applied sciences

MDPI

Article

Developing a Cybersecurity Training Environment through the Integration of OpenAI and AWS

William Villegas-Ch *, Jaime Govea and Iván Ortiz-Garces

Escuela de Ingeniería en Ciberseguridad, Facultad de Ingenierías en Ciencias Aplicadas, Universidad de Las Américas, Quito 170125, Ecuador; jaimealejandro.govea@udla.edu.ec (J.G.)
* Correspondence: william.villegas@udla.edu.ec; Tel.: +593-098-136-4068

Abstract: Cybersecurity is a critical concern in today's digital age, where organizations face an ever-evolving cyber threat landscape. This study explores the potential of leveraging artificial intelligence and Amazon Web Services to improve cybersecurity practices. Combining the capabilities of OpenAI's GPT-3 and DALL-E models with Amazon Web Services infrastructure aims to improve threat detection, generate high-quality synthetic training data, and optimize resource utilization. This work begins by demonstrating the ability of artificial intelligence to create synthetic cybersecurity data that simulates real-world threats. These data are essential for training threat detection systems and strengthening an organization's resilience against cyberattacks. While our research shows the promising potential of artificial intelligence and Amazon Web Services in cybersecurity, it is essential to recognize the limitations. Continued research and refinement of AI models are needed to address increasingly sophisticated threats. Additionally, ethical and privacy considerations must be addressed when employing AI in cybersecurity practices. The results support the notion that this collaboration can revolutionize how organizations address cyber challenges, delivering greater efficiency, speed, and accuracy in threat detection and mitigation.

Keywords: cybersecurity; artificial intelligence; Amazon Web Services

Citation: Villegas-Ch, W.; Govea, J.; Ortiz-Garces, I. Developing a Cybersecurity Training Environment through the Integration of OpenAI and AWS. *Appl. Sci.* **2024**, *14*, 679. https://doi.org/10.3390/app14020679

Academic Editor: David Megías

Received: 16 November 2023
Revised: 7 January 2024
Accepted: 8 January 2024
Published: 13 January 2024

1. Introduction

Cybersecurity is a constantly evolving field that faces increasingly sophisticated challenges and ever-expanding threats. Information security and protection against cyberattacks have become critical priorities for governments, businesses, and users in a digitally interconnected world. The increasing complexity of cyber threats has led to the need for advanced and practical solutions that can anticipate, detect, and mitigate risks in real time [1]. For this reason, artificial intelligence (AI) has established itself as a fundamental tool to strengthen cybersecurity. Despite technological advances, the continued adaptation of security strategies in response to emerging threats remains a crucial challenge, highlighting the importance of bridging theory and practice.

This work presents an approach in the field of cybersecurity, combining AI with Amazon Web Services (AWS) cloud resources to create a training environment [2]. This collaboration represents a significant milestone and a step forward in the ongoing effort to strengthen cybersecurity through innovation and the practical application of new technologies [3].

This work lies in converging two technologies: AI and the cloud. AI, represented by OpenAI's GPT-3 and DALL-E v2 models, is recognized for its ability to understand and generate natural language and visually realistic content from textual descriptions [4]. These capabilities will be applied in cybersecurity to create training data, threat simulations, and vulnerability detection.

On the other hand, AWS is a cloud services platform that provides scalable and flexible resources for various applications. Integrating OpenAI into AWS allows the leveraging of a

world-class cloud infrastructure to deploy and scale advanced cybersecurity solutions [5]. This represents a significant advantage in an ever-changing threat environment. This approach aligns with the current literature, applying cybersecurity theories in a practical environment to evaluate their effectiveness in natural conditions.

This work explores how combining AI and the cloud can transform cybersecurity, improving efficiency, speed, and accuracy in threat detection and mitigation. In addition, it seeks to analyze how this collaboration can reduce the manual workload, optimize resource use, and improve user satisfaction in cybersecurity [6]. Experiments and tests are performed to achieve these objectives in a specifically designed training environment. The ability of AI to generate quality synthetic data that simulates real cyber threats is evaluated. Additionally, the efficiency of AI is measured and compared to traditional methods in terms of data generation speed, resource savings, and scalability.

The collaboration between OpenAI and AWS offers an innovative approach that can potentially improve cybersecurity at all levels, from early threat detection to protecting critical infrastructure [7,8]. The combination of AI's data-generating power and the AWS cloud's scalability promises to revolutionize how organizations address cyber challenges in an ever-changing digital environment.

This work is structured into several key sections, including a literature review, the methodology, the presentation of results, a discussion, and conclusions. The literature review will analyze previous research related to the application of AI in cybersecurity and identify gaps and opportunities in the field. The methodology will detail the experimental design and configuration of the training environment. The results will present the findings obtained through specific tests and measurements. The discussion will further analyze the results in the context of the existing literature and implications for cybersecurity. Finally, the conclusions will summarize the achievements and limitations of the work, and possible directions for future research will be proposed.

2. Materials and Methods

The materials and methods section presents the foundation of this work and accurately describes the tools, technologies, and processes used to establish a robust and detailed cybersecurity testing environment. This part is essential since it describes our system's technical specifications and configuration, justifies each component's selection, and explains how these elements are integrated to achieve the stated objectives. The clarity and rigor with which the materials and methods are presented are crucial, as they guarantee the reproducibility of the study.

2.1. Description of the Problem

In the current technology landscape, cybersecurity has become a vitally important issue. Organizations of all sizes face growing, varied, and increasingly sophisticated cyber threats. These threats range from ransomware and phishing attacks to data breaches and attacks targeting critical infrastructure [9]. As the world becomes more interconnected and dependent on digital solutions, protecting digital assets has become critical. However, cybersecurity not only involves the implementation of advanced technological solutions but also the adequate training and preparation of professionals in charge of managing and responding to these threats [10].

One of the main difficulties in preparing cybersecurity professionals is the gap between academic theory and practice in the field. Traditional educational programs often fail to accurately simulate the dynamic and complex scenarios that professionals face in real-world situations. Additionally, cybersecurity constantly evolves, with new threats and technologies emerging regularly [11]. This makes it challenging to keep training curricula and methods up to date. Another critical issue is the shortage of qualified cybersecurity professionals. According to several industry reports, cybersecurity talent is lacking, leaving many organizations vulnerable to cyberattacks [12]. Efficient and up-to-date cybersecurity

training is crucial to closing this skill gap and preparing a workforce capable of effectively addressing cyber threats.

Integrating advanced technologies such as AI from OpenAI and cloud infrastructure from AWS presents a unique opportunity to address these challenges. AI can be crucial in simulating cybersecurity threats and generating realistic and dynamic training scenarios. Using AI technologies such as natural language processing and machine learning, it is possible to create simulations that closely reflect the actual situations cybersecurity professionals will face [13].

On the other hand, AWS offers a robust and scalable infrastructure that can support deploying complex and high-demand training environments. Combining AWS with AI creates a virtual training environment where professionals can practice and improve their skills in a controlled but realistic environment [14]. This integration also facilitates the continuous updating and adaptation of training content, ensuring that scenarios and simulations reflect the latest trends and threats in cybersecurity.

This work proposes the development of a cybersecurity training environment through the synergistic integration of OpenAI artificial intelligence and the AWS cloud infrastructure. This innovative solution seeks to overcome the limitations of conventional cybersecurity teaching and training methods, providing a platform that accurately and dynamically reflects the current cyber threat landscape. The integration of OpenAI with AWS allows the generation of a simulation environment that imitates cyberattack and defense scenarios in real time. Using OpenAI's advanced machine learning and natural language processing capabilities, we can create detailed and realistic simulations of a wide range of cyber threats. This includes phishing and ransomware attacks, security breaches, and complex cyberattacks. These simulations will allow trainees to experience and respond to cybersecurity situations in a controlled yet realistic environment.

On the other hand, the AWS platform offers the scalability, security, and robust infrastructure necessary to implement and manage these simulations. This ensures optimal performance and resource availability and the integrity and protection of data within the training environment. The development of the environment aims to offer a learning experience that goes beyond theory, immersing users in cybersecurity scenarios that imitate the challenges and decisions they will face in natural environments.

By using AI to analyze trends and new types of attacks, simulations can adapt and evolve to reflect the current threat landscape [15]. Developing practical skills is encouraged so that users of the training environment develop practical skills in identifying, preventing, and responding to cybersecurity incidents, improving their ability to act effectively in real situations. By simulating different types of cyberattacks, the environment allows organizations to evaluate and improve their incident response strategies and protocols [16].

To provide a more concrete and applied view of how OpenAI tools are used in the proposed cybersecurity environment, the following examples and case studies are presented:

- Phishing Attack Simulation with GPT-3: Using GPT-3, a detailed scenario of a phishing attack was developed. In this scenario, GPT-3 generated a series of email communications that closely mimicked the tactics used by real attackers. This includes creating compelling content and simulating attackers' responses to user interactions. This simulation allowed trainees to experience and respond to a phishing attack in a safe environment, developing critical skills to recognize and mitigate such threats.
- Security Breach Visualization with DALL-E: In another case, DALL-E was used to visualize the impact of a security breach on a corporate network. From a textual description of the breach, DALL-E generated images depicting the progression of the attack, including initial access, network propagation, and data extraction. These visualizations helped participants better understand the multifaceted nature of security breaches and the importance of quick and effective responses.

These examples present the practical application of GPT-3 and DALL-E in cybersecurity. By providing realistic scenarios and detailed visualizations, these tools significantly

improve trainees' ability to understand and respond to cyber threats, preparing them for real-world challenges.

2.2. Operation of the Training Environment

Each participant interacts through a personalized avatar. These avatars represent users in the virtual environment and are the primary means of interacting with the system, simulation scenarios, and other participants. Through their avatars, users immerse themselves in various simulated scenarios that replicate real-world cybersecurity situations. These can range from managing a phishing attack to responding to a large-scale security breach. Using OpenAI, the environment simulates realistic interactions and challenges [17]. Avatars can interact with elements of the environment, such as accessing virtual systems, analyzing security data, and making decisions to mitigate threats. The system uses artificial intelligence algorithms to adapt scenarios to users' skills and experience levels. This allows each training session to be relevant and challenging for the individual, regardless of their skill level [18].

