WIP: CyberSkiller: Empowering STEM Undergraduates for the Future of Cybersecurity

Gahangir Hossain
Information Science
University of North Texas
Denton, TX, USA
Gahangir.Hossain@unt.edu

Kinshuk
College of Information
University of North Texas
Denton, TX, USA
Kinshuk@unt.edu

***Abstract*— This innovative practice WIP paper relates to the current age of advanced technology, characterized by innovations like 6G networks and multi-cloud environments, safeguarding cyberinfrastructure is of paramount importance. Simultaneously, the escalating cyber threats have exposed a significant shortfall in cybersecurity expertise. This research endeavors to develop an online intelligent training system known as "CyberSkiller." The primary objective of CyberSkiller is to equip STEM undergraduates with the skills required to address the demands and opportunities within the cybersecurity domain. It seeks to empower individuals, leveraging their existing knowledge and backgrounds, through either augmenting their current skill set (up-skilling) or acquiring entirely new competencies (re-skilling). This work-in-progress paper presents the preliminary design of the CyberSkiller system.**

Keywords—Intelligent tutoring, Cybersecurity Workforce, Cybersecurity Training, STEM Undergraduates, Up-skilling, Re-skilling

I. Introduction

In the contemporary era dominated by cutting-edge technologies, such as 6G networks and multi-cloud environments, safeguarding cyberinfrastructure stands as an imperative task. The increasing complexity of cyber threats accentuates a critical deficiency in cybersecurity expertise. In response to this challenge, this research introduces an innovative solution in the form of an online intelligent training system "CyberSkiller." CyberSkiller emerges as a crucial initiative with the primary goal of arming STEM undergraduates with the essential skills requisite for navigating the intricate landscape of cybersecurity. It aims to address both the demands and opportunities inherent in this dynamic field, recognizing the indispensability of a well-equipped workforce. The system's core focus lies in empowering individuals by harnessing their existing knowledge and backgrounds. Whether through enhancing their current skill set (up-skilling) or acquiring entirely new competencies (re-skilling), CyberSkiller endeavors to mold adept professionals capable of tackling the evolving challenges of the cybersecurity domain. The aim is to implement continuous learning modules, hands-on training labs, and mentorship programs to ensure cybersecurity professionals can regularly update their skills, while fostering a collaborative learning environment through online platforms and industry partnerships. The system utilizes gamification elements and industry-recognized certification programs to create engaging and immersive up-skilling and re-skilling experiences, ensuring a versatile workforce capable of addressing evolving cybersecurity challenges. This paper provides an overview of the preliminary analysis and design of the CyberSkiller system. By leveraging intelligent training methodologies, the system aspires to bridge the existing gap in cybersecurity expertise. Through a comprehensive approach, CyberSkiller not only aims to fortify the cyber defenses of organizations but also serves as a catalyst for individual growth and adaptability in the face of an ever-evolving technological landscape. As the cyber threat landscape continues to evolve, CyberSkiller emerges as a promising solution to cultivate a skilled workforce capable of navigating and mitigating the complexities of the cybersecurity domain.

As outlined in [1], identifying instructional problems entails a combination of three fundamental components: need assessment, goal analysis, and performance assessment. Need assessment serves as a robust tool for categorizing needs into six types: normative, comparative, felt, expressed, anticipated, and critical incidents. Goal analysis involves delineating objectives, establishing primary and secondary goals, refining these objectives, and prioritizing them accordingly. Performance assessment aids in determining whether the training program effectively addresses the stated goals, or if an alternative intervention is more suitable. This comprehensive process of identifying instructional problems, depicted in Figure 1, often necessitates employing multiple techniques and iterations to precisely identify the issue. In this model, one may commence with either need assessment or performance assessment and subsequently incorporate goal analysis.

