Journal of Posthumanism
2025
Volume: 5, No: 7, pp. 1952–1979
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v5i7.2995

CyberCompanion: Enhancing Cybersecurity Awareness through AI-Powered Training and Personalized Learning

Murad Al-Rajab[1], Fathi M. AlKhatib[2], Ahmad Hasna[3], Ghanem AlMarar[4], Ahmed Yasser Sharif[5]

Abstract

The leading cause of cyber-attacks is the human error, which is caused by the lack of awareness in cybersecurity within organizations, this raises the importance for a solution for cybersecurity unawareness, especially within non-technical staff, and staff working outside the scope of cybersecurity. The training solutions offered in the present time fall short in terms of being clear and engaging to learners, plus they do not provide any personalization or adaptation to their learning levels. To solve this problem, we propose CyberCompanion, an innovative solution that provides a complete educational platform to empower employees with the necessary cybersecurity practical knowledge. CyberCompanion uses a generative AI advisor, which serves as a "companion" for the employees whenever they need advice or clarification about anything cybersecurity related. This generative AI cybersecurity advisor is called SIDCA, which is short for Smart Integrated Digital Cybersecurity Advisor, which utilizes advanced Natural Language Processing (NLP) techniques. Additionally, CyberCompanion enhances user engagement by using a personalized course recommendation algorithm, which also ensures every single learning gap is covered. As for the platform itself, we have implemented complete and advanced security measures to ensure employee data is safeguarded from unauthorized access. Such measures include Multi Factor Authentication (MFA) for users when logging in, Transport Security Layer (TLS) for secure communication, and hashing and salting, for password storage security and encryption. CyberCompanion proved to be highly engaging due to the interactive content for the employees, in addition to CyberCompanion having an interface for team leads to track the employees' performance, to ensure everyone is on track with their cybersecurity awareness leaning path.

Keywords: *Cybersecurity, AI-Powered Training, Personalized Learning, Cybersecurity Awareness; Cybersecurity Education; Generative AI in Training; Natural Language Processing (NLP) in Education.*

Introduction

Understanding the essential concepts in cybersecurity remains complex, especially to people outside the scope of Information Technology. The foundations of cybersecurity often appear intimidating to employees due to the field being extensively technical and ever evolving each day, this makes a gap within employees' awareness in cybersecurity, which current technologies fail to address. It is important to note that employees' cybersecurity awareness is an integral part of the organization's security posture, because employees of all levels and positions hold important responsibilities and access to certain assets. This lack of awareness proves to be much more dangerous within employees of higher positions, such as administrators, or CEO's.

[1] College of Engineering, Abu Dhabi University Abu Dhabi, UAE, Email: murad.al-rajab@adu.ac.ae
[2] College of Engineering, Abu Dhabi University Abu Dhabi, UAE, Email: 1078629@alumni.adu.ac.ae
[3] College of Engineering, Abu Dhabi University Abu Dhabi, UAE, Email: 1081911@alumni.adu.ac.ae
[4] College of Engineering, Abu Dhabi University Abu Dhabi, UAE, Email: 1082287@alumni.adu.ac.ae
[5] College of Engineering, Abu Dhabi University Abu Dhabi, UAE, Email: 1079518@alumni.adu.ac.ae

Cybersecurity is a fast-growing field, which poses a problem; non-specialists struggle to keep up with the latest cybersecurity trends. And this is especially riskier in large organizations, because if a single device were to be compromised, it can result in theft of valuable business data, asset losses, and even bankruptcy. According to statistics, 95% of data breaches happened in 2019 due to human error, with phishing and insider threats being the top common methods of cyberattacks [1][2]. This statistic further proves the importance of fostering suitable measures to raise awareness among all employees of an organization. To combat the rapid pace of evolving cybersecurity, a cybersecurity awareness platform should be updated regularly, like OWASP Top 10 list [3].

Key challenges include:

- A significant lack of cybersecurity awareness among employees, which creates vulnerabilities within organizations and increases the risk of exploitation.
- The rapid evolution of cybersecurity threats, making it challenging for non-specialists to stay informed.
- Limited access to tailored and practical resources for newcomers to cybersecurity.

As digital transformation accelerates, the financial and reputational stakes of cybersecurity breaches grow. High-profile cases, such as the $4.4 million ransom paid by Colonial Pipeline in 2021 and the $11 million ransom paid by JBS [4][5], illustrate the devastating potential of cyberattacks. With 94% of malware infections originating from malicious emails [1], robust employee awareness training can dramatically reduce the likelihood of such incidents. Cybercrimes are one of the leading causes of financial losses, which is estimated to be $2 trillion in 2019, which is projected to rise to $6 trillion by 2021, as the digital infrastructure keeps expanding and getting more complex, day by day [6]. This further emphasizes the importance and urgency of a revolutionary cybersecurity awareness solution.

To address the aforementioned issues, we present CyberCompanion, a smart, interactive web application designed to enhance cybersecurity knowledge and resilience among employees. One of the most important components of CyberCompanion is SIDCA, which is the Smart Integrated Digital Cybersecurity Advisor, a generative AI tool that provides users with real-time personal learning experience and advice. CyberCompanion uses multimedia, such as inforgraphics, videos, and quizzes, to ensure user is engaged throughout their learning journey. Additionally, CyberCompanion incorporates periodic assessments to ensure employees learning is going smoothly and their progress is in check, and all of this progress is visualized in an analytics dashboard, to make it easier for employees to follow up and track their progress.

The main contributions of CyberCompanion are as follows:

1. Developed a web-based platform, CyberCompanion, designed to enhance cybersecurity awareness and training among individuals and employees.

2. Implemented SIDCA, the Smart Integrated Digital Cybersecurity Advisor, a generative AI tool that provides real-time assistance and advice to cybersecurity-related queries.

3. Utilized machine learning (ML) algorithms to develop a personalized course recommendation system, tailoring training content based on individual users' strengths and weaknesses.

4. Developed a performance analytics dashboard to monitor and assess the progress of

users, enabling administrators to identify knowledge gaps and provide targeted support.

5. Addressed the critical issue of cybersecurity unawareness by providing an accessible and effective cybersecurity training solution that empowers employees to protect themselves and their organizations from cyber threats

CyberCompanion addresses the major cybersecurity awareness gap by providing employees with ease of access to the latest cybersecurity trends, in the most understandable and engaging way possible. This contributes to the overall security posture of the organization. The following sections outline the development, implementation, and evaluation of CyberCompanion, demonstrating its effectiveness as a transformative tool for cybersecurity education. The organization of the paper is also shown in the attached Figure 1

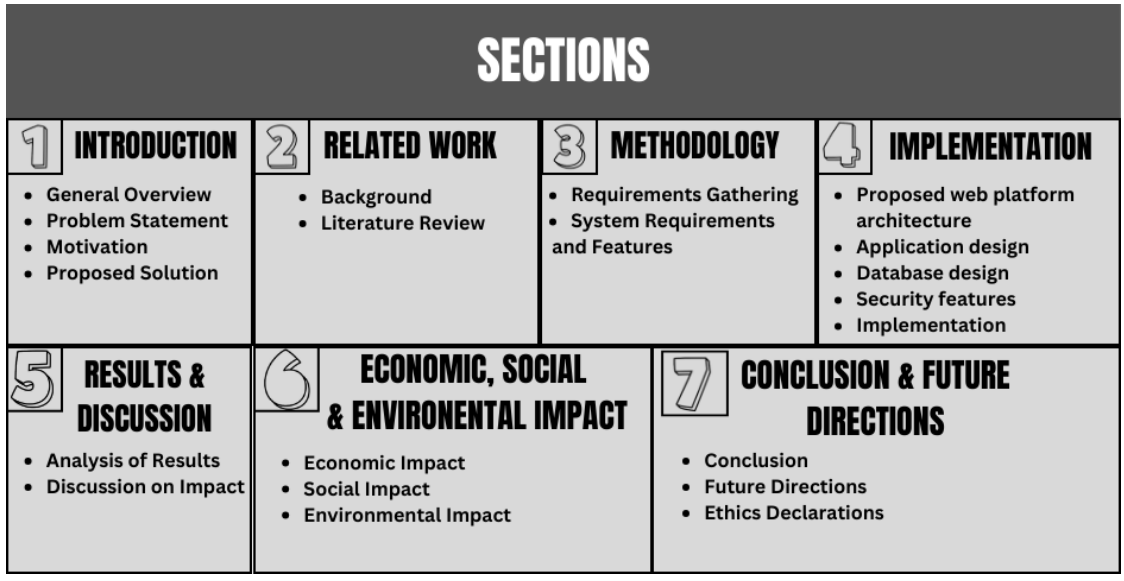

Figure 1. Paper Organization

Related Work

Generative AI enables users to simplify complicated concepts into easily comprehensible segments. This AI technology can create personalized visuals and interactive learning content, providing a seamless experience for employees' learning journey. The feedback and visuals prove to be essential for employees' information retention [7].

Adding generative AI into simulations is an effective way to enhance user understanding and visualization of complex concepts. This approach creates realistic training environments alongside hands-on labs, making it easier for users to grasp the material. The technical domain of generative AI includes various APIs used to implement and integrate numerous machine learning models, each designed for distinct tasks [8]. Developers can utilize a range of APIs, particularly those specializing in natural language processing (NLP), like OpenAI's API for ChatGPT, to create chatbots for human-like conversations [8]. ChatGPT is a powerful AI tool that can generate content, translate languages, and answer questions in an informative way, which assists users with many tasks in all sectors [9]. ChatGPT utilizes natural language processing algorithms to understand human language and generate relevant and informative answers [10]. OpenAI's API allows developers to integrate ChatGPT as a chat interface for any

application, product, or service [11]. ChatGPT uses various NLP algorithms, such as back-propagation, fine-tuning, and language modeling, to collectively pre-train the machine learning model [12].

Opting for cloud server hosting is typically the best choice for application hosting given its scalability, cost-effectiveness, and overall advantages in maintenance; hence, the need for its security arises. Firebase is our choice of web application hosting as it is very simple and more suitable for students [13]. It contains a bundle of options to manually secure a Firebase environment through security rules and policies that ensure the security of the web application's users and resources [14]. These security controls can be implemented alongside secure coding practices to ensure the application's security. HTML, CSS, and JavaScript are the most common front-end languages that developers use, and some frameworks include React, Bootstrap, and AngularJS, which can be used to create responsive web applications. React is useful as it has preset libraries, declarative programming, compatibility with mobile devices, optimized performance, and a large and active community that supports the React framework [15]. Traditionally, e-learning relied on static text pages offering limited functionality. Nowadays, however, e-learning environments have rapidly advanced through more dynamic pages using engaging multimedia, personalized learning, and even immersive virtual reality experiences [16].

