

# Intelligent modeling of personalized learning in cybersecurity training<sup>\*</sup>

Pavlo Skladannyi<sup>1,2,\*</sup>, Yuliia Kostiuk<sup>1,†</sup>, Oleksii Zhylytsov<sup>1,‡</sup>, Yuriy Savchenko<sup>1,‡</sup>, and Yevhen Antypin<sup>1,‡</sup>

<sup>1</sup> Borys Grinchenko Kyiv Metropolitan University, 18/2 Bulvarno-Kudriavska str., 04053 Kyiv, Ukraine

<sup>2</sup> Institute of Mathematical Machines and Systems Problems of the National Academy of Sciences of Ukraine, 42 Ac. Glushkov ave., 03680 Kyiv, Ukraine

## Abstract

The article focuses on the application of intelligent technologies and Artificial Intelligence (AI) for modeling personalized learning trajectories in the training of cybersecurity and information security specialists. An approach is proposed that combines adaptive educational systems, educational data analytics, and AI algorithms to build dynamic learning models. The developed conceptual model offers content personalization and individualization of the learning process, tailored to the applicant's profile, current level of competence, and prediction of educational outcomes. The use of recommendation systems, Bayesian networks, reinforcement learning algorithms, and XAI technologies allows for the automatic formation of optimal educational trajectories, increasing the effectiveness of training and compliance with modern cybersecurity market requirements. Particular attention is paid to the integration of ISO/IEC 27001, NIST Cybersecurity Framework standards, and recommendations for ensuring data protection and privacy in electronic educational environments. Mathematical modeling is based on a partially observable decision-making process (POMDP), which allows building dynamic learning trajectories in conditions of incomplete information about the applicant's knowledge. The developed prototype of an intelligent educational system implements content personalization, adaptive task complexity control, result prediction, and cyber threat scenario simulation.

## Keywords

artificial intelligence, intelligent technologies, personalized learning trajectories, adaptive learning, digital pedagogy, cybersecurity, information security, learning analytics, competency ontology, secure educational platforms

## 1. Introduction

The rapid development of digital technologies, the integration of AI into various fields of activity, and the constant growth in the number of cyber threats necessitate a transformation of educational approaches to training specialists in cybersecurity and information protection [1–4]. Traditional learning models based on static programs are unable to provide the necessary level of adaptability and personalization, which complicates the development of the competencies required to address modern information risks and respond to dynamic threats [5–8]. This problem is particularly relevant in the context of the digital economy and the transition to intelligent educational environments, where the training of specialists requires the integration of innovative technologies, data analytics, and AI systems. The problem lies in the lack of effective educational platforms capable of modeling personalized learning trajectories for students with different levels of training and individual professional goals [2, 7]. Most existing solutions focus on standardized curricula that do not account for the dynamic characteristics of students, their individual learning pace, cognitive abilities, motivational factors, and changes in the cyber threat environment [4, 6, 9, 10]. In addition, educational systems often do not integrate international information security standards—ISO/IEC

<sup>\*</sup> CPITS-II 2025: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, October 26, 2025, Kyiv, Ukraine

<sup>\*</sup> Corresponding author.

<sup>†</sup> These authors contributed equally.

✉ p.skladannyi@kubg.edu.ua (P. Skladannyi); y.kostiuk@kubg.edu.ua (Y. Kostiuk); o.zhylytsov@kubg.edu.ua (O. Zhylytsov); y.savchenko@kubg.edu.ua (Y. Savchenko); y.antypin@kubg.edu.ua (Y. Antypin)

ORCID 0000-0002-7775-6039 (P. Skladannyi); 0000-0001-5423-0985 (Y. Kostiuk); 0000-0002-7253-5990 (O. Zhylytsov); 0000-0003-3662-2787 (Y. Savchenko); 0000-0002-0371-2498 (Y. Antypin)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

27001, NIST Cybersecurity Framework, GDPR—which leads to insufficient preparation of future specialists for practical challenges in the field of data protection.