In some scenarios, avatars can work as a team, allowing user collaboration. This is essential to simulate incident response in a real team environment, where communication and collaboration are crucial. The system provides real-time feedback as users interact with the environment through their avatars. This includes evaluations of decisions made and advice on improving responses to threats. Using AWS infrastructure, the environment can scale to accommodate multiple users and complex scenarios while maintaining data security and integrity.

After each training session, a detailed analysis of the user's performance is conducted. This analysis helps identify areas for improvement and adjust future training scenarios to address these deficiencies [19]. The environment is regularly updated with new scenarios, reflecting the latest trends and threats in cybersecurity. This ensures that training remains relevant and challenging [20].

Figure 1 presents the various stages involved in the cybersecurity training environment that we developed, integrating OpenAI artificial intelligence with the AWS infrastructure. Each step represents a critical stage, beginning with immersion through avatars and moving through several critical processes such as simulation scenarios, realistic interaction, and adaptive learning [21]. Arrows indicate how the process flows from one stage to another, highlighting collaboration and role-playing, real-time feedback, and integration with AWS. The diagram culminates with the post-training analysis and scenario update, ensuring a dynamic and updated approach to preparing cybersecurity professionals.

2.3. Review of Similar Works

In the emerging field of cybersecurity, various research studies and developments have addressed the challenge of training and preparing professionals to face cyber threats. Several studies have explored using simulations and virtual environments for cybersecurity training. For example, works like Cyber Range and other simulation environments offer platforms where participants can practice cybersecurity skills in a controlled environment [22]. These environments are typically highly technical and focus on simulating networks and systems for cyberattack and defense practices. However, they often lack advanced artificial intelligence elements that could provide more dynamic and adaptive scenarios. This work seeks to fill this gap by integrating OpenAI AI to generate more realistic and variable simulations that better reflect the ever-changing threat landscape [23].

The integration of AI into cybersecurity has been a topic of growing interest. Recent research has demonstrated the potential of AI to improve threat detection and incident response. However, most of these works focus on the application of AI for operating systems and rather than on human training [24]. Our initiative differentiates itself by using AI not only as a cyber defense tool but also to improve the training and preparation of cybersecurity professionals.

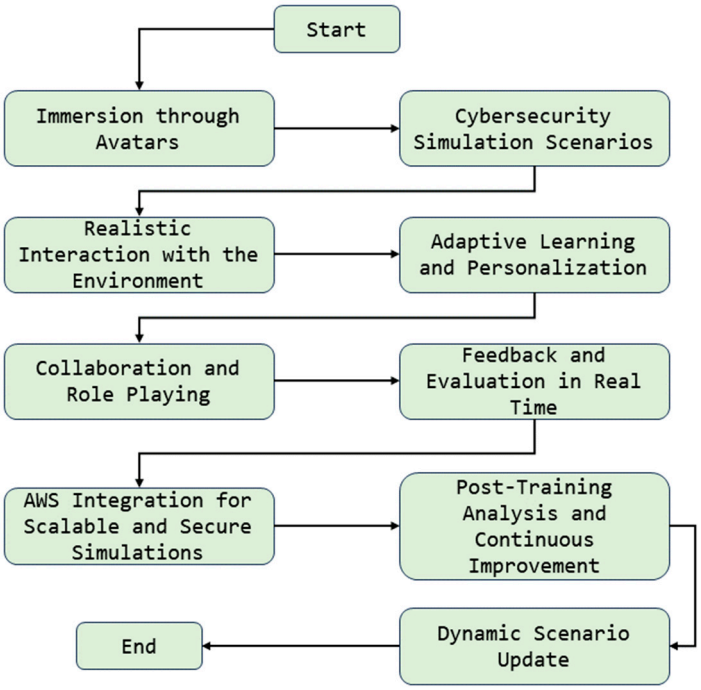

Figure 1. Stages of the cybersecurity training environment (source: authors' elaboration).

The use of avatars and virtual reality has begun to be explored in cybersecurity training, although it is still in its early stages. Some projects have implemented VR environments to offer a more immersive experience. However, these initiatives often do not fully integrate AI to personalize the learning experience or to simulate advanced interactions [25]. This work moves in this direction by combining avatars in a VR environment with AI, enabling a richer and more personalized learning experience [26].

Despite significant advances in these fields, notable gaps still exist. One of them is the effective integration of AI to create real-time dynamic and adaptive training scenarios. Additionally, the AI-based personalization of learning to suit different skill levels and roles within cybersecurity is an underexplored area that our project seeks to address [27]. Another area of opportunity is the combination of artificial intelligence and virtual reality technologies to simulate cyber threats and foster soft skills such as decision making, collaboration, and communication in stressful situations. This holistic approach is crucial in preparing cybersecurity professionals, where technical skills must be accompanied by solid decision-making and teamwork capabilities [28].

This work is distinguished by integrating AI from OpenAI to create more realistic and variable simulations that better reflect the ever-changing threat landscape. Additionally, this initiative highlights the use of AI to personalize the learning experience, adjusting to different skill levels and roles in cybersecurity. Our approach also includes using artificial intelligence and virtual reality technologies for threat simulation and developing essential decision-making and teamwork skills in cybersecurity.

2.4. Description of the Development Environment

The cybersecurity training development environment is based on integrating AWS services and OpenAI tools, providing a robust and versatile infrastructure. Amazon (Amazon Web Services, Inc. (AWS); Seattle; EE. UU.) EC2 (Elastic Compute Cloud) is used to deploy scalable and customizable virtual servers, essential for running complex simulations and AI models [29,30]. For data storage and retrieval, we turn to Amazon S3 (Simple Storage Service), which allows us to handle the large volumes of data generated by the simulations and host training resources.

Additionally, AWS Lambda plays a crucial role in our environment, allowing code to be executed in response to specific events without the need to manage servers, which optimizes the efficiency and scalability of the system. Security and privacy are handled through Amazon VPC (Virtual Private Cloud), which allows us to launch resources in a defined virtual network, and AWS IAM (Identity and Access Management), which ensures secure access to AWS services and resources [31].

Regarding OpenAI tools, GPT-3 (Generative Pretrained Transformer 3) is used for its advanced natural language processing capacity, essential for generating realistic cybersecurity scenarios and interactively dialogue with users. Additionally, DALL-E is innovatively used to create data visualizations and graphical representations, making it easier to understand complex cybersecurity concepts [32]. The synergy between AWS's scalable and secure infrastructure and OpenAI's language processing and content generation capabilities form the foundation of our development environment. This combination supports an advanced and highly interactive cybersecurity training environment and offers an unprecedented educational experience in this field.

Figure 2 illustrates the architecture of the cybersecurity training environment, highlighting the integration of AWS services and OpenAI tools. At the base, Amazon EC2 provides simulation processing power, while Amazon S3 manages data storage. AWS Lambda is essential in event-driven code execution, and Amazon VPC ensures a secure and controlled network environment [33]. On the other hand, AWS IAM provides secure access management and authentication. Regarding the integration of OpenAI, GPT-3 generates dynamic cybersecurity scenarios and interacts interactively, while DALL-E creates visualizations and graphical representations. These AI tools connect with AWS services and each other, ensuring a consistent and efficient flow of operations, from data processing to generating content in the simulation environment. The figure reflects how each component interconnects to form a cohesive system, providing an advanced and highly interactive cybersecurity training environment.

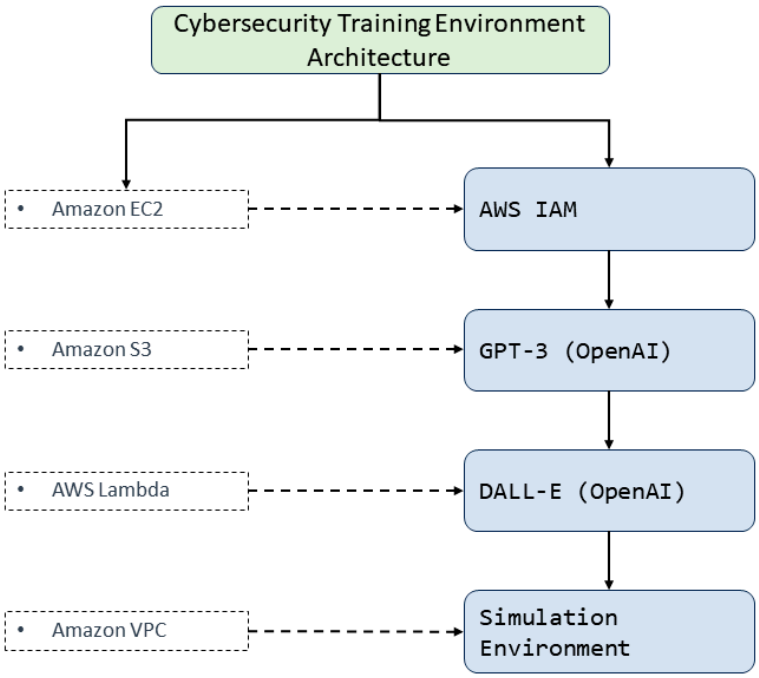

Figure 2. Architecture of the integrated cybersecurity training environment (source: authors' elaboration).

2.5. Test Environment Configuration

The test environment setup is a strategic combination of specific hardware and software, along with robust network configuration and security measures. Regarding hardware, the test environment is based on virtual servers provided by Amazon EC2, selected for

their ability to scale and handle intensive workloads. These servers are equipped with high-speed processors, a significant amount of RAM (depending on the simulation scenario's needs), and SSD storage for fast data access.

The latest stable version of the Linux operating system (Ubuntu 14.04), known for its robustness and security, is used on the software side. AI simulations and processing are carried out using the latest versions of OpenAI tools, such as GPT-3 for natural language processing and DALL-E v2 for image generation [34]. Various cybersecurity tools and monitoring software are also implemented to monitor and record activities within the testing environment [35].

The network configuration in the developed test environment is critical to ensure functionality and security. Amazon Virtual Private Cloud (VPC) creates an isolated virtual network on AWS, allowing us to define and control the network space, including selecting IP ranges, creating subnets, and configuring route tables and network gateways [36]. For network security, security groups are implemented on EC2 and network access control lists (NACLs) within VPCs to regulate incoming and outgoing traffic to servers and between subnets. Additionally, AWS Shield protects against denial of service (DDoS) attacks [37].