The rapid digitization of various businesses highlights the vital requirement for comprehensive cybersecurity education and awareness. To properly protect sensitive data, training approaches must constantly adapt to the evolving risks posed by cyberspace. Many creative methods have been investigated in the field of cybersecurity education to improve student learning and recall of cybersecurity concepts. In order to improve high school students' grasp of cybersecurity principles, G. Jin et al. [17] evaluated game-based learning using cybersecurity games created using Unity3D. Similarly, B. D. Cone et al. [18] introduced "CyberCIEGE," an interactive game designed for the U.S. Navy that creates a realistic cybersecurity training environment by employing a special simulation engine. Though their often narrow scope means that they may not fully cover all educational objectives, these immersive and interactive techniques have demonstrated potential in enhancing engagement and learning results. In their investigation of the idea of cyber defense as a system of collective action, S. M. Ho et al. [19] argued in favor of interactive learning environments as a means of promoting a cybersecurity-aware culture inside businesses. This is consistent with the results of the authors in [20], who showed that by imitating actual cybersecurity dangers and countermeasures, serious games such as Riskio may greatly increase the awareness of non-technical personnel. B. Srinivasa-Desikan [21] introduced a hybrid recommendation system that integrates collaborative filtering with the New Apriori algorithm to address the issues of data sparsity and the cold-start problem in customized learning settings. Learning materials may be more customized thanks to the incorporation of Natural Language Processing (NLP) technology like spaCy, which supports various languages and makes processing pipelines more efficient and scalable [22]. Additionally, cloud-based settings such as Google Colaboratory make machine learning models easier to deploy and test, making these technologies available for instructional use.

There are still shortcomings in the way these technologies are applied to cybersecurity education, notwithstanding these developments. Since many of the solutions available today are geared toward certain demographics, their applicability in larger contexts may be limited. Additionally, it might be difficult for educators to keep the curriculum relevant because cyber risks are evolving faster than educational content is being updated. Our proposed solution,

"CyberCompanion," aims to bridge the gap between theoretical knowledge and practical application by integrating these state-of-the-art educational technologies and techniques into a complete and adaptable learning environment. CyberCompanion uses generative AI to adjust to each learner's pace and learning style. This means that all staff members, no matter how much experience they have, get customized, pertinent training that equips them to reduce cybersecurity risks effectively. The body of research highlights the potential for individualized and interactive learning strategies to improve cybersecurity education. To keep up with the rapidly changing landscape of cyber risks, instructional content must be updated and modified on a regular basis. By offering a dynamic, scalable, and efficient training solution that evolves with the newest developments in cybersecurity, CyberCompanion aims to address these issues.

Methodology Approach and Data Analysis

Requirements Gathering:

This section provides a detailed description of the data analysis methodology and the requirements gathering tools along with the functional requirements of the proposed system.

We have launched two surveys for employees from several companies inside the UAE. The respondents consisted of employees, and team leads (such as managers, supervisors, and team leaders) working in the Information Technology cybersecurity field. We have asked the respondents about their experience and knowledge on cybersecurity, so we can further understand their demands, to collect more information about the requirements and features for our proposed system. Table 1 shows a summary of the requirements-gathering techniques that we have used.

Name of Study	Sample	Nature of Analysis	Statistics Used
Employee Survey	110	Quantitative	Google Forms
Team Lead Survey	12	Quantitative	Google Forms

Table 1. Requirements Gathering Illustration

The Employee Survey exposed some key findings. About 53% of participants mentioned they had received cybersecurity awareness training, though it did not include AI. A significant 80% of employees consider cybersecurity crucial for the organization's operations and data security. Additionally, 57% of respondents had either experienced or knew someone who had faced an attempted data theft. There is also strong interest in AI, with 80% of participants expressing a need for an AI cybersecurity chatbot to help them learn about cybersecurity more easily. Many responses stressed the importance of raising awareness and providing more cybersecurity training. On the other hand, the Employee Survey results found that email warnings are the most commonly mentioned method for assessing cybersecurity awareness in the workplace. Leaders noted that ensuring compliance with cybersecurity policies is a significant challenge. Additionally, team leaders strongly recommended providing employees with greater access to scanning and antivirus tools to enhance the organization's security.

System Requirements and Features

The main features and requirements of proposed system are explained in the table 2:

Requirement	Description
Real-Time Advisor "SIDCA"	A generative AI advisor available on-demand to answer questions, clarify doubts, and guide users through complex cybersecurity scenarios. Using SIDCA, the AI-driven advisor, employees are guided through a structured, step-by-step process for managing various security incidents. SIDCA's role is to foster a deeper, intuitive understanding of cybersecurity as employees progress through their training.
Course Recommendation System	A course recommendation algorithm utilizing natural language processing based on employee performance and assessment results. For instance, if an employee is weak in Phishing training, the system suggests courses based on that specific weakness.
Training	Employees begin their cybersecurity education through interactive learning modules, which include videos, knowledge checks, assessments, and labs. These modules cater to varying levels of prior knowledge and aim to progressively develop a comprehensive understanding of cybersecurity principles and practices.
Assessment	The CyberCompanion application evaluates the employees' cybersecurity awareness knowledge. After completion, it provides comprehensive insights of each user's knowledge gaps and strengths and can be scheduled periodically for reinforcement.
Performance Analytics	CyberCompanion provides a performance analytics dashboard for holistic overview, enabling team leads and system administrators to track the learning's effectiveness. These insights allow the ease of knowledge gap identification while ensuring resources are allocated effectively.
Password Strength	Provides real-time evaluation of users' password strength, immediately notifying them if their chosen password is weak. Feedback includes guidelines for stronger passwords, emphasizing length, complexity, and character variety.

Table 2. Functional Requirements of the Proposed System

Design and Implementation

This section discusses the proposed system architecture, software design, AI models, and the implementation.

The Proposed Web Platform Architecture

Figure 2 illustrates the general architecture of our proposed smart cybersecurity interactive web application, which integrates an AI-powered Digital Cybersecurity Advisor to enhance user knowledge and skills in combating cyber threats. This platform is structured with distinct roles: Employee, Team Lead, System Admin, and Cyber Specialist, each accessing tailored features through a user interface. The interface offers training content, progress tracking, analytics, account management, and content delivery. A core feature, the Digital Cyber Advisor, connects the user interface to backend technologies, including a secure server, ChatGPT API, and Firebase, ensuring secure data handling, user management, and real-time analytics. This setup supports personalized training modules, adaptive learning, and role-specific access control, promoting cybersecurity awareness and skills across different organizational roles.

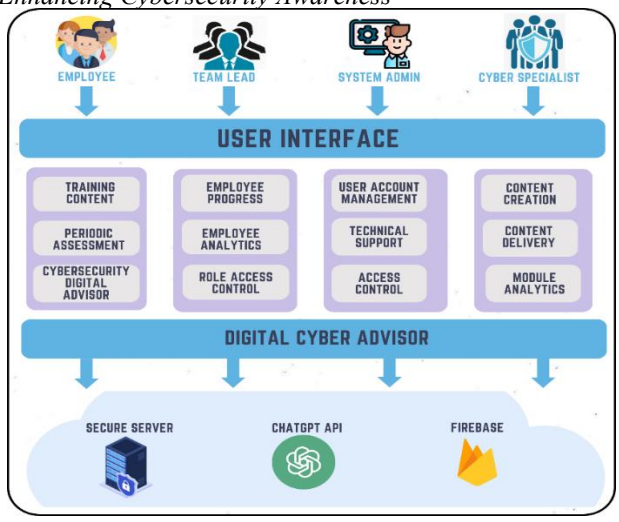

Figure 2. Web App Architecture

As shown in Figure 3, the CyberCompanion web platform's system design utilizes three-tier client-server architecture. The front end, back end, and database all make up the architecture, each with a distinct role in simplifying cybersecurity training. The front-end contains user-friendly interfaces for various roles, including employees, team leads, system admins, and cybersecurity specialists. The back end manages data processing and interactions with APIs between the application and database to provide training modules, skill assessments, and course recommendations. The database enables the efficiency of data storage and retrieval from within the application. Employees interact with CyberCompanion through their main dashboard, skill assessments, and chatbot, while team leads can monitor their progress and assign each individual employee assessments. Cybersecurity specialists manage and update training content based on the latest trends whereas system admins oversee all the identity and access management of the organization. This architecture allows efficient task execution and a holistic foundation for the platform. The proposed system workflow starts with the secure login of employees using multi-factor authentication, followed by an initial skill assessment to identify their cybersecurity awareness gaps. Based on assessment results, the AI-powered course recommendation algorithm suggests tailored courses targeting those specific gaps. Interactive training modules would contain videos, quizzes, infographics, and labs, with periodic assessments to assess progress. Throughout the learner's journey, employees can interact with SIDCA for real-time assistance, especially guidance and clarifications. Each employee's progress has performance data, such as training completion, assessment scores, and identified weaknesses which are tracked securely by analytics on the team lead's dashboard. Team leads can monitor employees' progress, assign skill assessments and additional courses, as well as address team-wide knowledge gaps. Cybersecurity specialists maintain the system by creating and updating courses. System administrators oversee user account management and ensure robust security measures. This dynamic, role-based system fosters personalized learning while promoting organizational cybersecurity resilience.