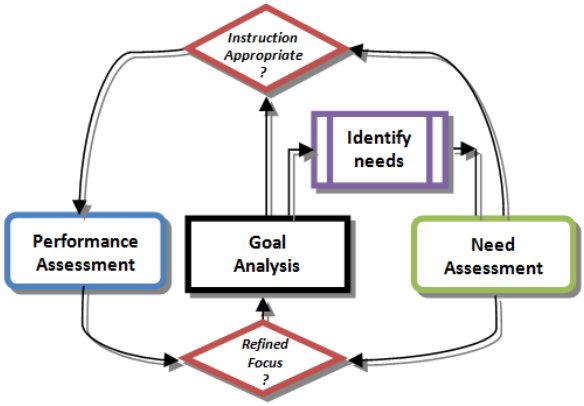

Fig 1. Instructional Problem Identification Model

The main goal of this research is to take advantage of intelligent tutoring systems in workforce development through online training and skill improvement. To realize this goal, the project focuses on four key tasks: (1) Researching intelligent tutoring systems, examining a wide range of tutoring tools, techniques, and resources suitable for cybersecurity training and skill enhancement. (2) Integrating virtual reality and other applied learning tools into practical tasks related to computing, science, and engineering curriculums, such as AI and machine learning. (3) Developing the CyberSkiller system to autonomously update itself with the latest information from various educational resources and web links. (4) Evaluating and enhancing the performance of the tutoring systems in fostering communities that effectively use online training materials and tools.

The rest of this paper is organized as follows. In Section II, we briefly review intelligent tutoring systems that can be used for cybersecurity skill development and re-skilling. Section III describes how to use reinforcement learning for CyberSkiller modeling and inference. In Section IV, we explain CyberSkiller's prospects and limitations for up-skilling and re-skilling, providing examples. Section V discusses the features that enable CyberSkiller to be integrated into career pathways. The conclusion is presented in Section VI.

II. Litarure Review

Cybersecurity training is vital for preparing the next generation of cybersecurity professionals and keeping current workers' skills up to date. This training can be done in a number of ways, such as through face-to-face or online college education, attending workshops, or internships in industries. To have self-phased training, online agent based tutoring systems, called intelligent tutoring systems (ITSs) could be effective. ITSs like ALEKS, Cognitive Tutor, and IBM's Watson, which are increasingly popular in both classroom and online settings, exemplify this trend [1]. While hands-on experience and team skill development are preferred for cyber training, virtual simulation-based solutions also have their place [2]. Meanwhile, for professionals who need to update their skills, returning to university can be challenging. The CyberSkiller (an ITS) can serve as a versatile tool, assisting students with varying skill levels. Whether the students have foundational knowledge or need more comprehensive training, the CyberSkiller aims to provide the necessary support and skill updates [3].

Investigation on various intelligent tutoring systems, such as AutoTutor [4, 5], MetaTutor [6, 7], Deep Tutor [8], and a virtual reality-based tutor [9] show that such tutors are capable of enhancing the training effectiveness . In addition, GAN based tutoring [10][11] and GPT based robust technologies [12][13][14], extended to chatbots [15], have shown huge prospects for future learning technology research.

The aim of this research is to design and develop CyberSkiller intelligent tutoring system tailored for cybersecurity training. This research establishes connections between different active learning methods, educational materials, web-based projects in information systems, and cybersecurity courses. The system facilitates flexible, personalized learning, enabling students to assess their skills, find relevant courses or modules, and reinforce their capabilities to tackle cybersecurity challenges. Specifically, this research strives to explore, develop, implement, and support a dynamic, personalized web-based intelligent learning environment for tutoring and training the next generation of cybersecurity experts.

III. CyberSkiller Models and Interfaces

The CyberSkiller system enables students to learn contemporary cybersecurity topics, assess their skills instantly, and engage in interactive learning and quizzes, aligned with W3Schools Online Web Tutorials [16]. Participation in online course materials not only keeps students updated but also rewards them with recognition badges. Organized, interactive, and readily accessible content, along with personalized options, transforms the learning experience. By integrating search results and competency scores (e.g., credit scores), students can