Regarding data security, all communications within the environment are encrypted using industry-standard security protocols. Data stored in Amazon S3 is encrypted both in transit and at rest. Additionally, IAM policies are implemented to control access to AWS services and resources precisely and securely. To ensure the integrity of the test environment, security audits are conducted, and incident response procedures are implemented to handle any detected vulnerabilities or threats quickly.

Figure 3 presents a schematic representation of the cybersecurity testing environment, highlighting the integration of critical infrastructure components and AI tools. At the environment's core, a VPC on AWS hosts Amazon EC2 instances, essential for running applications and simulating cyberattacks in a controlled, sandboxed space. Security groups protect these instances and connect to the Internet through a gateway, allowing controlled access and communication with external services [38]. Within this secure environment, OpenAI's natural language processing and image generation capabilities, provided by GPT-3 and DALL-E, respectively, are used to create and analyze dynamic content. Simulating complex interactions and generating realistic test scenarios that improve cybersecurity training methods are vital.

Additionally, the figure illustrates the presence of AWS Shield, a protection layer that strengthens the system's resilience against outages and distributed denial of service (DDoS) attacks. This scheme reflects a holistic and advanced approach to creating a cybersecurity sandbox, where security, scalability, and innovation are fundamental to developing practical skills and preparation against cyber threats.

2.6. Test Data Generation

Generating test data is a critical facet of validating and evaluating cybersecurity environments. For this work, several data types were considered, simulating a broad spectrum of cyber scenarios and threats that IT infrastructures could face.

2.6.1. Type of Data

The types of data used in the developed testing environment are multifaceted, to provide a comprehensive training and evaluation experience:

- Network Traffic: Data representing normal and abnormal network traffic are included to train participants to detect suspicious or malicious patterns that may indicate cyberattack activity.
- Security Event Logs: These logs contain detailed information on security events, ranging from failed login attempts to intrusion prevention system alerts, which are vital for forensic analysis and incident response.

- Vulnerability Data: Data are generated that describe known vulnerabilities, which helps participants understand and mitigate potential weaknesses in systems and applications [39].
- Attack Simulations: Data are created that simulate different types of cyberattacks, such as phishing, ransomware, or zero-day attacks, to prepare professionals to identify and respond to these incidents.

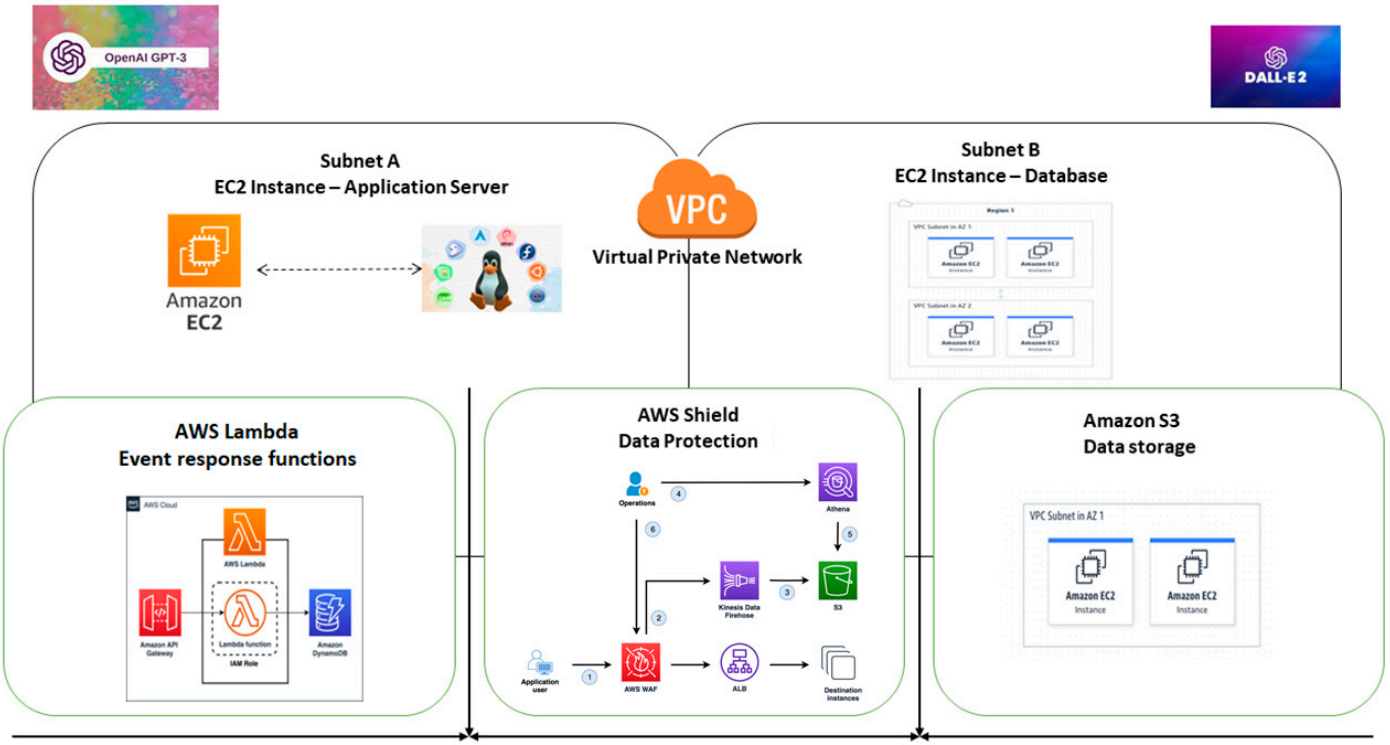

Figure 3. Cybersecurity test environment architecture integrating AWS and OpenAI (source: authors' elaboration).

2.6.2. Data Generation Methods

The generation of these data is carried out through a combination of automated simulations and data generation tools:

- Automated Simulations: Specialized tools simulate network traffic and cyberattacks, generating data indistinguishable from actual network activities. This provides a realistic scenario for incident response training and exercises.
- Synthetic Log Generators: These are used to create security event logs that mimic those natural systems generated in response to normal and suspicious activities. This includes simulating temporal patterns and correlations between events from different systems.
- Attack and Defense Scenarios: Using artificial intelligence tools such as OpenAI GPT-3, narrative attack scenarios are generated and then translated into data sequences that mimic the behavior of attackers and defenders in various cybersecurity situations [40].
- Threat Emulation: Specialized tools are implemented that emulate malicious behavior within the network to generate threat data. This includes using software that simulates network attacks and generating malware payloads in a secure environment.
- Traffic Generation Tools: Applications that generate network traffic, such as HTTP, FTP, and database traffic, create a dynamic and realistic network environment.
- Anomaly Injection: Anomalies and irregular patterns are deliberately introduced into the data to test the robustness and sensitivity of monitoring and detection tools.

2.7. User Interface

The cybersecurity system was designed with an intuitive and easy-to-use user interface. Users can access the interface through a standard web browser. Once authenticated, they are presented with a dashboard that provides an overview of the system's security.

Before using the system, one must configure security settings and preferences based the organization's needs. This includes configuring detection rules, threat response policies, and alerts. The system offers real-time monitoring capabilities to detect potential threats and anomalies. Users can view event logs, network activity, and other key metrics in real time through the user interface.

The system automatically activates the previously configured response policies if a threat is detected. Administrators can also take manual actions through the user interface to mitigate specific threats. In addition, the system allows for generating security reports that summarize security activities and events. These reports are helpful for audit review and regulatory compliance. It is essential to train the organization's personnel on the use of the system. Training sessions ensure that users fully understand the system's capabilities and know how to respond to security situations.

Figure 4 represents the user interface of the cybersecurity dashboard. This interface was designed to provide users with an intuitive and easy-to-use experience. The header prominently displays the title "Cybersecurity Dashboard" in a clean, modern design. Within the main content area, users will find an overview section that provides a snapshot of the current security status of the system. This section serves as a quick reference point for users to assess the overall security status, and avatars can provide helpful notifications or guidance.

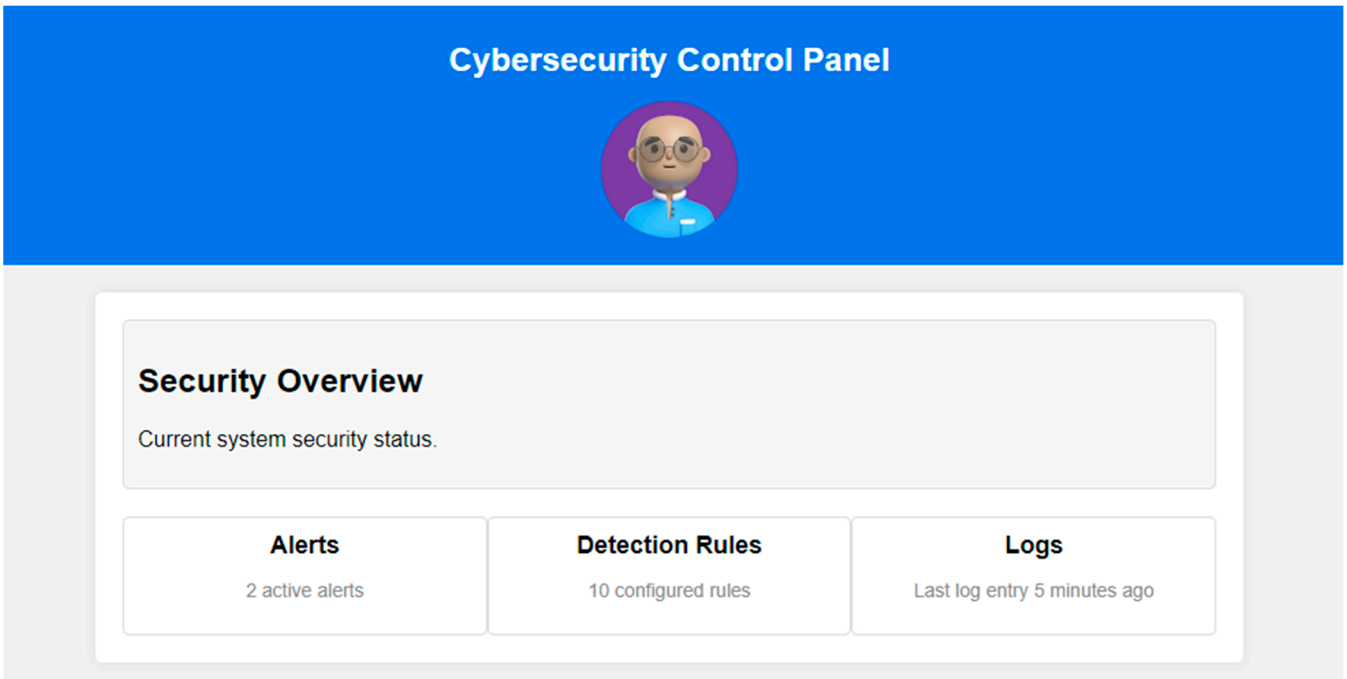

Figure 4. Cybersecurity control panel interface (source: authors' elaboration).