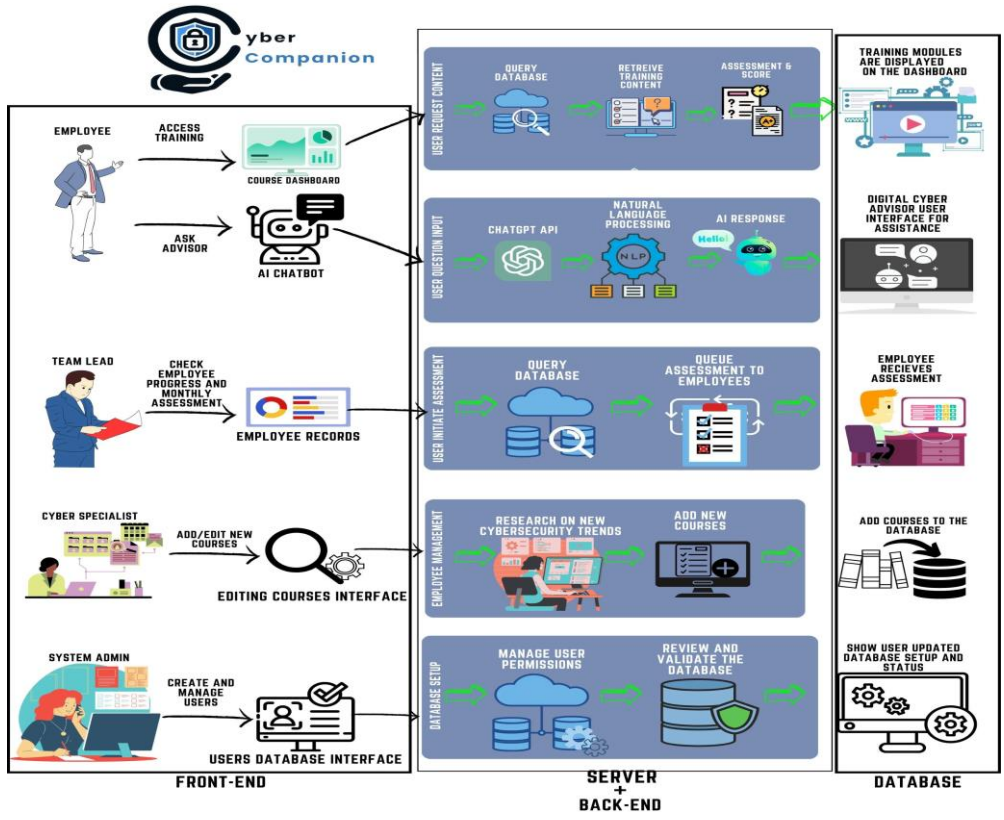

Figure 3. System Design Workflow

Application Design

Real-Time Advisor "SIDCA" using Generative AI

One of the main features in our proposed application is the Cyber companion "SIDCA" which allows employees to interact in real-time with a generative AI cybersecurity AI advisor chatbot about their cybersecurity concerns or prompts. The entire process of the generative AI feature is shown in Figure 4. The web application will initially collect the user's response and send the question into ChatGPT's API using an HTTP request with the prompt from the user, among other things. Upon receiving the response from the ChatGPT API, the web application parses the data in JSON format to extract the generated text and processes it even further to ensure that it is clear and meaningful. This is done through a series of tasks such as detokenization, embedding, feeding them into their machine learning model, generating the response, and then finally detokenizing the response to convert the tokens back into text [23]. This text is then sent using the API to our web application, which will be displayed to the user to view as an answer to his concern or prompt.

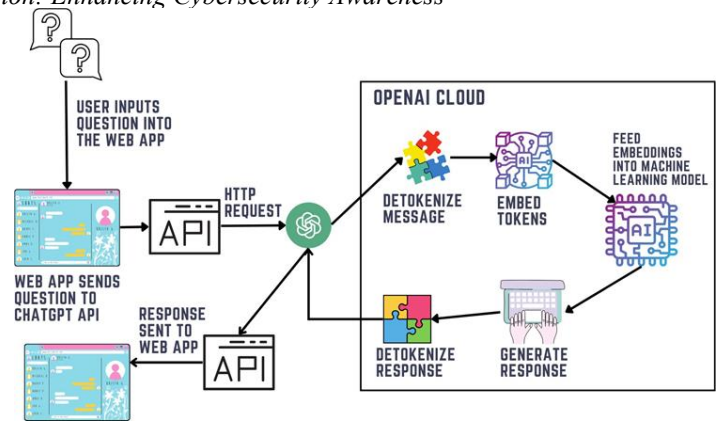

Figure 4. ChatGPT's Generative AI Illustration

Course Recommendation System Design

The Recommendation System is initiated by the cybersecurity specialist who is responsible to creating or updating courses and assessment questions, ensuring the database contains up-to-date materials. The employees then take an assessment to evaluate their cybersecurity knowledge, and the system analyzes their performance to identify weaknesses. Using these results, the system dynamically matches the employee's weak areas with course content and recommends personalized training modules. This ensures a personalised learning experience based on real-time assessment data.

A) Dataset Description:

The "Coursera Courses Dataset 2021" dataset [24] was used to enrich the SpaCy model which is available on Kaggle and contains around 3,400 unique records with 7 attributes. The dataset is proven to be valuable in ML as it has a usefulness of 8.24 out of 10 on Kaggle with many users utilizing it for their applications [24]. A sample of the dataset is illustrated in Table 1. To ensure consistency, we dropped the university, difficulty level, course rating, course URL, and skills attributes while keeping only the course name and course description attributes as the algorithm will be used to input the course name and description from the courses in Firestore to generate the embeddings.

Course Name	University	Difficulty Level	Course Rating	Course URL	Skills
Write A Feature Length Screenplay For Film Or Television	Michigan State University	Beginner	4.8	https://www.coursera.org/learn/write-a-feature-length-screenplay-for-film-or-television	Drama Comedy peering screenwriting film Document Review dialogue
Business	Courser	Be	4.	https://www.coursera	Finance business plan

Strategy: Business Model Canvas	a Project Network	ginner	8	.org/learn/canvas-analysis-miro	persona (user experience) business model canvas Planning
Silicon Thin Film Solar Cells	cole Polytechnique	Advanced	4.1	https://www.coursera.org/learn/silicon-thin-film-solar-cells	chemistry physics Solar Energy film lambda calculus Electrical Engineering
Finance for Managers	IESE Business School	Intermediate	4.8	https://www.coursera.org/learn/operational-finance	accounts receivable dupont analysis analysis Accounting Finance Operations
Retrieve Data using Single-Table SQL Queries	Coursera Project Network	Beginner	4.6	https://www.coursera.org/learn/single-table-sql-queries	Data Analysis select (sql) database management systems online
Building Test Automation Framework using Selenium and TestNG	Coursera Project Network	Beginner	4.7	https://www.coursera.org/learn/building-test-automation-framework-using-selenium-and-testng	maintenance test case test automation screenshot project helper class selenium reusability debugging php
Doing Business in China Capstone	The Chinese University of Hong Kong	Advanced	3.3	https://www.coursera.org/learn/doing-business-in-china-capstone	marketing plan Planning Marketing consumption (economics) wait (system call) business plan
Programming Languages, Part A	University of Washington	Intermediate	4.9	https://www.coursera.org/learn/programming-languages	inference ml (programming language) higher-order function functional
The Roles and Responsibilities of Nonprofit Boards of Directors	The State University of New York	Intermediate	4.3	https://www.coursera.org/learn/nonprofit-gov-2	Planning Peer Review fundraising strategic planning resource concept testing leadership
Business Russian Communication. Part 3	Saint Petersburg State University	Intermediate	Not Calibrated	https://www.coursera.org/learn/business-russian-communication-3	Russian market (economics) tax exemption cooperation Communication

Table 1: The Coursera Courses Dataset Sample

B) Dataset Preprocessing:

Preprocessing is a critical step in NLP workflows, transforming raw data into a structured and clean format suitable for analysis. For example, given the input text “Phishing emails can be detected using detection methods!”, the preprocessing process begins with tokenization, which breaks down the text into individual units or tokens, such as words. This segmentation enables a focused analysis of text fragments. Next, all tokens are converted to lowercase to ensure the NLP model treats variations like "Learn" and "learn" as identical, eliminating conflicts caused by case differences and ensuring uniformity. Following this, lemmatization is applied, reducing tokens to their base forms while maintaining grammatical context. For instance, “detection” and “detected” are both reduced to “detect.” This step minimizes redundant word forms while preserving the semantic meaning of the text. Additionally, all punctuation and special characters are removed to reduce noise and focus the analysis on relevant words, improving semantic interpretation. Stop words (e.g., “the,” “is,” “at”) that add little value to understanding the content are also filtered out. These steps are then applied to preprocess all course names and descriptions retrieved from the Firestore database. This ensures the embeddings accurately capture the most semantic meaning of the text for improved analysis and embedding generation. The entire workflow ensures data accuracy, semantic coherence, and relevance for machine learning models, as illustrated in Figure 6.

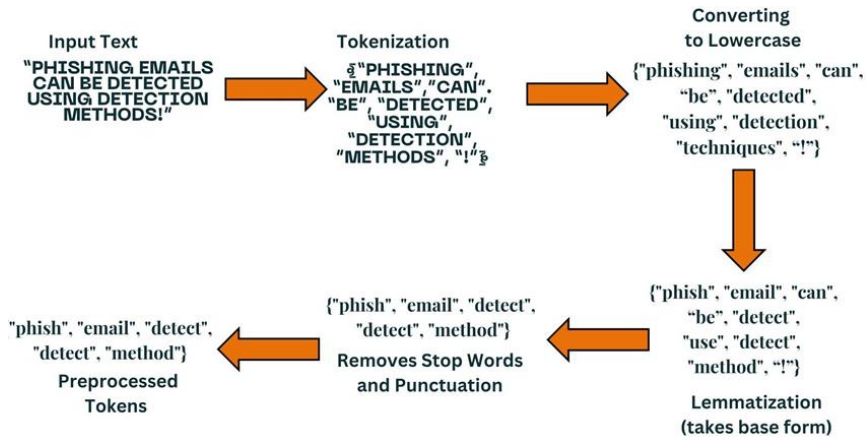

Figure 6. Preprocessing Illustration

C. NLP and Recommendation System

The design of the course recommendation system utilizes NLP methods to generate personalized course recommendations based on the performance of the employees in assessments, specifically targeting areas of identified weakness. The process begins with data preprocessing described earlier, where raw data from course titles and descriptions are cleaned and standardized. The pre-processed tokens are then processed using SpaCy, an efficient and advanced open-source NLP library [25], to generate word vectors using pre-trained models. SpaCy provides powerful tools like tokenization, part-of-speech tagging, named entity recognition, and dependency parsing, forming the foundation for understanding text structure and building semantic relationships [26]. Additionally, SpaCy supports advanced functionalities like word vectors, similarity detection, and text classification, making it ideal for tasks such as semantic search,

clustering, and recommendation systems. This overall process forms the foundation of the course recommendation system and is illustrated in Figure 7.