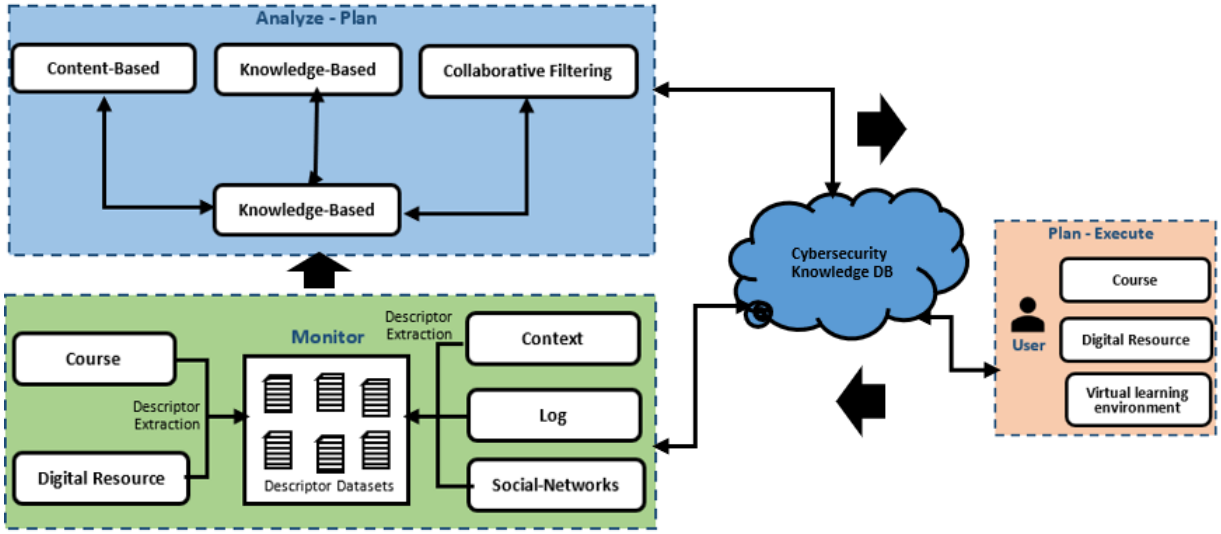

Fig 2. The proposed cloud based CyberSkiller system.

swiftly identify and address knowledge gaps. The proposed system is constructed amalgamating elements from recent works [17 -19] to facilitate a hybrid structure, which is shown in Figure 2.

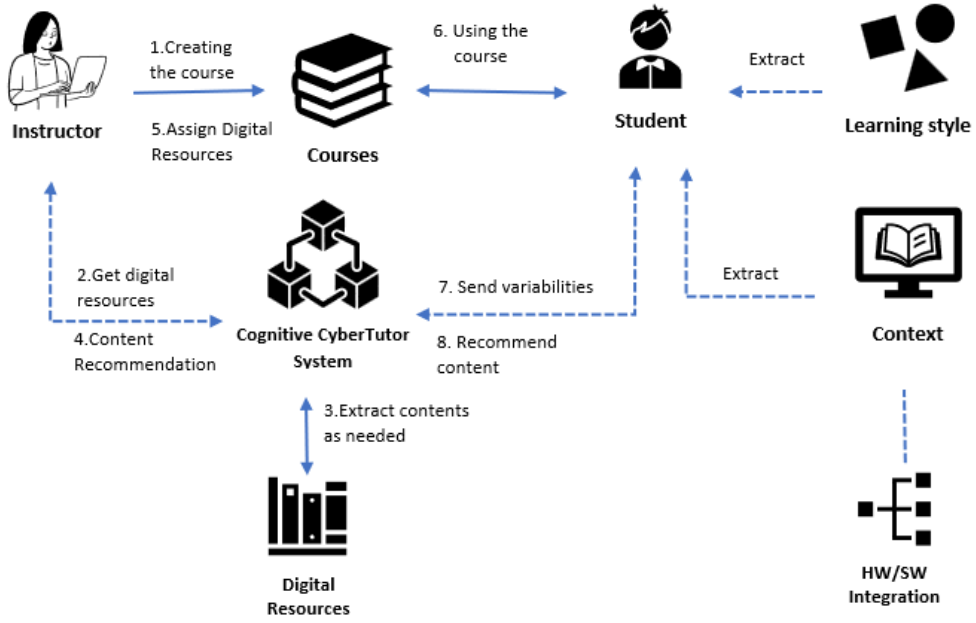

Fig 3. Steps in learning with Cognitive Systems based CyberSkiller

The CyberSkiller agent framework is abstract, flexible, and applicable to a variety of web-based problems in multiple ways. For example, consider using reinforcement learning to align recent cybersecurity topics in demand in the job market with the topics a learning agent has already mastered and those the learner can quickly grasp. In this context, actions might involve assessing the learner's current knowledge and providing learning instructions that enhance their skills to meet career demands. Rewards could be given for each successfully acquired new topic. To maintain a smooth learning environment, a small, negative reward can be applied at each time step to reflect ongoing progress and effort (Figure 3).

