick-guide/ (accessed Feb. 21, 2024).

[73] "Cortex XSOAR: Security Orchestration and Automation - Palo Alto Networks." https://www.paloaltonetworks.com/cortex/cortex-xsoar (accessed Feb. 21, 2024).

[74] U. Tariq, I. Ahmed, A. K. Bashir, and K. Shaukat, "A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review," *Sensors 2023, Vol. 23, Page 4117*, vol. 23, no. 8, p. 4117, Apr. 2023, doi: 10.3390/S23084117.

[75] D. Pedreschi, F. Giannotti, R. Guidotti, A. Monreale, S. Ruggieri, and F. Turini, "Meaningful explanations of black box ai decision systems," *33rd AAAI Conf. Artif. Intell. AAAI 2019, 31st Innov. Appl. Artif. Intell. Conf. IAAI 2019 9th AAAI Symp. Educ. Adv. Artif. Intell. EAAI 2019*, pp. 9780–9784, 2019, doi: 10.1609/AAAI.V33I01.33019780.

[76] S. Al-Mansoori and M. Ben Salem, "The Role of Artificial Intelligence and Machine Learning in Shaping the Future of Cybersecurity: Trends, Applications, and Ethical Considerations," *Int. J. Soc. Anal.*, vol. 8, no. 9, pp. 1–16, Sep. 2023, doi: 10.1057/s41284-021-00321-2.

[77] G. L. Sanclemente, "Reliability: understanding cognitive human bias in artificial intelligence for national security and intelligence analysis," *Secur. J.*, vol. 35, no. 4, pp. 1328–1348, Dec. 2022, doi: 10.1057/S41284-021-00321-2/TABLES/1.

[78] N. Joseph, *The Role of Artificial Intelligence in Predictive Cybersecurity Analytics*. 2023.