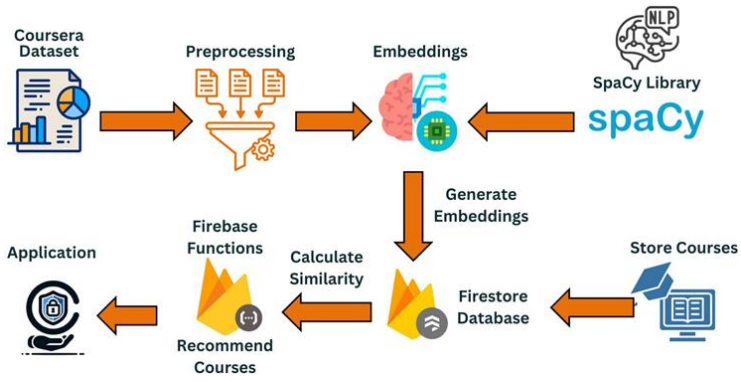

Figure 7. Course Recommendation Model Illustration

These word vectors, or embeddings, encapsulate the semantic relationships within the input data by transforming textual information into high-dimensional numerical vectors. Once generated, the embedding of the sentence is averaged and stored in the Firestore database alongside corresponding course information, as shown in Figure 8. To identify and recommend relevant courses, the system calculates similarity scores based on the Euclidean distance between embeddings, with scores ranging from 0 to 1. A higher score indicates greater similarity, enabling the system to suggest courses most aligned with employees' learning needs.

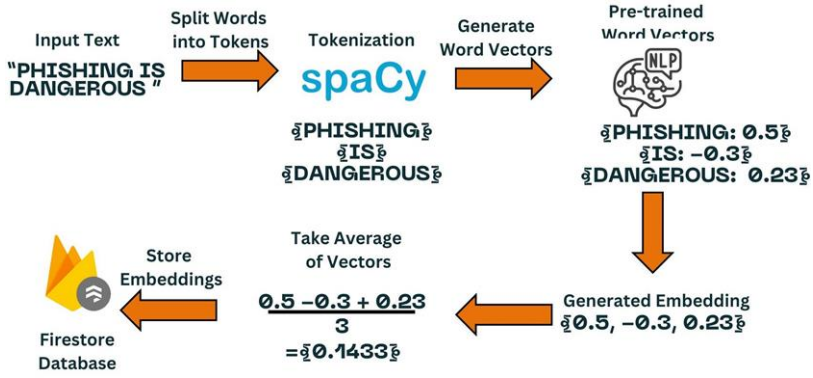

Figure 8. Embedding Generation Illustration

Employees Performance Analytics

The entire process of the employees' performance that team leads can use in our web application is shown in Figure 9. Assuming the team lead is logged in, he can request the performance of a desired employee. After initially requesting the performance data, the database searches for the requested data and then retrieves the employees' performance data. The performance data include but are not limited to the employee performance score, employee profile, number of training modules taken, skill assessment results, areas of improvement, training module duration, and completion status. After the data is retrieved from the database, the data is sent to the front-end user interface and displayed in a user-friendly document that the team lead can download.

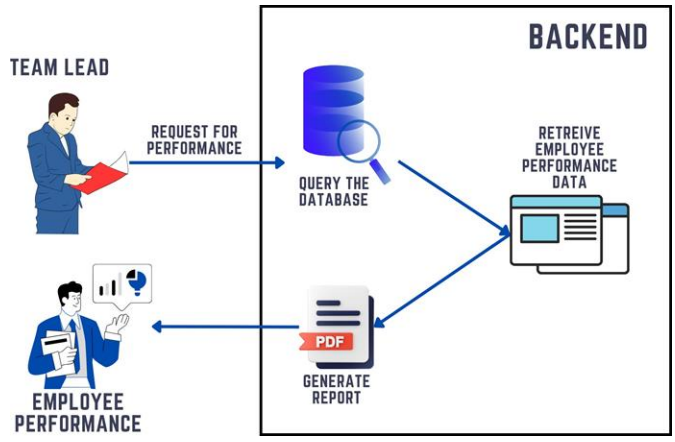

Figure 9. Employee Performance Illustration

Database Design

To ensure a robust data management system, we designed a secure and scalable NoSQL database using Firebase Firestore. The later utilizes a NoSQL structure, enabling real-time updates, which makes it efficient for our application. Its flexibility allows for easy modification and extension of the data structure without being constrained by a fixed schema. Firestore will be used to store user credentials, such as email addresses and hashed passwords, along with courses, skill assessments, and user performance. Data is organized into collections and documents, enabling a flexible and scalable schema. Each time a user attempts to log in, the system verifies their identity by checking the hashed password stored in the database for authentication and authorization. For added security, we securely store the multi-factor authentication (TOTP) secrets in the Firestore database for each user, enabling further authentication through the Google Authenticator app. The use of this database ensures system security while supporting the accessibility of personalized learning which enables the system to remain secure and the personalized learning to be accessible.

Security Features:

Multi-factor authentication (TOTP)

After the user has entered their correct credentials, they are sent an authentication code to their mobile phone via the Google Authenticator app. The user is then given a TOTP (Time-based One Time Passcode) in their authenticator app that changes every 30 seconds. If the code entered in the login page matches, the user is authenticated. This method is quite a popular choice as it requires the user to provide two or more pieces of evidence to prove their identity to log in.

One-way users can Configure TOTP MFA, is by installing Google Authenticator, then scanning a QR code from the app they want to log into, then, the account is set up on Google Authenticator, and the user now can receive the TOTP on the app. However, another security concern, is that users should be careful when scanning QR codes, as there's a risk of falling into a Qishing attack, meaning that the user might lose access to their account, we should only scan QR codes from trusted sources. TOTP MFA is explain in Figure 10.

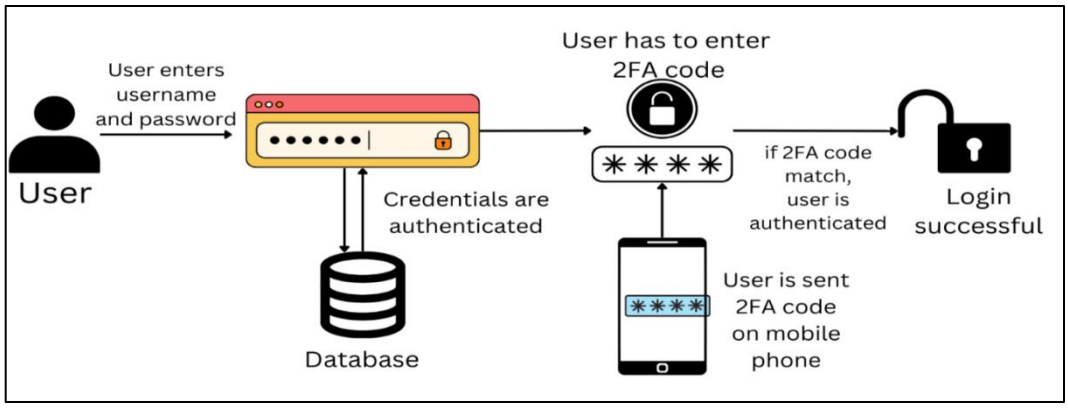

Figure 10. Multi-factor Authentication using TOTP

Compliance with OWASP Top 10 standards.

The OWASP Top 10 list is illustrated in Figure 11, and is described as follows:

- Broken Access Control: Use least privilege and RBAC to restrict user access and minimize breach impact.
- Cryptographic Failures: Use secure encryption algorithms like SHA-256 to prevent vulnerabilities.
- Injection: Sanitize inputs to block harmful characters and prevent attacks like SQL injection and XSS.
- Insecure Design: Follow security principles, perform regular assessments, and update systems.
- Security Misconfiguration: Conduct regular audits to identify and fix vulnerabilities.
- Vulnerable Components: Keep components updated and test for compliance with security standards.
- Authentication Failures: Use secure methods like Two-Factor Authentication and strong password hashing.
- Data Integrity Failures: Ensure data integrity with cryptographic hashes and digital signatures.
- Logging Failures: Monitor logs for anomalies and report suspicious activity.
- SSRF Prevention: Validate and sanitize inputs to block malicious server-side requests.

Figure 11. OWASP Top 10 list

Data encryption strategies.

we can implement TLS (Transport Security Layer). When the user initiates the communication, the user requests an encrypted certificate from the server, and then the user verifies the legitimacy of the digital certificate. After the certificate is verified, the user sends back an encrypted key to the server. The server can decrypt the key and send encrypted traffic, completing the TLS Handshake. However, we should note that this falls into the risk of Man-in-the-middle attack or knows nowadays as On-path attack. This type of cyber-attack happens when a hacker intercepts traffic and steals login tokens, or encryption keys, which steals the identity of the intended user. To mitigate this risk, users must ensure that the communication channel is encrypted, so that when an attacker intercepts it, the data will be unreadable and with no use to the attacker, ensuring that our communication is safe. TLS is visualized in Figure 12.

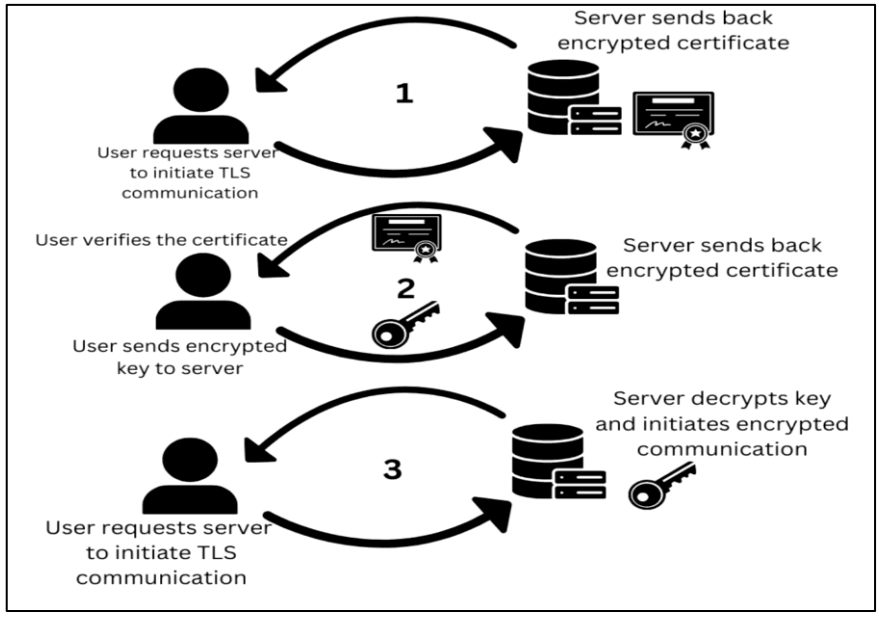

Figure 12. TLS explained.