The CyberSkiller investigates students' cognitive load to ensure they are not overwhelmed by course materials when using the online training for immediate access to information and hands-on experience. Furthermore, intelligent cybersecurity games are integrated, allowing students to practice proper actions in real-time. Learning can happen at students' convenience. Various systems, such as cyber ranges and testbeds [17] and cyber security training for critical infrastructure protection [20], offer insights. Monsalve-Pulido, Julián, et al. [18], propose an intelligent learning system with an efficient recommendation system, while Weyns, Danny, et al. [19], describe an expert system for efficient recommendation and orchestration of information in the learning environment.

A. CyberSkiller Learning Environment

As part of the machine learning component in the CyberSkiller design, the reinforcement learning (RL) technique is adopted. RL is motivated by increased interaction between the learner (student) and the learning environment. This technique frames the problem and allows learning from interactions to achieve the goal.

In the context of artificial intelligence and RL terminology, the learner is considered as a "learning agent (LA)" and everything outside the agent as the "learning environment (LE)." The interaction process continues through learning activities, where the agent selects actions and the learning environment responds, presenting new situations to the agent. As part of RL, the learning environment generates feedback known as "rewards (R)" to motivate the learner. Rewards can be coded as numeric values corresponding to the learning gain. A complete formal explanation of the learning task and learning environment can be defined as follows –

In a discreate time steps ($t = 0, 1, 2, 3, \ldots n$), the learning agent (LA) and the CyberSkiller learning environment (LE) interacts. The LA receives some representation of the LE, which is defined as states S, where $S_t \in S$. Based on this, a learning action is considered as $a_t \in A(S_t)$, where $A(S_t)$ is the set of learning instructions available for the learner in state S_t. As the process progresses to advance towards a new state S_{t+1}, the LA receives the feedback in terms of reward, r_{t+1}, were, $r_{t+1} \in R$, as shown in figure 4.

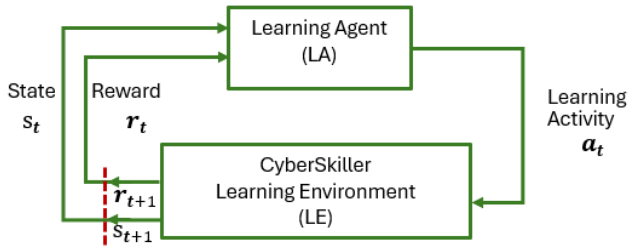

Fig 4. Learning activity and environment with steps

For the automated recommender system "CyberSkiller," skill improvement through item difficulty sequence modeling is achieved by adopting a practical model [16], which is a crucial initial step. In line, multi-faceted item features are utilized to develop a robust skill model which is capable of handling sparse data. Reinforcement learning (RL) can be employed to estimate item difficulty on a scalable basis.

IV. CyberSkiller : Up-skilling and Re-Skilling

CyberSkiller helps in "Up-skilling" and "Re-skilling" in the context of cybersecurity workforce development.

Up-skilling: CyberSkiller aims to enhance or improve the skills and knowledge of individuals in a specific cybersecurity area, often to keep up with the evolving demands of their current job or industry. In the context of STEM undergraduates and cybersecurity, up-skilling would help them further develop the existing knowledge and skills the students have in STEM fields, such as computer science or engineering, to become proficient in cybersecurity.

Re-skilling: The CyberSkiller helps in re-skilling that involves acquiring entirely new skills or knowledge in a different like the cybersecurity domain. In the context of STEM undergraduates and cybersecurity, re-skilling would help students transition from their existing STEM background to start a career in the new but demanding field of cybersecurity. CyberSkiller organizes online learning activities and resources that involve learning new tools, methodologies, and concepts specific to the field of cybersecurity.

In both cases, the goal of CyberSkiller is to prepare STEM undergraduates to acquire necessary knowledge and skills to meet the demands and opportunities in the field of cybersecurity. This can be done either by building upon their current skills (up-skilling) or acquiring new ones (re-skilling).

For instance, using CyberSkiller, one can learn cybersecurity topics, such as cryptographic systems. To understand cryptographic systems, the learner first needs to grasp the basics, including definitions of sender, receiver, encryption, decryption, plaintext, ciphertext, and keys, as shown in figure 5. It may also include in-depth encryption and decryption algorithms based on the learning agent's current knowledge level. After mastering these fundamental concepts, the learner can proceed to study the types of cryptography, such as symmetric and asymmetric encryption.