The article proposes a conceptual model of an intelligent educational system that implements the modeling of personalized learning trajectories using AI technologies and smart analytics of educational data [3, 5, 8]. The approach is unique in that it combines a dynamic profile of the learner's competencies, ontology of knowledge in the field of cybersecurity, and AI algorithms for content adaptation using recommendation systems, reinforcement learning, Bayesian networks, and XAI technologies for explainability of learning decisions [11–13]. This approach allows for the automatic construction of optimal educational trajectories, adapting them to the individual needs of learners and current cyber threat scenarios.

The goal of the study is to develop and implement an intelligent educational platform that provides automated modeling of personalized learning trajectories in the training of cybersecurity and information security specialists. The proposed approach aims to integrate adaptive educational techniques, learning data analytics, and international cyber defense standards to enhance the effectiveness of professional competency development [14–16].

The practical significance of the work lies in the creation of an intellectual system of personalized learning capable of improving the quality of specialist training by integrating dynamic competency models, practice-oriented tasks, cyberattack simulation scenarios, and adaptive recommendations based on AI [17–19]. The implementation of the proposed model in educational platforms will ensure personalized training of future specialists, develop flexible skills for responding to cyber threats, increase the level of information security, and ensure that training programs meet international requirements and the challenges of the modern digital environment.

The scientific contributions of the study consist of the development of an integrated POMDP model for personalizing learning trajectories with optimization of the  $\pi_\theta$  policy based on reinforcement learning, which takes into account knowledge dynamics, cyber threat risks, and data privacy [11]. An ontological graph of competencies with Bayesian a posteriori distributions and explainable AI (XAI) based on SHAP/LIME for transparent recommendations has been proposed [5, 12]. For the first time, differential privacy  $\epsilon_{\text{tot}}$  budget control has been integrated, and an SPI metric covering MITRE ATT&CK tactics has been introduced to assess students' readiness for cyber incidents. An experiment involving 120 participants demonstrated the effectiveness of the system ( $F1=0.89$ ,  $RMSE=0.12$ ,  $LG=0.43$ ,  $SPI=0.87$ ), which is 40% better than the results obtained with traditional LMS.

## 2. Literature review

Contemporary research pays considerable attention to the use of intelligent technologies and AI in the field of training specialists in cybersecurity and information security [4, 13, 18]. Approaches focused on personalized learning and adaptive educational trajectories demonstrate high effectiveness in improving knowledge acquisition and developing professional competencies.

For example, Seda et al. [1] presented an approach to adaptive learning in practical cybersecurity training that allows the complexity of tasks to be automatically adjusted according to the level of student preparation, thereby increasing the effectiveness of practical training. Further development of this direction is presented in the work of Vykopal et al. [2], where a “smart environment” for adaptive learning of cyber skills was created, combining Learning Analytics with AI algorithms for personalizing learning tasks.

Recent studies demonstrate the potential of Generative AI in education. For example, Wang et al. [3] proposed the CyberMentor platform, which utilizes large language models (LLMs) to provide personalized mentoring to students in the field of cybersecurity and to select educational resources adaptively. A similar approach was used by Elkhodr & Gide [5], who explored the integration of Generative AI into pedagogical strategies. This made it possible to improve students' critical thinking skills and develop practical skills in cybersecurity policy development.

The Triplett study [4] analyzes innovative solutions for AI-oriented cyber education, including the use of intelligent tutors, virtual laboratories, and competency assessment systems, which provide a personalized approach to shaping learning trajectories. Jawhar et al. [6] focus on the use of AI platforms for individualized learning and raising awareness in the field of cybersecurity, demonstrating the effectiveness of adaptive personalization models.

The problem of risk modeling in distributed information systems is also relevant. Palko et al. [9] proposed a cyber risk assessment model that considers the dynamics of threats and facilitates the integration of student training with realistic attack scenarios. In addition, a systematic review by Barrera Castro et al. [7] analyzed the barriers to personalized learning using AI and identified the main challenges related to data protection, content adaptation, and the construction of personalized learning models.