Implementation

Development Environment

The development environment for the proposed system incorporates various tools and technologies to ensure smooth implementation. Visual Studio Code serves as the primary IDE. Firebase, a comprehensive framework by Google, is used for its NoSQL database, hosting features, analytics, and scalability. Google Authenticator is integrated for two-factor authentication, providing time-based one-time passcodes for enhanced security. Google Colaboratory, a cloud-based platform, supports Python coding for data analysis and machine learning, aiding the course recommendation system. OpenAI's GPT API powers SIDCA, the interactive cybersecurity advisor chatbot.

Multiple APIs and libraries are utilized to meet user requirements and maintain security. Firebase APIs, including Firestore, Firebase functions, Firebase storage, and Firebase tools, handle authentication, data storage, and serverless function execution. Firestore's real-time NoSQL database enables data syncing and offline access, while Firebase functions execute backend code in response to events. Firebase storage manages secure uploads and downloads of user-generated content, and Firebase tools provide a CLI for deploying and managing applications. OpenAI's GPT-3.5 Turbo API enables SIDCA to offer human-like interactions and cybersecurity guidance. Node.js APIs such as bcryptjs for password hashing and salting and speakeasy for generating and verifying time-based passcodes add further security.

An Illustrative Example:

The application begins with the login page where the employee is prompted to enter their login credentials in order to log in. There is no registration as we assume that the system admin creates the users for the organization and gives out the users to employees with default passwords that must be changed once they log in. After successfully authenticating their password, they will be moved to the two-factor authentication page which will prompt the employee to enter the time-based one-time password that is sent to their google authenticator app.

After successfully authenticating their user account, the employee is shown their dashboard in which they have multiple options including navigating to the SIDCA page, checking their recommended courses based on their results from their previous skill assessment, and checking the currently assigned skill assessments they have to complete that are assigned by their respective team lead.

The employee can choose which path to use out of the three. Upon clicking on the ask SIDCA button, the employee will be sent to the SIDCA AI advisor in order to receive assistance. Upon selecting the recommended course out of the options that are recommended by the NLP recommendation algorithm, the employee will be redirected to the respective course details page in which they can view the course details and then proceed to start the course. Upon selecting the assigned skill assessment, the employee will be redirected to conduct the skill assessment in which they will be tested on their cybersecurity knowledge through a series of multiple-choice questions based on a variety of topics.

After completing the courses, the employee will be able to view that they have completed the course. After completing the skill assessment, the employee is shown their score as well as recommended courses based on the incorrect questions, they have completed which demonstrates room for improvement for specific topics they are weak in. The employee can also

view these recommended courses in their dashboard. After the completion of the skill assessment questions, the recommended courses will be displayed enabling the employees to improve on their weaknesses dynamically, effectively enhancing their cybersecurity posture and making it simple to access. The employee user journey is shown in Figure 13 and the cybersecurity specialist journey is shown in Figure 14.

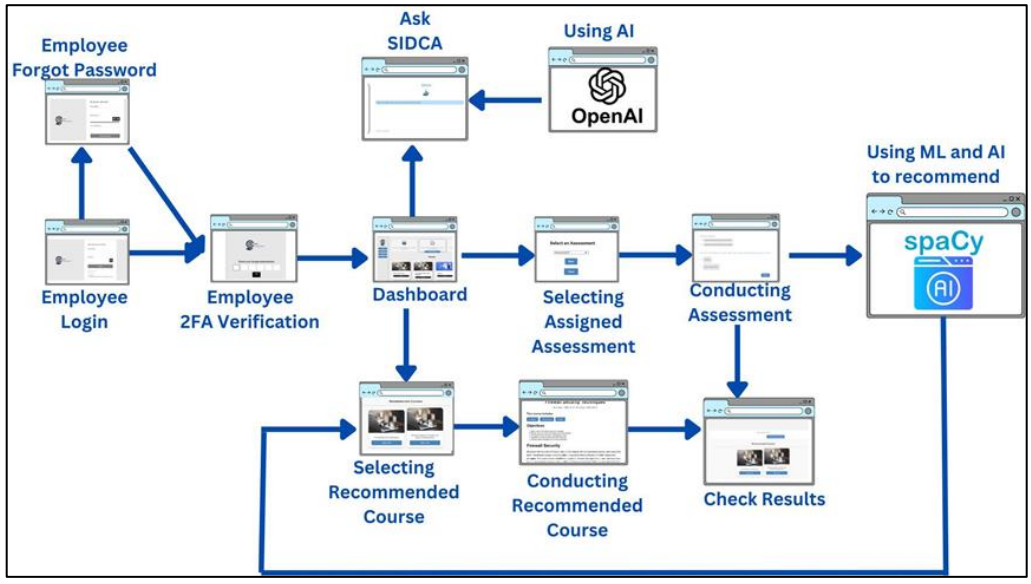

Figure 13. Employee User Journey

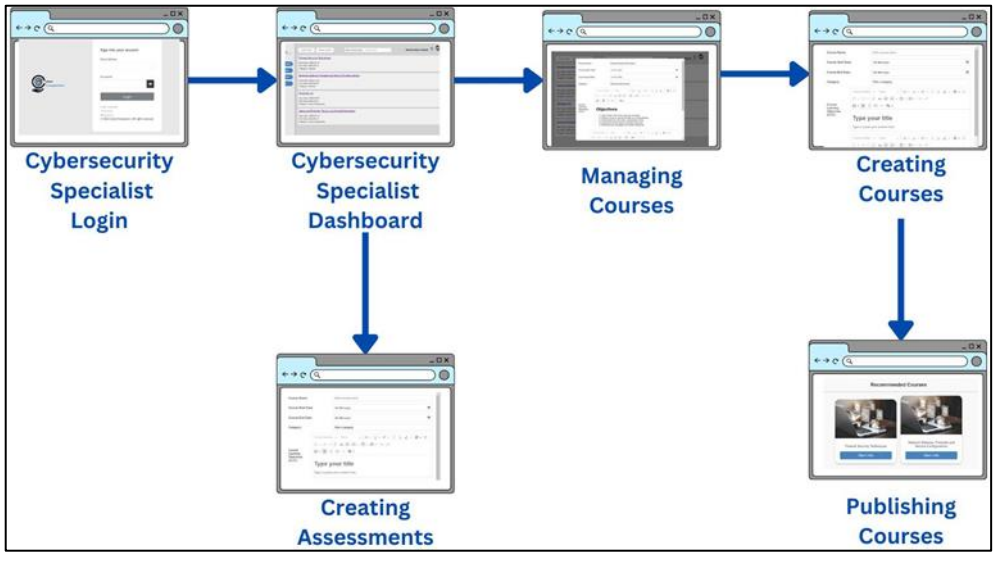

Figure 14. Cybersecurity Specialist User Journey

The screen in Figure 15 illustrates the main dashboard for the cybersecurity specialist. It shows the main options such as creating course, managing course content, searching for course created, assessments, and updating profile. On the other hand, Figure 16, illustrates the process when the cybersecurity specialist creates a course. He must fill in options such as the course title, start and

end dates, and category, allowing him to specify which cybersecurity category the course falls under, for instance, Phishing, Physical Security, Malware Analysis, Social Engineering, Network Security, Incident Response, Cloud Security, Mobile Device Security, Secure Coding Practices, or Data Privacy. Additionally, he must define the Course Learning Outcomes (CLOs) to measure the objectives of the course, the course description, and the course content. The content of the course can be created using various multimedia formats, such as text, audio, video, presentations, and attached files.

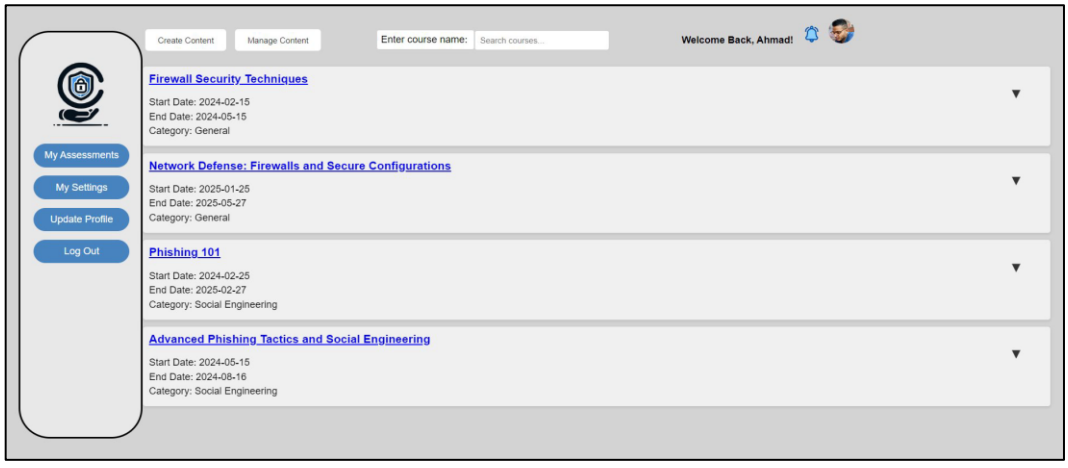

Figure 15. The Cybersecurity Specialist Main Dashboard

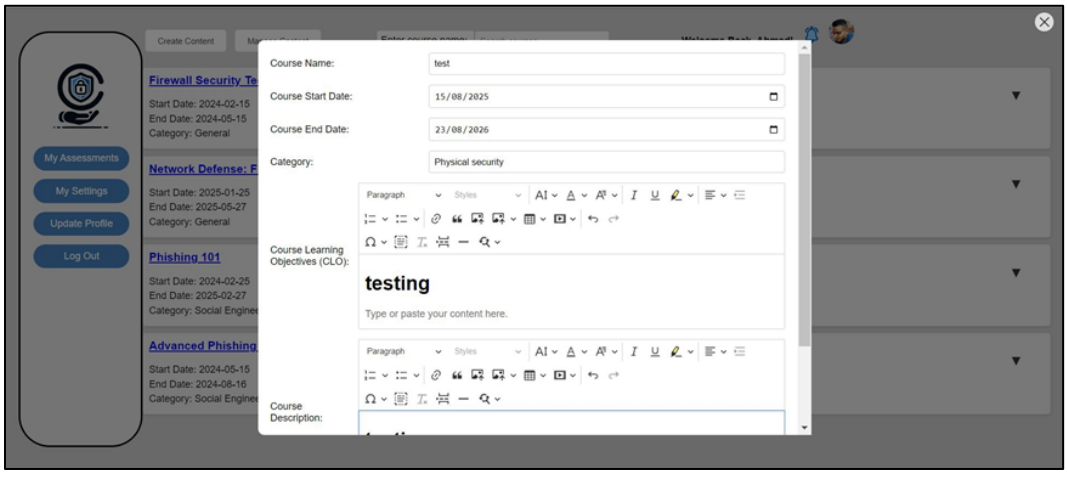

Figure 16. Create Course Content

Upon logging in, the employee is directed to the main dashboard shown in Figure 17, which provides access to key features: SIDCA (Smart Integrated Digital Cybersecurity Advisor), Taking Assessments, Viewing Recommended Courses, and Taking a Course. Each feature is designed to support the learning journey of the employees and enhance their cybersecurity skills.