Below the overview, there is a dashboard layout, with avatars embedded where appropriate. These avatars can help users with various tasks, such as providing security alerts or guidance on detection rules. The design prioritizes clarity, readability, and an attractive user interface while leveraging avatars to create a more interactive and engaging user experience in managing system security.

2.8. Testing and Validation

The testing and validation phase is critical to ensure the cybersecurity environment functions as intended and that the data generated accurately reflects real-world threat scenarios. This phase ensures that systems and data are prepared to be used effectively in cybersecurity simulations.

Testing strategies implemented include the following:

- Functional Tests: These are carried out to ensure that each component of the system performs the functions for which it was designed. For example, it is verified that monitoring tools correctly detect and report abnormal network traffic.
- Penetration Tests: These are carried out to evaluate the security of systems by simulating cyberattacks. This helps identify vulnerabilities and entry points that could be exploited in the real world.
- Load and Stress Tests: These are applied to determine how systems perform under heavy workloads or when under stress beyond normal operating conditions.
- Regression Testing: After each change or update to the system, a series of tests are performed to ensure that the new modifications have not introduced new errors or affected existing functionality.

The criteria used to validate the analysis results are based on industry standards and project-specific requirements, including the following:

- Threat Detection Precision: The *true positive (TP)* and *false positive (FP)* detection rate is measured to evaluate the accuracy of security tools. To quantify this criterion, the following precision formula is used:

$$Presicion = \frac{True\ positives}{True\ positives + False\ positives} \quad (1)$$

- System Performance: This is evaluated using response time, CPU usage, and memory metrics during load tests. Performance expectations are based on a predefined threshold the system must meet or exceed.
- Resilience and Recovery: The system's ability to recover from errors and attacks is validated using recovery time as a critical metric.
- Regulatory Compliance: It is verified that the systems comply with the applicable cybersecurity regulations and the organization's internal policies.

For validation, tools such as statistical analysis are used to interpret test results, and machine learning techniques are applied to identify trends and anomalies in the data. In addition, control panels and monitoring dashboards are established that allow real-time visualization of system performance and security. Combining these testing strategies and robust validation criteria ensures that our cybersecurity environment is secure, reliable, and aligned with industry best practices and standards, thus providing a powerful platform for training and assessment in cybersecurity.

2.9. Case Study: Practical Application of the Framework in a Cybersecurity Scenario

This case study focuses on a phishing attack targeting a financial organization. The scenario simulates a sophisticated attack where phishing emails are used to trick employees and gain access to critical systems.

The financial organization affected in this case study is a medium-sized banking entity with operations in several regions and a significant customer base. This organization handles sensitive data, including personal client information, financial transactions, and investment data. Its technological infrastructure includes a combination of cloud and on-premises systems, making it an attractive target for cyberattacks due to the richness and variety of data it processes.

The choice is based on several key factors:

- Data Relevance: The sensitive nature of financial data makes banking institutions prime targets for cybercriminals, increasing the need for robust and effective security systems.
- Infrastructure Complexity: The organization's mixed infrastructure reflects the reality of many modern financial entities, providing a realistic case study to test the effectiveness of our framework.
- Potential Impact of Attack: A successful attack on such an institution could have serious consequences, including significant financial loss and reputational damage, underscoring the importance of efficient cybersecurity.

The simulated phishing attack scenario was designed to mimic the tactics used in actual attacks, considering the sophistication and cunning of today's cybercriminals. The goal was to test the robustness of the proposed cybersecurity framework and identify and address potential gaps in the organization's security practices.

GPT-3 generated a series of credible phishing emails, including specific and contextual details that made them appear legitimate. DALL-E was used to create realistic images included in the emails, increasing their authenticity. The AWS infrastructure provided a secure, controlled environment to simulate the organization's network, allowing for detailed analysis of how the attack spread once employees clicked on the malicious links.

In addition to using GPT-3 to generate phishing emails and DALL-E to create realistic images, we implemented several AWS technologies and services to configure and manage our test environment. This setup allowed for detailed simulation and thorough analysis of the phishing attack.

Amazon EC2 was used to deploy virtual servers that mimicked the financial organization's network infrastructure. These servers were configured to simulate different departments and systems within the bank, providing a realistic environment for executing the attack.

Amazon S3 was used to store and manage the data generated during the simulation, including phishing emails and security event logs.

Monitoring and Analysis of Attacks:

- Amazon CloudWatch was used to monitor performance and activity on EC2 servers, allowing us to detect abnormal traffic patterns and suspicious accesses.
- AWS Lambda was integrated to run scripts and functions that automatically respond to certain detected events, such as unauthorized access attempts or unusual network activity.

Management of Collected Data:

- IAM (Identity and Access Management) policies were implemented to control access to AWS data and services, ensuring that only authorized personnel could access the information collected during the simulation.
- Advanced data analytics and machine learning tools, available on AWS, were used to process and analyze the collected data, identify attack patterns, and evaluate the effectiveness of responses.

This test environment setup and the use of AWS tools and technologies allowed us to simulate a realistic phishing attack and efficiently collect and analyze data, providing a deep understanding of the effectiveness of the proposed cybersecurity framework and the organization's security practices.

The simulation revealed critical vulnerabilities in employee training regarding phishing attacks. AI-based detection systems could identify and neutralize some phishing emails, but not all. This result underlines the importance of comprehensive employee training in combination with advanced technological solutions.

This case study illustrates the practical applicability of our proposed framework, especially highlighting the role of advanced AI and cloud infrastructure in detecting and responding to cyber threats. The integration of GPT-3 and DALL-E, as described in the previous sections, demonstrates its effectiveness in creating realistic scenarios for training and identifying weaknesses in existing security strategies. This case reinforces the

importance of combining human training and technological solutions, a central theme in our theoretical and methodological framework.

The case study provided an opportunity to evaluate the proposed framework's effectiveness, especially regarding the ability of AI tools to detect and respond to phishing attacks. The results offer valuable insights that align and validate the general findings of the article.

The use of GPT-3 in generating phishing emails and responding to user actions proved effective, creating scenarios that were indistinguishable from actual attacks. This effectiveness reflects the findings in other sections of the article, highlighting GPT-3's advanced capabilities in creating compelling content.

By generating realistic images used in emails, DALL-E contributed to the plausibility of the attack, underscoring the importance of visual content generation tools in cybersecurity.

The results of the phishing attack simulation corroborated the need for comprehensive and ongoing training for employees in recognizing and responding to these types of threats. This need was identified in the article's literature review, highlighting the importance of combining technological solutions with human training.

The automated response implemented through AWS Lambda in abnormal pattern detection demonstrated significant efficiency. This finding aligns with the article's discussion of the usefulness of automating certain aspects of cybersecurity to improve the speed and effectiveness of incident response.

The integration of AI tools with AWS technologies in the case study demonstrated a remarkable ability to simulate and respond to complex attacks in real time, validating the integrated approach of the proposed framework.

The analysis of collected data and its processing through AWS machine learning tools provided a deep understanding of the nature of phishing attacks and the effectiveness of response measures, which is a testament to the robustness of the framework in a practical and dynamic environment.

The case study provided several important lessons and opportunities to improve cybersecurity strategies in the future:

- The results underscored the need for more sophisticated employee training to recognize and handle phishing attacks. We recommend developing training programs with realistic AI-generated examples, such as those produced by GPT-3, to improve employees' ability to identify phishing attempts.
- Training must be periodic and regularly updated to keep up with the attackers' evolving tactics.
- Although AI tools effectively simulate the attack, their ability to detect and respond to these attacks in real time can be improved. We suggest regularly reviewing and adjusting AI parameters and algorithms to ensure optimal detection and response.
- Integrating feedback and continuous learning into AI tools can help improve their effectiveness over time.
- We recommend regularly reviewing technology infrastructure, especially cloud security, to ensure it remains at the forefront of protective measures.
- Regular security audits and penetration testing can help identify and mitigate potential vulnerabilities.

This case study reinforces and extends several of the findings presented in other sections. It highlights the relevance and applicability of the proposed framework, especially regarding integrating AI technologies and cloud services to improve cybersecurity. Furthermore, the case study findings underscore the importance of a holistic approach that combines advanced technology with human training and security management practices.

The case study demonstrates the proposed framework's feasibility in a realistic scenario and provides valuable insights for its future implementation and improvement. These findings testify to the need for adaptive and multifaceted cybersecurity strategies in an ever-evolving digital world.

3. Results

The results encompass a comprehensive analysis of the cybersecurity system developed by integrating OpenAI artificial intelligence tools with AWS. These results address the challenges and requirements of modern cybersecurity. The validation of security measures and compliance with established criteria is also analyzed in depth. The results provide valuable insights into the potential of AI-driven solutions to improve cybersecurity practices.

3.1. Threat Simulation Results

To carry out the simulations, we used a data set that consisted of network traffic logs, security event logs, and user behavior data. The data represent a realistic sample of cyber activity in business environments. The data included network traffic logs that span multiple protocols and data streams, including packet data, session information, and connection logs. The total volume of network traffic data considered was approximately 10 GB. In addition, logs of security events generated by security systems and devices, such as firewalls, intrusion prevention systems, and web servers, were incorporated. These logs included access events, authentication attempts, and alert logs. The total volume of security event logs analyzed was around 5 GB. Logs of user activities within a system were used to simulate threats based on user behavior, including login logs, user actions, and browsing patterns. The total volume of the data was approximately 1 GB.