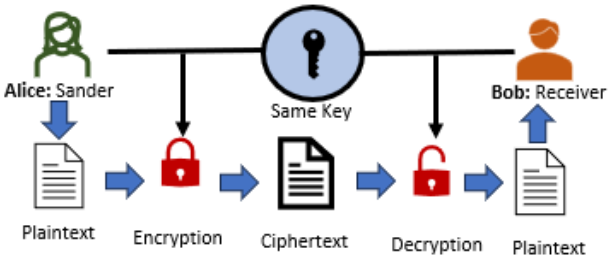

Fig 5. Cryptographic system

For re-skilling, the learning environment (LE) includes relevant mathematics and basic computing skills, such as algorithms and propositional logic. To evaluate the usability and cognitive load of CyberSkiller, the following set of questions has been designed for students with responses requested on a Likert scale ranging from strongly agree to strongly disagree:

1. How do you feel overall about the CyberSkiller instruction?
2. How enjoyable has CyberSkiller-based learning been for you?
3.How difficult was the learning content for you?
4. How difficult was it to learn the material?
5. How much did you concentrate while learning?

The first two questions assess the usability of the system, while the last three questions collectively evaluate the cognitive load of the learner. The total cognitive load is measured as the combination of intrinsic cognitive load (3rd question), extraneous cognitive load (4th question), and germane cognitive load (5th question).

Based on the usability and cognitive load assessments, the CyberSkiller system is fine-tuned for optimal performance. In addition, a pre and a post skill test help in recommending whether the student required up-skilling or re-skilling modules in CyberSkiller.

V. CyberSkiller : Prospects to Cybersecurity Career

The field of cybersecurity has emerged as a critical necessity across various sectors, including government, industries, healthcare, business, and academia. However, there is an increasing shortage of skilled professionals in this domain. Given the growing demand for cybersecurity expertise in both public and private sectors, there is an urgent need for skilled personnel. Therefore, learning and improving cybersecurity skills is imperative to meet the future demands of the cybersecurity workforce.

CyberSkiller system includes various learning activities within basic, intermediate, and advanced modules related to foundational cybersecurity concepts to prepare for careers in cybersecurity. It includes introducing fundamental terms and definitions related to ethics, security, and privacy concepts, along with best practices for protecting cyber assets. Additionally, various hands-on software tools are included as part of the curriculum. Hands-on projects and real-world application case studies are essential components, enabling students to apply theoretical knowledge to practical scenarios.

The research has also investigated potential challenges associated with integrating the rapidly changing cybersecurity modules. By preparing students with the skills and mindset necessary for success in cybersecurity careers, this research underscores the urgency and importance of embracing future cybersecurity careers. Question arises, which is the most appropriate professional development program for employees' reskilling and upskilling, considering the individual and organizational influence factors? For instance, conducting pre and post evaluation can shed some light, as shown in Table 1.

TABLE I: PERFORMANCE BASED UP-SKILLING AND RE-SKILLING SELECTION

	Skill Set Difficulty (Pre or post design)	Evaluation Score (Pre)	Evaluation Score (Post)
1.	Well below average	1	1
2.	Below average	5	5
3.	Slightly below average	5	10
4.	Slightly above average	20	16
5.	Above average	20	24
6.	Well above average	6	10

Table I illustrates an example of skill evaluation used to classify training needs into up-skilling or re-skilling based on pre-test and post-test scores. If a pre-test score falls below a predefined required threshold, the system classifies the skill set as needing re-skilling; otherwise, it qualifies for up-skilling. In the example shown in Table 1, individuals 1 and 2 require re-skilling. Additionally, goodness-of-fit tests can be applied to fine-tune the assessment process.

VI. CONCLUSION

Designing a learning environment with the reinforcement learning framework is inherently goal-directed, comprising three essential components: the actions taken by the learning agent, the states influencing these actions, and the rewards guiding the agent's goal. While this three-stage framework may not address all decision-learning problems comprehensively, the RL-based approach has demonstrated widespread usefulness and applicability. Hence proposed in CyberSkiller recommender system with recommending upskilling and re-skilling decisions. This is particularly vital in the continuously evolving realm of computing technology, ensuring that professionals possess the latest cybersecurity knowledge and expertise to tackle emerging threats and challenges. Cybersecurity training demands practical exposure to computing, information systems, and network literature.

The CyberSkiller project aims to deliver adaptable web-based learning materials, applications, and tools applicable to various practical courses. This initiative empowers students to bridge the gap between unfamiliar and familiar concepts within advanced cybersecurity topics. The system, designed to be self-adjusting with deep learning capabilities, categorizes cybersecurity information and misinformation, thereby enriching the knowledge base.