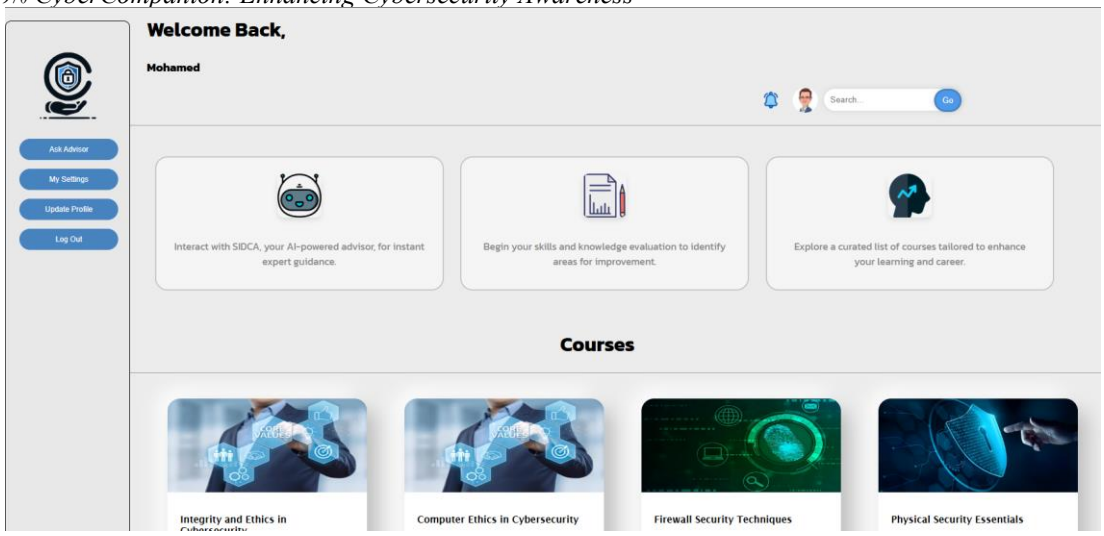

Figure 17. The Employee Main Dashboard

The SIDCA advisor is an interactive chatbot powered by advanced AI and prompt engineering. Employees can ask SIDCA various cybersecurity-related questions, seek guidance on specific topics, or request suggestions for improving their skills. SIDCA has been rigorously tested for robustness against typos, prompts in different languages (English, Arabic, French, and others) and queries outside the scope of proposed system. The testing, as shown in Figures 18, 19, 20, and 20 respectively, demonstrates that SIDCA consistently delivers accurate and relevant responses due to prompt engineering and fine-tuning.

Figure 18. SIDCA User Interface

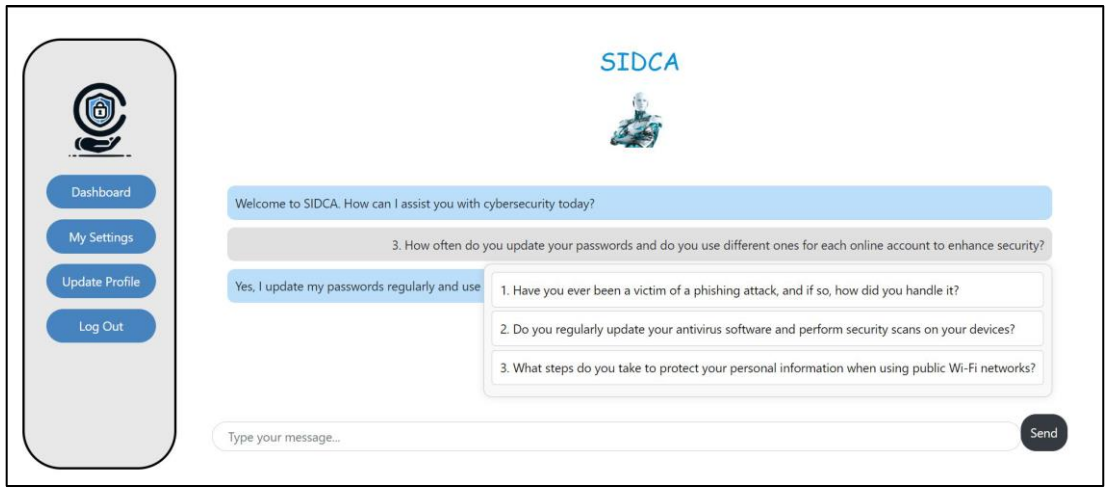

Figure 19. Prompt Suggestions Working

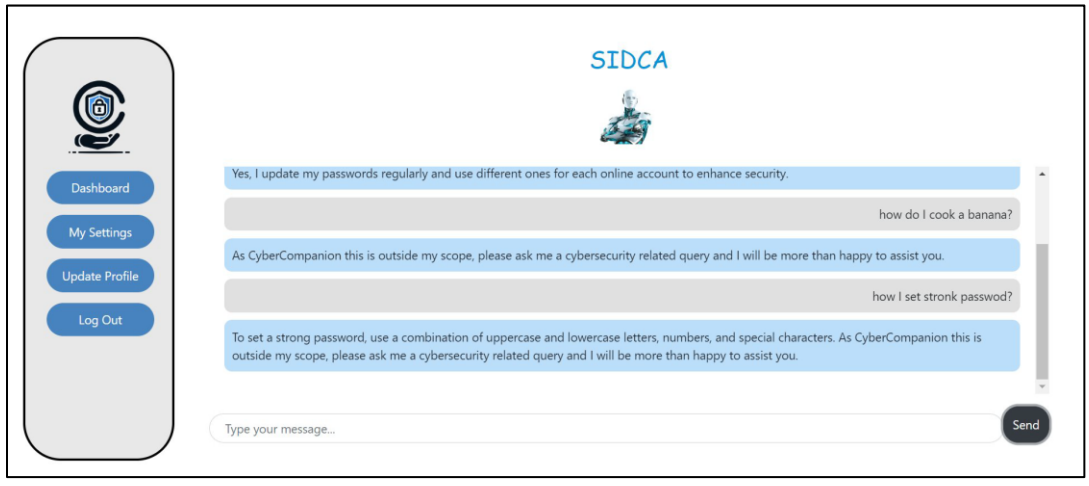

Figure 20. Prompt Evaluated Against Typos and Against Out-Of-Scope

Figure 21. Prompt Evaluated Against Languages

Furthermore, employee can select and initiate assessments directly from the dashboard. As shown in Figure 22, the system allows the employee to choose an assessment, which then proceeds to the test cases. The assessments evaluate the knowledge of the employee and identify areas of weakness by analyzing incorrect answers. These results form the foundation for personalized course recommendations.

The recommendations, displayed alongside the assessment score as illustrated in Figure 23, are dynamically generated by the system. Even if an employee skips the recommended courses initially, they will remain visible on the dashboard for easy access. Employees can also view detailed course information, including descriptions and objectives, retrieved from the database as shown in Figure 24, ensures that the course details are accurate and reliable.