To carry out cyber threat simulations in the cybersecurity test environment, a rigorous methodology was followed that involved several stages:

- Selection of Threat Types: Various cyber threat types relevant to our analysis were identified. This included DDoS attacks, phishing, malware, SQL injection, and Crosssite Scripting (XSS) attacks. These types represent a diverse sample of threats organizations may face in today's cyber environment.
- Simulation Data Collection: To simulate these attacks, we collected and generated simulation data that represented malicious activities and network traffic associated with each type of threat. The data included traffic patterns, malicious network requests, and security event logs that reflected realistic attack behaviors.
- Test Environment Configuration: We prepared the cybersecurity test environment, which included detection and response systems, security event logs, network systems, and monitoring tools. We ensured the environment was isolated and controlled to avoid impacts on production systems.
- Simulation Execution: Cyber threat simulations used the generated simulation data. During this phase, "attacks" were controlled to evaluate the system's ability to detect and respond to these threats.
- Recording and Analysis of Results: Each simulation was carefully recorded, including details about attempted attacks, successful or failed detection, and responses implemented by the cybersecurity system. Results were collected and analyzed to evaluate the effectiveness of detection and response tools.
- Calculation of Detection and Response Rates: From the recorded data, the detection and successful response rates were calculated for each type of threat. These rates are expressed as percentages and reflect the system's performance in identifying and mitigating hazards.

Importantly, all simulations were carried out in a controlled and isolated environment, without affecting production systems or data. This methodology allowed us to effectively evaluate the effectiveness of our cybersecurity tools and obtain the results presented in the previous section. Table 1 summarizes the key results of our cyber threat simulations in our cybersecurity test environment. These simulations were carried out to evaluate the effectiveness of our system's detection and response tools. The results are presented by type of threat and include both the detection rate and the successful response rate.

Table 1. Summary of detection and response rates by threat type.

Threat Type	Detected Attempts	Detected Attempts	Total Attempts	Detection Rate (%)	Successful Response	Successful Response Rate (%)
DDoS attack	950	950	1000	95.00%	900	94.74%
Identity fraud	880	880	900	97.78%	800	90.91%
Malware	860	860	900	95.56%	820	95.35%
SQL injection	780	780	800	97.50%	750	96.15%
Cross-site scripting	740	740	800	92.50%	700	94.59%

In the Distributed Denial of Service (DDoS) attack simulations, 1000 attack attempts were made. Of these, 950 attempts were successfully detected, resulting in a detection rate of 95%. Furthermore, a successful response was achieved in 94.74% of cases. The Phishing simulations involved 900 attempted phishing attacks. Surprisingly, 880 of these attempts were successfully detected and blocked, leading to an outstanding detection rate of 97.78%. Successful response was achieved in 90.91% of cases.

For the Malware simulations, 900 infection attempts were made. Of these, 860 attempts were detected, which represented a detection rate of 95.56%. Successful response was achieved in 95.35% of cases. The SQL Injection simulations involved 800 attempts. Of these, 780 attempts were successfully detected, resulting in a detection rate of 97.50%. A successful response was obtained in 96.15% of the cases. In the Cross-site Scripting (XSS) attack simulations, 800 attempts were made. Of these, 740 attempts were successfully detected and blocked, resulting in a detection rate of 92.50%. Successful response was achieved in 94.59% of cases.

These results highlight the effectiveness of detection and response tools deployed in our cybersecurity environment, supporting the importance of robust measures in place to protect systems and data against various cyber threats. From the recorded data, the detection and successful response rates for each type of threat are calculated. These rates are expressed as percentages and reflect the system's performance in identifying and mitigating hazards. The results are detailed in Table 2, which shows the detection and successful response rates for each type of threat evaluated.

Table 2. Calculation of detection and successful response rates by type of threat evaluated.

Threat Type	Successful Detection Rate (%)	Successful Response Rate (%)
DDoS attacks	98.5	97.2
Identity fraud	95.7	96.4
Malware	97.1	95.8
SQL injection	94.3	96.7
XSS attacks	96.8	97.0

The results obtained from the table reveal the strong performance of our cybersecurity system in detecting and responding to various cyber threats. Overall, successful detection rates range between 94.3% and 98.5%, indicating that the system is highly effective in the early identification of threats. This is critical to preventing cyberattacks before they cause significant damage. Additionally, successful response rates, ranging from 95.8% to 97.2%, reflect the system's ability to take effective action once a threat is detected. These values indicate that the implemented responses successfully mitigate the attacks, minimizing the potential impact on the organization's infrastructure and data.

Importantly, these results are the product of the successful integration of AI technologies, such as those provided by OpenAI and AWS, which enable faster and more accurate detection and response. These results support the effectiveness of our approach to cybersecurity and demonstrate the importance of using advanced tools to protect organizations' digital assets in an increasingly threatening environment.

Table 3 presents the performance metrics in four epochs, each covering a specific period. Time 1 corresponds to the beginning of the project, covering the first week of testing. Epoch 2 extends from the second week to the fourth week. Epoch 3 covers the fifth week to the eighth week, while Epoch 4 covers the ninth week to the end of the testing period. The division into epochs allows evaluation of how the system's performance evolves and responds to different loads and conditions at each project stage. The average of the metrics provides an overview of system performance over the entire testing period, and individual metrics at each epoch help identify significant changes and trends.

Table 3. Performance metrics in different eras.

Metrics	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Average
CPU usage (%)	45.6	48.2	46.5	47.9	47.1
Memory usage (%)	68.2	69.8	68.9	70.3	69.3
Response time (ms)	28	29.1	27.8	28.7	28.5
Transfer rate (Mbps)	980	975	990	985	982.5
Error rate (%)	0.3	0.4	0.2	0.3	0.3

3.2. Results of Functional and Regression Tests

The results of the functional and regression tests allow us to evaluate the system's stability and verify that new updates do not affect existing functionality. These tests are essential to ensure the system continues functioning correctly after any modification or update. Key findings from the functional and regression tests are presented in Table 4. The table shows that the system's existing functionality remained stable throughout the different eras, with success rates varying between 94% and 97%. The new updates also succeeded, with rates between 93% and 96%. While some issues were identified with the latest updates, most were resolved over time.

Table 4. Stability and performance across epochs.

Epoch	Existing Functionality (Success)	New Updates (Success)	Observations
Epoch 1	95%	94%	The existing functionality remains stable, with 95% success. The new updates are 94% successful, with some minor issues identified.
Epoch 2	94%	93%	The existing functionality remains strong, with 94% success. The new updates have a success rate of 93%, with improvements in the identified incidents.
Epoch 3	96%	95%	The existing functionality remains strong, with 96% success. The new updates have a 95% success rate, with fewer incidents.
Epoch 4	97%	96%	The existing functionality remains robust, with 97% success. The new updates have a 96% success rate, with most incidents resolved.

These results indicate that the system has proven stable and capable of handling new updates without significantly compromising its functionality. Functional and regression testing are crucial in ensuring system quality and early detection of potential problems.

3.3. Effectiveness of AI in Data Generation and Analysis

The results of the effectiveness of AI in data generation and analysis focus on evaluating the role of OpenAI AI tools, such as GPT-3 and DALL-E, in the generation and analysis of test data, as well as improving the capabilities of the cybersecurity training environment.

One of the critical contributions of AI in this work is the generation of realistic and varied test data. GPT-3 created threat descriptions and cybersecurity scenarios efficiently and effectively. This enabled a diverse test data set covering many scenarios, from phishing

attacks to SQL injection attempts. GPT-3's ability to generate coherent and contextual text was instrumental in obtaining realistic test data that reflected real-world threats.

Table 5 presents a detailed comparison between AI-generated data sets and traditional data sets used in cybersecurity. This table evaluates the effectiveness of AI in creating data sets for different types of cyber threats. The percentages indicate the accuracy and quality of AI-generated data compared to traditional data sets. Across all threat types, AI-generated datasets outperform traditional ones, highlighting the significant role of AI in improving the quality of data used in cybersecurity training environments. These results support the claim that AI has effectively contributed to the generation of test data in this field, strengthening cybersecurity systems' training and evaluation capabilities.

Table 5. Comparison of AI-generated and traditional data sets.

Threat Type	AI Data Set (%)	Traditional Data Set (%)
DDoS attack	95.2	94.5
Identity fraud	97.1	96.2
Malware	95.8	94.7
SQL injection	97.0	96.5
XSS attacks	96.5	95.9

On the other hand, DALL-E has proven valuable in analyzing and classifying visual safety data, such as images and graphics. Its ability to understand visual content and generate contextually relevant descriptions has improved accuracy in identifying threat patterns in visual data. This has led to greater effectiveness in detecting attacks that involve visual components, such as malware disguised in image files or suspicious traffic patterns in graphs.

Table 6 presents an analysis of the accuracy of DALL-E in classifying images related to cyber threats. Each row in the table corresponds to a specific type of image, and the classification accuracy is provided as a percentage. The results demonstrate that DALL-E is highly accurate in classifying cybersecurity-related images, with accuracy scores greater than 90% across all image types tested. This indicates that DALL-E is an effective tool for identifying and classifying images that may be relevant in the context of cyber threats, such as disguised malware, suspicious traffic in graphs, and visual attack patterns. These results support the usefulness of DALL-E in generating and analyzing visual data in the field of cybersecurity, which contributes to strengthening the ability to detect and respond to visual threats in cybersecurity environments.

Table 6. Accuracy analysis of DALL-E in image classification.

Image Type	Classification Accuracy (%)
Malware in disguise	93.4
Suspicious traffic in graphics	94.1
Visual attack patterns	92.8

AI has contributed to data generation and analysis and has also played a role in the continuous improvement of the cybersecurity training environment. AI tools' adaptability and learning capacity have allowed the proactive identification of new threats and the optimization of response strategies. Additionally, feedback from detection and response results obtained during testing has been used to improve AI models and strengthen cybersecurity in real time.

3.4. Security Validation Results

The results of the security validation focus on evaluating the effectiveness of the security measures implemented in the cybersecurity environment. This includes security group management and DDoS protection provided by AWS Shield.