ACKNOLEDGEMENT

We utilized ChatGPT for grammar checking and revising the text of this paper.

REFERENCES

[1] S. Spark, "How 'Intelligent' Tutors Could Transform Teaching," EdWeek, 2017,

[2] J. Hu, M. Schmitt, C. Willems, and C. Meinel, "A tutoring system for IT security," in IFIP World Conf on Information Sec Education, 2003, pp. 51-60.

[3] "Tutoring vs. training: What's the Difference?" [Online].

[4] A. C. Graesser, K. Wiemer-Hastings, P. Wiemer-Hastings, R. Kreuz, and the Tutoring Research Group, "AutoTutor: A Simulation of a Human Tutor," Cognitive Systems Research, vol. 1, no. 1, pp. 35-51, 1999.

[5] A. C. Graesser, S. D'Mello, X. Hu, Z. Cai, A. Olney, and B. Morgan, "AutoTutor," in Applied Natural Language Processing: Identification, Investigation, and Resolution, IGI Global, 2012, pp. 169-187.

[6] E. Penalosa and S. Castaneda, "Meta-Tutor: an Online Environment for Knowledge Construction," International Journal of Continuing Engineering Education and Lifelong Learning, vol. 18, no. 3, pp. 283-297, 2008.

[7] R. Azevedo, R. S. Landis, R. Feyzi-Behnagh, M. Duffy, G. Trevors, J. M. Harley, and G. Hossain, "The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor," in International Conference on Intelligent Tutoring Systems, 2012, pp. 212-221.

[8] V. Rus, N. Niraula, and R. Banjade, "DeepTutor: An effective, online intelligent tutoring system that promotes deep learning," in Twenty-ninth AAAI conference on artificial intelligence, 2015.

[9] J. Hu, "A virtual machine architecture for IT-security laboratories," Doctoral dissertation, Universität Potsdam, 2006.

[10] Chui, K. T., Liu, R. W., Zhao, M., & De Pablos, P. O. (2020). Predicting students' performance with school and family tutoring using generative adversarial network-based deep support vector machine. IEEE Access, 8, 86745-86752.

[11] Gan, Wenbin. "Dynamic Learner's Knowledge Assessment by Incorporating Learner and Domain Modeling in Intelligent Tutoring Systems." PhD diss., The Graduate University for Advanced Studies, 2022.

[12] Vujinović, Aleksandar, Nikola Luburić, Jelena Slivka, and Aleksandar Kovačević. "Using ChatGPT to Annotate a Dataset: A Case Study in Intelligent Tutoring Systems." Machine Learning with Applications (2024): 100557.

[13] Limo, Fernando Antonio Flores, David Raul Hurtado Tiza, Maribel Mamani Roque, Edward Espinoza Herrera, José Patricio Muñoz Murillo, Jorge Jinchuña Huallpa, Victor Andre Ariza Flores et al. "Personalized tutoring: ChatGPT as a virtual tutor for personalized learning experiences."Przestrzeń Społeczna (Social Space) 23, no. 1 (2023): 293-312.

[14] Virvou, Maria, and George A. Tsihrintzis. "Is ChatGPT Beneficial to Education? A Holistic Evaluation Framework Based on Intelligent Tutoring Systems." In 2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA), pp. 1-8. IEEE, 2023.

[15] W3Schools [https://www.w3schools.com/

[16] N. Chouliaras, G. Kittes, I. Kantzavelou, L. Maglaras, G. Pantziou, and M. A. Ferrag, "Cyber ranges and testbeds for education, training, and research," Applied Sciences, vol. 11, no. 4, p. 1809, 2021.

[17] N. Chowdhury and V. Gkioulos, "Cyber security training for critical infrastructure protection: A literature review," Computer Science Review, vol. 40, p. 100361, 2021.

[18] J. Monsalve-Pulido, et al., "Autonomous recommender system architecture for virtual learning environments," Applied Computing and Informatics, 2020.

[19] D. Weyns, et al., "Towards Better Adaptive Systems by Combining MAPE, Control Theory, and Machine Learning," in 2021 International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS), IEEE, 2021.

[20] Umemoto, K., Milo, T., & Kitsuregawa, M. (2020). Toward recommendation for upskilling: Modeling skill improvement and item difficulty in action sequences. In 2020 IEEE 36th International Conference on Data Engineering (ICDE) (pp. 169-180). IEEE.