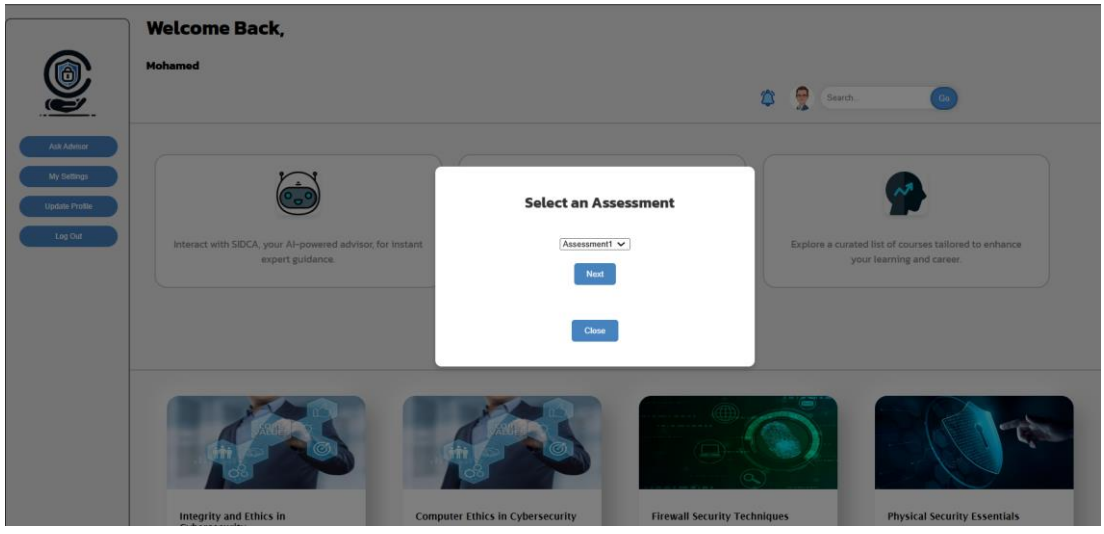

Figure 22. Select Assessment Screen

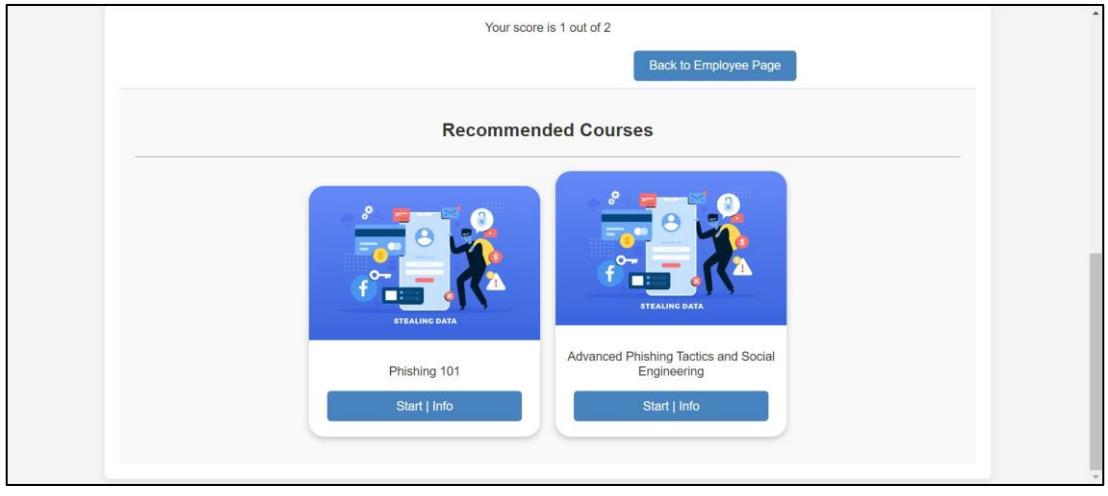

Figure 23. Recommended Courses Dashboard

Figure 24. Course Detail Screen

Finally, once logged-in, the team lead is directed to the main dashboard, which provides access to key features: Updating Profile, Assigning Skill Assessments, and Viewing Employee Scores. These features enable the team lead to manage their profile, monitor team performance, and facilitate targeted skill development. The team lead can assign skill assessments to employees directly from the dashboard. As demonstrated in Figure 25, once a skill assessment is assigned, it becomes accessible to the selected employee. This feature allows team leads to target specific skills or knowledge areas, helping employees enhance their cybersecurity competencies. The assessments assigned here align with the functionalities described previously in the employee workflow.

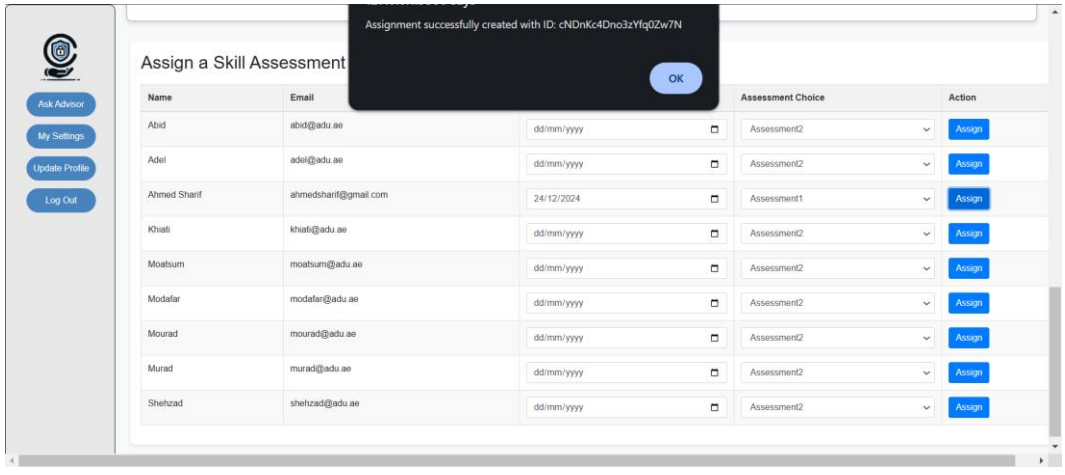

Figure 25. Assigning Skill Assessments to Employees

Additionally, the team leads can view employee performance through a results dashboard. Figure 26 illustrates how scores are retrieved from the database and reflected on the dashboard. Valid scores are displayed as integers, providing team leads with a clear overview of their team's performance. This feature supports effective monitoring and enables data-driven decisions for further skill development.

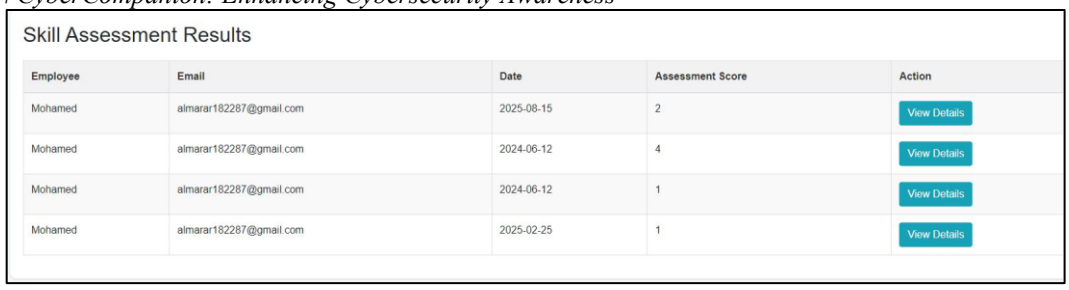

Figure 26. Viewing Employee Performance Scores

Results and Discussions

The proposed system is designed to offer an intuitive experience by interacting with an AI smart advisor named (SIDCA) and a course recommendation system powered by AI and ML techniques, specifically tilizing NLP and generative AI. The proposed system provides more interactive, personalized, and intuitive course recommendations to enhance the learning experience. The various development tools and libraries integrate coherently together enabling the system to allow users to ask SIDCA any user query they want to ask while getting tailored course recommendations. The ML algorithm used for the course recommendation system can handle similarity comparisons, it can also capture the semantic relationship between words which is the widely and commonly used method for recommendation systems.

The evaluation metrics for this recommendation system would involve a qualitative assessment involving the manual revision of the course name, description, and the incorrect question that the employee has completed. It is useful as we are able to capture any context that quantitative assessments might miss and since the recommended courses are only two, it is relatively easy to validate. For example, as shown in Figure 27, if the employee takes an assessment and answers questions about firewalls incorrectly, the page will display different courses related to firewalls to help strengthen their knowledge in the areas where weaknesses were identified during the assessment. This ensures that the recommendation system works successfully with high accuracy as it has recommended the correct courses out of all the courses in the database. Even if the employee proceeds back to the employee dashboard page without starting any of the recommended courses, the employee can still find these two same recommended courses in the dashboard ensuring error handling in the user perspective and ensures this feature is fault tolerant.

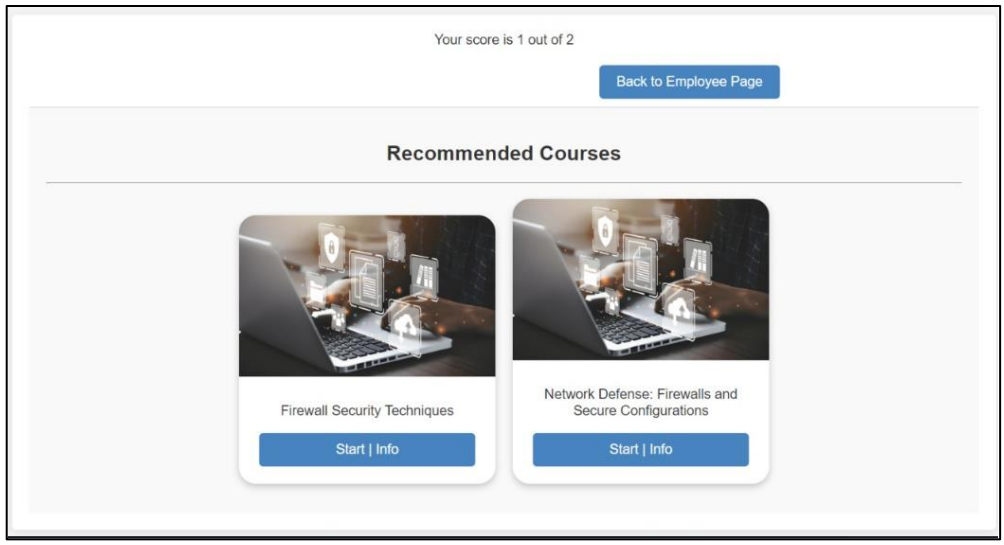

Figure 27**.** Displaying Recommended Courses Based on Assessment Results

Furthermore, if the employee interacts with the smart advisor, SIDCA, the user may attempt to input incorrect statements, statements with typos, out-of-scope statements, or non-English statements. All of these are handled thoroughly through the fine tuning and prompt engineering that was discussed earlier in section 4. This ensures that this feature is very convenient to help the user enhance their security posture through fun and interactive learning.

Table 2. summarizes the benchmarking of the proposed system with its features compared to others existing applications. We identified that while applications like "Hydran Cybersecurity Advisor" [27] and "SCIPS" [28] offer password analysis, they lack the interactive and AI-driven learning environment that our proposed system provides. Specifically, "CyberAware" [29] is aimed at K-6 children and employs game-based learning to educate on cybersecurity. However, despite its innovative approach, "CyberAware" may have limitations, such as its primary focus on a younger audience, potentially missing critical engagement with older or more advanced users. Additionally, its content update frequency and adaptability to diverse learning environments outside of its target demographic could be areas for improvement [29]. On the other hand, "OpenSecurityTraining" provides cybersecurity training through its website at no cost with no required sign-up [30]. Even though it is free, it is too simplistic and lacks dynamic content and interactivity. Our proposed system offers comprehensive features that serve a much wider user base, ensuring that employees across various sectors can benefit from a simplified cybersecurity training platform.