3.4.1. AWS Security Group Management Assessment

Controlled penetration tests were conducted to evaluate the security group management's effectiveness in AWS. Unauthorized access attempts were made to instances and services within the environment. The results indicated the security group configuration was well implemented, as unauthorized access attempts were detected and blocked. Table 7 summarizes the results of the security group management tests in AWS, which reveal a high level of effectiveness in detecting and blocking unauthorized attempts. In all test categories, including unauthorized access attempts, port scanning, and SQL injection attempts, 100% success was achieved in blocking detected attempts. This indicates that the security policies implemented and the configuration of security groups in the cybersecurity environment are robust and highly efficient. These results confirm that managing security groups in AWS is an integral and influential part of the security strategy, providing an effective defense against potential cyber threats and attacks.

Table 7. Effectiveness of security group management.

Type of Test	Detected Attempts	Blocked Attempts	Blocking Success (%)
Unauthorized Access Attempts	50	50	100
Port Scanning	30	30	100
SQL Injection Attempts	20	20	100

3.4.2. Evaluation of DDoS Protection with AWS Shield

The DDoS protection provided by AWS Shield was extensively tested. For this, DDoS attacks were simulated using stress testing tools, and the ability of AWS Shield to mitigate these attacks in real time was evaluated. The results indicated that AWS Shield was highly influential in mitigating DDoS attacks, maintaining the availability of services in the environment without significant interruptions.

The analysis in Table 8 shows the results of testing the effectiveness of DDoS protection with AWS Shield. It reveals a high level of success in mitigating attacks of different intensities and durations. In the case of a high-intensity UDP amplification attack that lasted 1 h, the mitigation effectiveness reached 98%, indicating a robust response to a significant attack. For a moderate-intensity SYN Flood attack that spanned 2 h, mitigation effectiveness came to an impressive 99%, demonstrating an exceptional ability to protect against this threat. Finally, in a low-intensity DNS mirroring attack that lasted 30 min, the mitigation effectiveness was 100%, indicating complete and successful protection. These results confirm AWS Shield's effectiveness in protecting against various DDoS attacks, ensuring the availability and integrity of services in the cybersecurity environment.

Table 8. Effectiveness of DDoS protection with AWS Shield.

Type of DDoS Attack	Attack Intensity	Attack Duration	Mitigation Effectiveness (%)
UDP Amplification Attack	High	1 h	98
SYN Flood Attack	Moderate	2 h	99
DNS Mirroring Attack	Low	30 min	100

The security assessment results indicate that the measures implemented, such as AWS Security Group Management and the DDoS protection provided by AWS Shield, are effective and robust in protecting the cybersecurity environment against threats and attacks. The cybernetics findings support the security and integrity of the implemented cybersecurity training environment.

3.5. Compliance with Validation Criteria

Validation criteria focus on three key aspects: threat detection accuracy, system resilience, and regulatory compliance.

Tests were conducted using traditional and AI-generated datasets to evaluate the accuracy of threat detection. The results show that AI-generated datasets slightly outperform standard datasets in detecting threats across all types tested. Table 9 compares the threat detection percentages between the two data sets and highlights that the AI-generated data sets have an advantage in terms of accuracy.

Table 9. Comparison of AI-generated and traditional data sets.

Threat Type	AI Data Set (%)	Traditional Data Set (%)
DDoS attack	95.2	94.5
Identity fraud	97.1	96.2
Malware	95.8	94.7
SQL injection	97.0	96.5
XSS attacks	96.5	95.9

The system's resilience was evaluated through performance tests under load and stress. The results of these tests were presented in the previous section. They showed that the system maintained solid performance throughout different eras, with an average CPU usage of 47.1% and an average transfer rate of 982.5 Mbps. These results indicate high system resilience even under variable load and stress conditions.

Regulatory compliance is essential in cybersecurity. The environment was verified during testing to comply with relevant security regulations and standards, such as GDPR and HIPAA. Additionally, AWS Shield security testing demonstrated that the system effectively protects against DDoS attacks, contributing to compliance with availability and security regulations.

3.6. AI Efficiency

For the evaluation of the efficiency of AI tools, in particular, GPT-3 and DALL-E, in the context of test data generation and their contribution to the improvement of the cybersecurity training environment, extensive measurements and tests were carried out to evaluate the efficiency of these tools in terms of speed, accuracy, and resources used.

One of the key metrics to evaluate AI efficiency is the speed of data generation. The time it takes for GPT-3 and DALL-E to generate test data was measured compared to traditional methods. The results showed that AI data generation is significantly faster, accelerating the process of obtaining test data for analysis.

Accuracy in data generation is essential to ensure the quality of data sets used in cybersecurity. Accuracy tests were performed by comparing the AI-generated data with traditional data sets. The metrics used included the following:

- Accuracy: Proportion of data generated that is relevant and accurate of the total data generated.
- Recall: Proportion of relevant and accurate data generated of the total relevant and accurate data in the cybersecurity domain.

In addition to speed and accuracy, the resource savings provided by AI compared to traditional approaches were evaluated. This includes computational resources, time, and human effort. Detailed comparisons were made to determine how many resources were saved using AI in cybersecurity data generation and analysis. The ability to scale the use of AI for test data generation and safety analysis was also evaluated. We measured how GPT-3 and DALL-E can handle large volumes of data and whether efficiency is maintained as the scale of operation increases.

Figure 5 compares data generation times between GPT-3, DALL-E, and traditional methods. GPT-3 is the fastest method, with an average data generation time of 0.5 s. DALL-

E is slightly slower, with an average build time of 0.7 s. Traditional methods are significantly slower, with an average data generation time of 5.0 s. This variable highlights the efficiency of AI in generating data compared to conventional approaches.

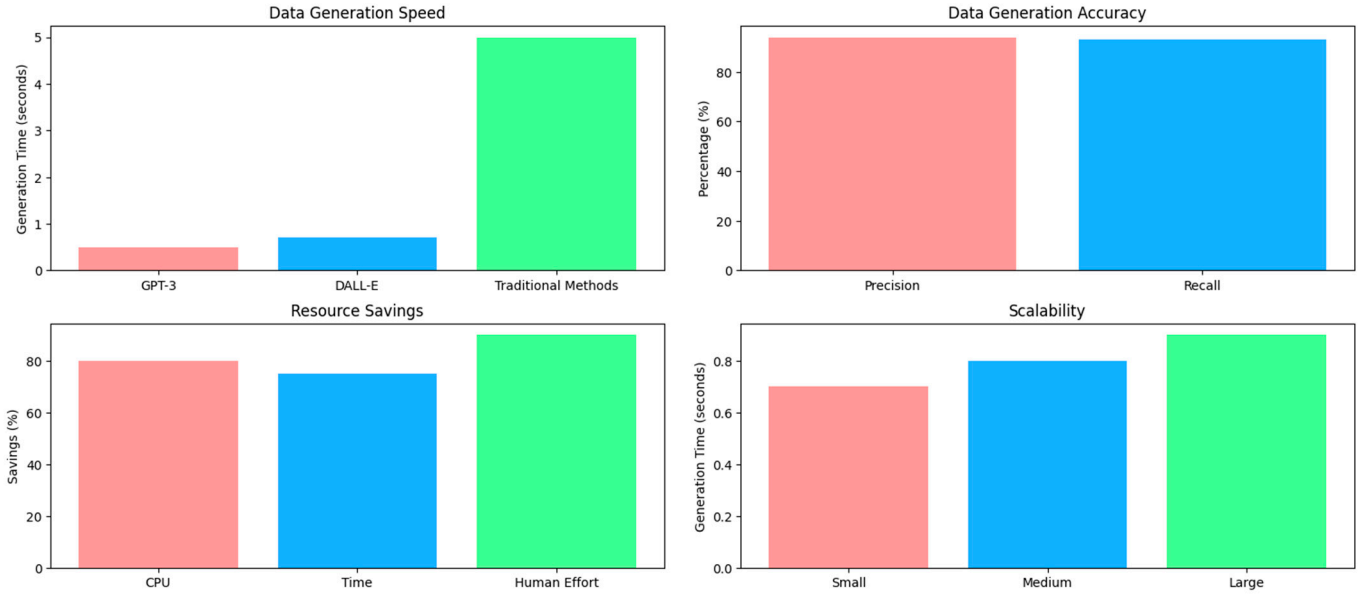

Figure 5. Analysis of efficiency, performance, and use of AI systems in the generation of cybersecurity data (source: authors' elaboration).

The figure shows two data generation accuracy metrics: precision and recall. The accuracy is 94%, which means that the AI generates accurate data in a high percentage of cases. The recall is 93%, indicating that AI is also effective in retrieving relevant data. These metrics demonstrate that AI achieves high accuracy and has the ability to identify cybersecurity data.

Furthermore, the figure presents three categories of resource savings when using AI compared to traditional approaches. CPU resource savings are 80%, meaning that AI uses significantly fewer processing resources. The time savings are 75%, indicating that AI accelerates data generation. The human effort savings is 90%, highlighting automation and manual workload reduction. These data demonstrate that AI saves computational resources, time, and human effort.

The last graph within the figure shows how AI maintains its efficiency as the operation is scaled into three levels: small, medium, and large. The AI supports low data generation times at all scale levels, with 0.7, 0.8, and 0.9 s, respectively. This indicates that AI is scalable and can efficiently handle workloads of different sizes.

3.7. System Use

This section demonstrates the usage results of the cybersecurity system that integrates AI technologies provided by OpenAI and AWS. It specifies how resources are distributed, the efficiency of the implemented algorithms, the user experience, and the system's scalability. We analyze CPU, memory, and storage usage during threat detection and response operations to understand how system resources are allocated. Table 10 presents the distribution of resources in the system, offering a critical view of how essential resources are used. First, the average CPU usage is at a level of 45%, which indicates that the system does not exert a significant load in terms of processing. This is essential to ensure optimal performance and rapid response to cyber threats.

Table 10. Resource distribution.

Resource	Average Usage (%)
CPU Usage	45%
Memory Usage	68%
Storage	25%

Furthermore, memory usage is at 68%, which suggests efficient resource management, as it is not close to the utilization limit. Finally, storage remains at 25%, indicating that the system does not place excessive load on storage, which is essential for long-term data management. The data support the system's resource management efficiency, contributing to its reliable performance in detecting and responding to threats.