Application	User Performance Tracking	Password Analysis	Interactive Quizzes & Training	Generative AI	Cybersecurity Awareness & Training	Limitations
Cybersecurity game-based	Yes	No	Yes	No	No	**Might be expensive**

learning [17]						**to maintain and difficult to keep engaging**
Riskio [20]	Yes	Yes	Yes	No	Yes	**Functionality depends on content**
Hydran Cybersecurity Advisor [27]	No	Yes	No	No	Yes	**Limited feedback and scalability**
SCIPS [28]	Yes	No	Yes	No	Yes	**Limited to cyber security incidents and English only**
CyberAware [29]	Yes	No	Yes	Yes	Yes	**Focus on a younger audience only**
OpenSecurityTraining [30]	No	No	No	Yes	No	**Lacks interactive training content and has limited functionality**
QRadar Advisor [31]	No	No	Yes	No	Yes	**Complex to use and might be expensive**
Our proposed System CyberCompanion	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**---**

Table 2. Similar Applications Benchmark

Economic, Social and Environmental Impact

The proposed system "CyberCompanion" takes the economic, social, and environmental impacts into consideration during development aligning with the UAE's sustainability goals. It is designed to bridge the cybersecurity awareness gaps and contribute to the economy by minimizing financial losses from data breaches. It promotes inclusivity and accessibility in cybersecurity education, enhancing digital literacy across society while ensuring individuals remain secure in their daily lives. As a result, a more informed and vigilant society is formed, reducing the main vulnerability of human error and promoting a culture of proactive cybersecurity defense.

The system boosts cybersecurity skills and awareness meaning that from an economic perspective, it can reduce costs associated with data breaches and improve efficiency within organizations. The system also enables ease of access for cybersecurity education, making it accessible to a wider range of sectors and promoting proactive cybersecurity resilience. By providing individuals with the minimum required knowledge to protect themselves and their organizations online, CyberCompanion contributes to a safer cyber space, fostering a culture of trust and security.

CyberCompanion is designed to function as a digital platform effectively making environmental impact negligible. The commitment to online service delivery enables the reduction of carbon emissions unlike traditional and in-person training methods. In the future, CyberCompanion plans to explore more sustainable operation practices such as supporting users' access through energy-efficient devices which further align the project with environmental sustainability goals.

Conclusion and Future Directions

The proposed system "CyberCompanion" addresses a significant awareness gap in cybersecurity by providing an interactive, AI-powered learning platform designed to bridge the cybersecurity awareness gaps. Through comprehensive features such as SIDCA, personalized course recommendations, and real-time performance analytics, the system delivers a holistic and scalable solution for cybersecurity awareness challenges. Utilizing advanced technologies such as GenAI, NLP, and ML, CyberCompanion dynamically adapts the training material to its users, ensuring a personalized learning experience. The platform integrates robust security measures to protect from cyberattacks, including encryption, multi-factor authentication, and compliance with regional and international standards such as the OWASP Top 10 standard and NIST, ensuring reliability and data protection. CyberCompanion reduces the risk of data breaches caused by human error by empowering employees with the cybersecurity knowledge required to strengthen organizational resilience. Future developments could include the integration of augmented and virtual reality (AR & VR) for immersive training sessions and advanced AI-driven analytics for better personalization. These enhancements will establish CyberCompanion as a vital basis in encouraging a proactive secure community, ensuring its relevance in an ever-evolving threat landscape.

Ethics Declarations

Competing interests: The authors declare no competing interests.

References

P. UK. People-centric cybersecurity: A study of IT Security leaders in the UAE — proofpoint uk. [Online]. Available: https://www.proofpoint.com/uk/resources/ white-papers/people-centric-

cybersecurity-study-it-security-leaders-uae. Accessed: Jan. 20, 2024.
R. Vardhman and L. Tonogbanua. How many cyber attacks happen per day in 2024? [Online]. Available: https://techjury.net/blog/how-many-cyber-attacks-per-day/. Accessed: Feb. 12, 2024.
OWASP, "OWASP Top Ten." [Online]. Available: https://owasp.org/www-project-top-ten/. Accessed: Aug. 25, 2024.
B. News, "Colonial Pipeline: US recovers most of ransom, justice department says," [Online]. Available: https://www.bbc.com/news/business-57394041. Accessed: Mar. 18, 2024.
B. News, "Meat Giant JBS pays 11m in Ransom to Resolve Cyber-Attack," [Online]. Available: https://www.bbc.com/news/business-57423008. Accessed: Apr. 22, 2024.
Cybersecurity Ventures, "Annual Cybercrime Report 2017." [Online]. Available: https://cybersecurityventures.com/annual-cybercrime-report-2017/. Accessed: Jun. 2, 2024.
Y. Dai, A. Liu, and C. P. Lim, "Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education," Procedia CIRP, vol. 119, pp. 84–90, 2023. [Online]. Available: https://doi.org/10.1016/j.procir.2023.05.002. Accessed: May 3, 2024.
R. Gozalo-Brizuela and E. C. Garrido-Merch´an, "A survey of Generative AI Applications," arXiv (Cornell University), 1 2023. [Online]. Available: https://arxiv.org/abs/2306.02781. Accessed: Jun. 15, 2024.
M. Javaid, A. Haleem, R. P. Singh, S. Khan, and I. H. Khan, "Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system," BenchCouncil transactions on benchmarks, standards and evaluations, vol. 3, no. 2, p. 100115, 6 2023. [Online]. Available: https://doi.org/10.1016/j. tbench.2023.100115. Accessed: Jul. 8, 2024.
J. Kocoń et al., "ChatGPT: Jack of all trades, master of none," Information Fusion, vol. 99, p. 101861, 2023. [Online]. Available: https://doi.org/10.1016/j.inffus.2023.101861. Accessed: Aug. 19, 2024.
OpenAI. OpenAI Platform Documentation: Overview. https://platform.openai.com/docs/overview. Accessed: Jun. 24, 2024.
A. Radford, K. Narasimhan, and et al. Improving Language Understanding by Generative Pre-Training. [Online]. Available: https://www.mikecaptain.com/ resources/pdf/GPT-1.pdf. Accessed: Oct. 14, 2024.
Firebase. A complete foundation for your web app. [Online]. Available: https://firebase.google.com/products/hosting. Accessed: Nov. 2, 2024.
Thanam. (2024, 3) Engineering Education. [Online]. Available: https://www.section.io/engineering-education/ how-to-secure-firebase-apps-with-firebase-security-rules/. Accessed: Dec. 10, 2024.
GeeksForGeeks. "what is full stack development?". [Online]. Available: https: //www.geeksforgeeks.org/what-is-full-stack-development/. Accessed: Jan. 23, 2024.
B. Taylor. "the future of elearning: How technology is transforming education". [Online]. Available: https://elearningindustry.com/ the-future-of-elearning-how-technology-is-transforming-education. Accessed: Feb. 8, 2024.
Jin, G., Tu, M., Kim, T.-H., Heffron, J., and White, J., "Evaluation of Game-Based Learning in Cybersecurity Education for High School Students," Journal of Education and Learning, vol. 12, no. 1, pp. 150–158, Feb. 2018. [Online]. Available: https://doi.org/10.11591/edulearn.v12i1.7736. Accessed: Apr. 25, 2024.
Cone, B. D., Irvine, C. E., Thompson, M. F., and Nguyen, T. D., "A video game for cyber security training and awareness," Computers and Security, vol. 26, no. 1, pp. 63–72, Feb. 2007. [Online]. Available: https://doi.org/10.1016/j.cose.2006.10.005. Accessed: Apr. 15, 2024.
S. M. Ho and M. Gross, "Consciousness of cyber defense: A collective activity system for developing organizational cyber awareness," Computers and security, vol. 108, p. 102357, 9 2021. [Online].

Available: https: //doi.org/10.1016/j.cose.2021.102357. Accessed: Apr. 5, 2024.
S. Hart, A. Margheri, F. Paci, and V. Sassone, "RisKiO: a serious game for cyber security awareness and education," Computers and Security, vol. 95, p. 101827, Aug. 2020. [Online]. Available: https://doi.org/10.1016/j.cose.2020.101827. Accessed: Jul. 5, 2024.
B. Srinivasa-Desikan, Natural language processing and computational linguistics. Elsevier, 2018. Accessed: Feb. 27, 2024.
spaCy. ”spacy: Linguistic features”. [Online]. Available: https://spacy.io/usage/ linguisticfeatures#vectors-similarity. Accessed: Jan. 30, 2024
OpenAI. OpenAI Platform. [Online]. Available: https://platform.openai.com/ docs/introduction/key-concepts. Accessed: May 17, 2024.
K. KAPOOR. ”coursera courses dataset 2021”. [Online]. Available: https: //www.kaggle.com/datasets/khusheekapoor/coursera-courses-dataset-2021. Accessed: May. 20, 2024.
K. Mehreen. A guide to top natural language processing libraries - KDNuggets. [Online]. Available: https://www.kdnuggets.com/2023/04/ guide-top-natural-language-processinglibraries.html. Accessed: Jan. 15, 2024.
spaCy. ”spacy: Linguistic features”. [Online]. Available: https://spacy.io/usage/ linguisticfeatures#vectors-similarity. Accessed: Jan. 30, 2024
H. C. Security. ”cyber security advisor as a service — hydra cyber security”. [Online]. Available: https://hydracybersecurity.io/cyber-security-advisor-as-a-service. Accessed: Jul. 30, 2024.
S. O’Connor, S. Hasshu, J. Bielby, S. Colreavy-Donnelly, S. Kuhn, F. Caraffini, and R. Smith, "SCIPS: A serious game using a guidance mechanic to scaffold effective training for cybersecurity," Information Sciences, vol. 580, pp. 524–540, 2021. [Online]. Available: https://doi.org/10.1016/j.ins.2021.08.098. Accessed: Sep. 25, 2024.
F. Giannakas, G. Kambourakis, and S. Gritzalis, "CyberAware: A mobile game-based app for cybersecurity education and awareness," in IEEE International Conference on Interactive Mobile Communication, Technologies and Learning (IMCTL), 2015, pp. 20–25. [Online]. Available: https://doi.org/10.1109/imctl.2015.7359553. Accessed: Jun. 22, 2024.
"O. S. Training. ”open security training”. [Online]. Available: https://
opensecuritytraining.info/Training.html. Accessed: Aug. 12, 2024.
IBM. ”ibm qradar advisor with watson”. [Online]. Available: https://www.ibm.com/downloads/cas/52GBXLK8. Accessed: Oct. 3, 2024.