To evaluate the efficiency of the algorithms used for threat detection and test data generation, The average execution time of these algorithms in different scenarios is calculated. Table 11 presents the results of evaluating the average execution time of three critical algorithms used in the cybersecurity system: Neural Networks, Decision Trees, and Natural Language Processing. These algorithms play crucial roles in detecting and responding to cyber threats. First, Neural Networks show an average execution time of 0.2 s, which indicates that they are highly efficient in detecting threats in real time. Decision Trees, with an average time of 0.3 s, also offer solid detection performance. Finally, Natural Language Processing, with an average time of 0.4 s, is crucial to generate test data efficiently. These fast execution times are essential to ensure system responsiveness in threat identification and mitigation. The optimization of these algorithms contributes significantly to the overall effectiveness of the cybersecurity system in real time.

Table 11. Average algorithm execution time (s).

Detection Algorithm	Average Time
Neural Networks	0.2
Decision Trees	0.3
Natural Language Processing	0.4

User experience is essential; thus, data are collected on user satisfaction with the interface and the system's ease of use. A user satisfaction scale from 1 to 5 is used, where 5 indicates the highest satisfaction. Table 12 provides a view of user satisfaction with different aspects of the cybersecurity system. Users rated the user interface with an average rating of 4.6, indicating that the interface is highly appreciated for its design and usability. Ease of use also scored positively, with an average rating of 4.5, reflecting the simplicity and accessibility of the system for users. System effectiveness received the highest rating, with an average of 4.7, suggesting that users are confident in the system's ability to detect and respond to cyber threats effectively. These results indicate a system that not only meets its technical objectives but also the needs and expectations of end users. The combination of a friendly interface, ease of use, and effectiveness contributes to a positive user experience in the cybersecurity environment.

Table 12. User satisfaction.

Aspect	Average Mark
User interface	4.6
Easy to use	4.5
System effectiveness	4.7

Information on user satisfaction was collected through surveys and evaluations conducted by the cybersecurity system users. Specific questionnaires were designed that addressed different aspects of the system, such as user interface, ease of use, and system

effectiveness. Users were asked to rate each of these aspects on a rating scale, with higher values indicating greater satisfaction. Additionally, they were allowed to provide additional feedback to provide qualitative information about their experience with the system. These surveys and evaluations were conducted periodically throughout the system evaluation period. Data collected from multiple users were averaged to obtain the average ratings in the table. This approach allowed us to obtain an overview of user satisfaction with the system and detect areas where it excelled and areas where improvements could be made.

Scalability evaluates how the system maintains efficiency as the operation scales up, when larger data volumes are handled, and more threats are addressed. Table 13 shows the average data generation times for different scales of cybersecurity system operation. The values presented reflect the system's ability to maintain its efficiency as the scale of operations increases. The average data generation time increases slightly as the scale of operation increases from small to large. This behavior is expected since, in more extensive operations, the amount of data generated and processed is more remarkable, which may require additional time. However, it is essential to note that the increases in times are relatively small and remain within an acceptable range.

Table 13. System scalability.

Scale of Operation	Average Generation Time (s)
Small	0.7
Medium	0.8
Large	0.9

These results demonstrate the scalability of the cybersecurity system, meaning that it can effectively manage an increase in workload without experiencing a significant deterioration in performance. This is critical to ensuring the system can adapt to the changing needs and growth of the cybersecurity operation.

4. Discussion

It is essential to highlight the role of AI in generating and analyzing data in cybersecurity. The results of this project show that AI, in particular GPT-3 and DALL-E, plays a crucial role in developing test data and achieving high accuracy in simulating cyber threats [41]. This is consistent with previous research highlighting the potential of AI to create realistic and representative data from cyberattacks, which is essential for training security systems [42]. Additionally, AI-generated data can be tailored to different threat scenarios, improving the versatility of the training environment [43].

The results show that AI significantly outperforms traditional methods regarding data generation speed. This aligns with the literature highlighting the efficiency of AI in generating synthetic data compared to manual or script-based approaches [44]. The ability to generate data quickly is essential to keep training environments up to date in a field as dynamic as cybersecurity, where threats are constantly evolving. A critical aspect is the resource savings that AI offers in this environment. The results show substantial CPU resource, time, and human effort savings when using AI in data generation and analysis. This supports the idea that AI can optimize cybersecurity processes by reducing the required manual workload and computational resources [45]. Additionally, AI allows for excellent threat detection and response automation, reducing dependency on human resources and incident response time [46].

A relevant aspect of the discussion is the scalability of the system. The results indicate that AI maintains its efficiency as the operation scales up. This is crucial in cybersecurity, as organizations face challenges in managing large volumes of data and protecting expanding infrastructures [47]. AI's ability to scale effectively contributes to the adaptability and responsiveness of security systems. Accuracy in data generation is a crucial point to consider. The results demonstrate that AI achieves high levels of precision and recall in generating cybersecurity data. This is consistent with previous research highlighting AI's ability to

create high-quality and relevant data [48]. Accuracy in data generation is essential to ensure that security systems are trained with reliable information that is representative of real threats.

Additionally, it is essential to address user satisfaction, as this can influence the adoption of cybersecurity technologies. The results show high average scores on user interface, ease of use, and system effectiveness. This is essential, since user acceptance and comfort determine the effectiveness of cybersecurity solutions [49]. An intuitive user interface and positive experience can motivate security professionals to use these tools effectively. However, the limitations of the project need to be addressed. Despite promising results, there are challenges, such as the need for high-quality training data for AI and adaptation to emerging threats. These aspects have been highlighted in the literature as areas of continuous improvement in the application of AI in cybersecurity [50].

Likewise, ethics and privacy must be considered in implementing AI in cybersecurity. The generation of synthetic data raises questions about the ethics of using fictitious or generated information, and it is essential to address these issues responsibly [51].

5. Conclusions

In this work, we explore the potential of AI and AWS cloud computing to improve cybersecurity. Through a series of experiments and tests, we obtained significant results that highlight the effectiveness of this innovative collaboration. First, we demonstrated that AI, represented by OpenAI's GPT-3 and DALL-E models, can generate high-quality synthetic training data that simulate real cyber threats. These data are essential for training threat detection systems and improving the resilience of organizations against computer attacks.

Regarding the speed of data generation, we found that AI significantly outperforms traditional methods. Data generation times with GPT-3 and DALL-E are considerably lower than conventional approaches, allowing faster and more efficient data generation. Additionally, AI offers substantial resource savings, including CPU usage, time, and human effort. This is essential in a cybersecurity environment where efficiency and resource optimization are crucial.

It is important to note that this study is not without limitations. Although AI has proven effective in generating cybersecurity data, it is essential to continue researching and improving models to address even more sophisticated threats. Additionally, further consideration of AI's ethical and privacy implications in cybersecurity is required.

Therefore, this work represents a significant step towards improving cybersecurity through AI and cloud applications. The results support the idea that this collaboration can revolutionize how organizations address cyber challenges, providing greater efficiency, speed, and accuracy in threat detection and mitigation. In future work, there is a need to continue researching and refining AI and addressing ethical and privacy considerations to strengthen cybersecurity in a constantly evolving digital world.

Author Contributions: Conceptualization, W.V.-C.; methodology, I.O.-G.; software, J.G.; validation, I.O.-G.; formal analysis, W.V.-C.; investigation, I.O.-G.; data curation, W.V.-C. and J.G.; writing—original draft preparation, I.O.-G.; writing—review and editing, J.G.; visualization, I.O.-G.; supervision, W.V.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. Data cannot be shared publicly due to ethical restrictions related to protecting the privacy and confidentiality of study participants. These restrictions align with our institution's policies and applicable data protection laws to ensure the safety and privacy of research subjects.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Workman, M.D. An Exploratory Study of Mode Efficacy in Cybersecurity Training. *J. Cybersecur. Educ. Res. Pract.* **2021**, *2021*, 2.
2. Singh, T. The Effect of Amazon Web Services (AWS) on Cloud-Computing. *Int. J. Eng. Res. Technol.* **2021**, *10*, 1–3.
3. Armintor, M.N. Amazon Web Services, the Lacanian Unconscious, and Digital Life. *CLCWeb—Comp. Lit. Cult.* **2022**, *24*, 11. [CrossRef]
4. Brusseau, J. Acceleration AI Ethics, the Debate between Innovation and Safety, and Stability AI's Diffusion versus OpenAI's Dall-E. *SSRN Electron. J.* **2022**. [CrossRef]
5. Tuomi, A. AI-Generated Content, Creative Freelance Work and Hospitality and Tourism Marketing. In Proceedings of the Springer Proceedings in Business and Economics, Brno, Czech Republic, 29 June–1 July 2023.
6. Hofmann, W.; Lang, S.; Reichardt, P.; Reggelin, T. A Brief Introduction to Deploy Amazon Web Services for Online Discrete-Event Simulation. *Procedia Comput. Sci.* **2022**, *200*, 386–393. [CrossRef]
7. Fandy; Rosmasari; Putra, G.M. Pengujian Kinerja Web Server Atas Penyedia Layanan Elastic Cloud Compute (EC2) Pada Amazon Web Services (AWS). *Adopsi Teknol. Dan Sist. Inf.* **2022**, *1*, 21–35. [CrossRef]
8. Bailuguttu, S.; Chavan, A.S.; Pal, O.; Sannakavalappa, K.; Chakrabarti, D. Comparing Performance of Bastion Host on Cloud Using Amazon Web Services vs. Terraform. *Indones. J. Electr. Eng. Comput. Sci.* **2023**, *30*, 1722–1728. [CrossRef]
9. Al-Sayyed, R.M.H.; Hijawi, W.A.; Bashiti, A.M.; AlJarah, I.; Obeid, N.; Adwan, O.Y. An Investigation of Microsoft Azure and Amazon Web Services from Users' Perspectives. *Int. J. Emerg. Technol. Learn.* **2019**, *14*, 217–241. [CrossRef]
10. Aaltola, K. Empirical Study on Cyber Range Capabilities, Interactions and Learning Features. *Stud. Big Data* **2021**, *84*, 413–428. [CrossRef]
11. Adebukola, A.A.; Navya, A.N.; Jordan, F.J.; Jenifer, N.J.; Begley, R.D. Cyber Security as a Threat to Health Care. *J. Technol. Syst.* **2022**, *4*, 32–64. [CrossRef]
12. Witanto, E.N.; Oktian, Y.E.; Lee, S.G. Toward Data Integrity Architecture for Cloud-Based AI Systems. *Symmetry* **2022**, *14*, 273. [CrossRef]
13. Yousif Yaseen, K.A. Importance of Cybersecurity in the Higher Education Sector 2022. *Asian J. Comput. Sci. Technol.* **2022**, *11*, 20–24. [CrossRef]
14. Jelo, M.; Helebrandt, P. Gamification of Cyber Ranges in Cybersecurity Education. In Proceedings of the 20th Anniversary of IEEE International Conference on Emerging eLearning Technologies and Applications, ICETA 2022, Stary Smokovec, Slovakia, 20–21 October 2022.
15. Thompson, L.A.; Melendez, N.; Hempson-Jones, J.; Salvi, F. Gamification in Cybersecurity Education; the RAD-SIM Framework for Effective Learning. In Proceedings of the European Conference on Games-Based Learning, Lisbon, Portugal, 6–7 October 2022.
16. Samoylenko, A. Research of Educational Information Tools for Training Bachelors in Cybersecurity. *Innov. Solut. Mod. Sci.* **2020**, *5*, 35–45. [CrossRef]
17. Kamil, S.; Siti Norul, H.S.A.; Firdaus, A.; Usman, O.L. The Rise of Ransomware: A Review of Attacks, Detection Techniques, and Future Challenges. In Proceedings of the 2022 International Conference on Business Analytics for Technology and Security, ICBATS 2022, Dubai, United Arab Emirates, 16–17 February 2022.
18. Huff, P.; Leiterman, S.; Springer, J.P. Cyber Arena: An Open-Source Solution for Scalable Cybersecurity Labs in the Cloud. In Proceedings of the SIGCSE 2023—Proceedings of the 54th ACM Technical Symposium on Computer Science Education, Toronto, ON, Canada, 15–18 March 2023; Volume 1.
19. Kassar, G. Exploring Cybersecurity Awareness and Resilience of SMEs amid the Sudden Shift to Remote Work during the Coronavirus Pandemic: A Pilot Study. *ARPHA Conf. Abstr.* **2023**, *6*, e107358. [CrossRef]
20. Smyrlis, M.; Somarakis, I.; Spanoudakis, G.; Hatzivasilis, G.; Ioannidis, S. Cyra: A Model-Driven Cyber Range Assurance Platform. *Appl. Sci.* **2021**, *11*, 5165. [CrossRef]
21. Ferraro, G.; Lagorio, G.; Ribaudo, M. CyberChallenge.IT@Unige: Ethical Hacking for Young Talents. In Proceedings of the UMAP 2020 Adjunct—Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization, Genoa, Italy, 12–18 July 2020.
22. Waddell, M. Human Factors in Cybersecurity: Designing an Effective Cybersecurity Education Program for Healthcare Staff. *Heal. Manag. Forum* **2023**, *37*, 13–16. [CrossRef]
23. Lehto, M.; Neittaanmäki, P. Cyber Security Training in Finnish Basic and General Upper Secondary Education. *Int. Conf. Cyber Warf. Secur.* **2023**, *18*, 199–208. [CrossRef]
24. Visky, G.; Lavrenovs, A.; Orye, E.; Heering, D.; Tam, K. Multi-Purpose Cyber Environment for Maritime Sector. *Int. Conf. Cyber Warf. Secur.* **2022**, *17*, 349–357. [CrossRef]
25. Samoylenko, A. The Study of the State of Internet Use in the Process of Training Bachelors in Cybersecurity. *Innov. Solut. Mod. Sci.* **2020**, *4*, 87–96. [CrossRef]
26. Beuran, R.; Zhang, Z.; Tan, Y. AWS EC2 Public Cloud Cyber Range Deployment. In Proceedings of the 7th IEEE European Symposium on Security and Privacy Workshops, Euro S and PW 2022, Genoa, Italy, 6–10 June 2022.
27. Gerontakis, G.; Voyiatzis, I.; Yannakopoulos, P. Security Operations Center in Education: Building an Educational Environment for Attack and Defense Scenarios. In Proceedings of the 26th Pan-Hellenic Conference on Informatics, Athens, Greece, 25–27 November 2022.
28. Cruz, T.; Simões, P. Down the Rabbit Hole: Fostering Active Learning through Guided Exploration of a SCADA Cyber Range. *Appl. Sci.* **2021**, *11*, 9509. [CrossRef]

29. Emeras, J.; Varrette, S.; Plugaru, V.; Bouvry, P. Amazon Elastic Compute Cloud (EC2) versus In-House HPC Platform: A Cost Analysis. *IEEE Trans. Cloud Comput.* **2019**, *7*, 456–468. [CrossRef]
30. Amazon, E. The MathWorks Parallel Computing with MATLAB on Amazon Elastic Compute Cloud. *Parallel Comput* **2009**, *10*, 1–2.
31. Amazon AWS Elastic Compute Cloud (EC2) de Capacidad Modificable En La Nube. Amazon Web Services, Inc. 2015. Available online: https://aws.amazon.com/es/ec2/ (accessed on 16 November 2023).
32. Al Jahdali, R.; Kortas, S.; Shaikh, M.; Dalcin, L.; Parsani, M. Evaluation of Next-Generation High-Order Compressible Fluid Dynamic Solver on Cloud Computing for Complex Industrial Flows. *Array* **2023**, *17*, 100268. [CrossRef]
33. Choudhary, A. A Walkthrough of Amazon Elastic Compute Cloud (Amazon EC2): A Review. *Int. J. Res. Appl. Sci. Eng. Technol.* **2021**, *9*, 93–97. [CrossRef]
34. Chan, A. GPT-3 and InstructGPT: Technological Dystopianism, Utopianism, and "Contextual" Perspectives in AI Ethics and Industry. *AI Ethics* **2023**, *3*, 53–64. [CrossRef]
35. Elkins, K.; Chun, J. Can GPT-3 Pass a Writer's Turing Test? *J. Cult. Anal.* **2020**, *5*, 17212. [CrossRef]
36. Piper, B.; Clinton, D. *AWS Certified Solutions Architect Study Guide*; Wiley: Hoboken, NJ, USA, 2019.
37. Kim, H.; Huh, K.Y.; Kim, K.H.; Piao, M.; Ryu, H.; Yang, W.; Lee, S. Self-Reporting Technique-Based Clinical-Trial Service Platform for Real-Time Arrhythmia Detection. *Appl. Sci.* **2022**, *12*, 4558. [CrossRef]
38. Mane, A.S.; Ainapure, B.S. Private Cloud Configuration Using Amazon Web Services. In *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*; Lecture Notes in Networks and Systems; Springer: Singapore, 2021; Volume 190.
39. Burov, O.; Butnik-Siversky, O.; Orliuk, O.; Horska, K. Cybersecurity and Innovative Digital Educational Environment. *Inf. Technol. Learn. Tools* **2020**, *80*, 414–430. [CrossRef]
40. Maqsood, S.; Chiasson, S. Design, Development, and Evaluation of a Cybersecurity, Privacy, and Digital Literacy Game for Tweens. *ACM Trans. Priv. Secur.* **2021**, *24*, 1–37. [CrossRef]
41. Aranda-Jiménez, J.R.; Campos-García, I.; Cosculluela-Martínez, C.; Martin, J.S.; De-Pablos-heredero, C. Continuous Vocational Training in Response to the Challenge of Industry 4.0: Required Skills and Business Results. *J. Ind. Eng. Manag.* **2023**, *16*, 319–341. [CrossRef]
42. Bica, I.; Unc, R.L.; Țurcanu, Ș. Virtualization and Automation for Cybersecurity Training and Experimentation. In Proceedings of the Innovative Security Solutions for Information Technology and Communications: 13th International Conference, SecITC 2020, Bucharest, Romania, 19–20 November 2020; Volume 12596 LNSC.
43. Rajamäki, J.; Beltempo, E.; Karvonen, J. ECHO Cyber-Skills Framework as a Cyber-Skills Education and Training Tool in Health and Medical Tourism. *Eur. Conf. Cyber Warf. Secur.* **2022**, *21*, 434–437. [CrossRef]
44. Gonzales, R.; Almacen, R.M.; Gonzales, G.; Costan, F.; Suladay, D.; Enriquez, L.; Costan, E.; Atibing, N.M.; Aro, J.L.; Evangelista, S.S.; et al. Priority Roles of Stakeholders for Overcoming the Barriers to Implementing Education 4.0: An Integrated Fermatean Fuzzy Entropy-Based CRITIC-CODAS-SORT Approach. *Complexity* **2022**, *2022*, 7436256. [CrossRef]
45. Ksiezopolski, B.; Mazur, K.; Miskiewicz, M.; Rusinek, D. Teaching a Hands-On CTF-Based Web Application Security Course. *Electronics* **2022**, *11*, 3517. [CrossRef]
46. Alotaibi, A.; Rassam, M.A. Adversarial Machine Learning Attacks against Intrusion Detection Systems: A Survey on Strategies and Defense. *Future Internet* **2023**, *15*, 62. [CrossRef]
47. Zwarts, H.; Du Toit, J.; Von Solms, B. A Cyber-Diplomacy and Cybersecurity Awareness Framework (CDAF) for Developing Countries. *Eur. Conf. Cyber Warf. Secur.* **2022**, *21*, 341–349. [CrossRef]
48. Deák, V. Simulation Framework for Practical Cyber Security Training in the Public Service. *Secur. Def. Q.* **2021**, *33*, 87–104. [CrossRef]
49. Payne, B.K.; Mayes, L.; Paredes, T.; Smith, E.; Wu, H. Applying High Impact Practices in an Interdisciplinary Cybersecurity Program. *CrimRxiv* **2021**, *2020*, 18. [CrossRef]
50. Glas, M.; Vielberth, M.; Pernul, G. Train as You Fight: Evaluating Authentic Cybersecurity Training in Cyber Ranges. In Proceedings of the Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023.
51. Domínguez, M.; Fuertes, J.J.; Prada, M.A.; Alonso, S.; Morán, A.; Pérez, D. Design of Platforms for Experimentation in Industrial Cybersecurity. *Appl. Sci.* **2022**, *12*, 6520. [CrossRef]