Despite significant progress in applying AI technologies for personalizing learning, modern approaches leave several unresolved issues. First, there is a lack of comprehensive systems where personalization is formulated as a POMDP with a policy  $\pi_\theta$ , trained with security risks (Risk, Viol) in the reward function [2, 20, 21]. Second, there are no integrated practical readiness metrics (SPI) that directly correlate with the coverage of MITRE ATT&CK tactics and techniques required for training cybersecurity specialists [4]. Third, there are no solutions that simultaneously use XAI constraints (threshold  $\theta_{\text{XAI}}$ ) and differentially private accounting (DP budget  $\epsilon_{\text{tot}}$ ) to ensure transparency of decisions and protection of personal data [7, 18, 22]. Additionally, there are almost no experiments on realistic data in virtual laboratories and Capture the Flag (CTF) scenarios, with percentage gains assessed on samples of more than 100 students over several weeks of training.

Despite active developments in the field of adaptive training, AI mentoring, Generative AI, and cyber risk modeling, most studies focus on individual aspects of the problem. Currently, there are no comprehensive models that integrate dynamic personalization, AI adaptation algorithms, cyber threat modeling scenarios, and personal data protection into a single platform. Such integration is necessary to create effective intelligent educational systems capable of providing high-quality training for cybersecurity and information security specialists in line with modern requirements and challenges. The approach combines POMDP modeling, RL policy, and an ontological graph of competencies to optimize dynamic learning trajectories, considering MITRE ATT&CK risks and scenarios. Unlike existing knowledge tracing methods (DKT, DKVMN, AKT) and RL-ITS, XAI-constraint and DP-budget are integrated to ensure transparency and privacy of decisions [18, 23]. The proposed system implements dynamic monitoring of  $\epsilon_{\text{tot}}$  for the first time and complies with ISO/IEC 27001 and GDPR standards, eliminating the limitations of current approaches.

### 3. Research methods

The research is based on a combination of theoretical, methodological, algorithmic, and experimental methods aimed at developing an intellectual model for modeling personalized learning trajectories in the training of cybersecurity and information security specialists. At the theoretical stage, an analysis of current research in the fields of AI, adaptive learning, and intelligent educational systems [1, 2, 4, 20, 21, 23] was conducted, as well as the requirements of international standards ISO/IEC 27001, ISO/IEC 27032, and NIST Cybersecurity Framework [13, 18]. This made it possible to define the pedagogical principles of personalization, the criteria for developing competencies, and the requirements for data protection in educational environments.

The methodology is based on systemic and competency-based approaches to creating knowledge ontology and defining professional competencies [7, 23], as well as on an adaptive approach that takes into account the individual characteristics of students [3–5, 8, 13]. To form personalized trajectories, learning data analytics and modern AI algorithms are utilized, specifically recommendation systems, Bayesian networks, reinforcement learning, and XAI for decision explainability.

The experimental part involved creating a prototype of an intelligent educational platform and testing it on students majoring in cybersecurity and information security [1, 2, 6]. The effectiveness of the system was assessed based on indicators of competency achievement, personalization accuracy, and recommendation quality, which confirmed the advantages of the developed model compared to traditional educational approaches [4, 7, 13, 18, 21, 23]. Thus, the methods used made it possible to create an intelligent educational system that combines AI technologies, adaptive learning, and international cybersecurity standards, ensuring the personalization of the educational process and increasing the effectiveness of training specialists in the field of cybersecurity and information security.

The experiment lasted 12 weeks and involved 120 students majoring in “Cybersecurity and Information Protection,” who were randomly divided into experimental and control groups of 60 participants each. Data was collected from the LMS platform, virtual laboratories, CTF environments, and activity logs [2, 22], which ensured comprehensive tracking of educational outcomes and behavioral signals. Participants met uniform selection criteria and had basic knowledge of network technologies.

The prototype was implemented in Python 3.11 using PyTorch, BNTToolkit, Scikit-learn, SHAP, and LIME [8, 12, 20, 23, 24]. To ensure reproducibility, random states (global seed), library versions, and data structure were fixed; training was performed using a 5-fold cross-validation scheme with identical splits for all models [18, 23, 25, 26]. Separately, pseudocodes for updating the belief state, content selection, and DP budget accounting are provided, allowing the results to be reproduced without disclosing personal data (GDPR, ISO/IEC 27001).

## 4. Main material

In the current context of digital transformation in education, the role of intelligent technologies and AI in building personalized adaptive trajectories is growing [18, 27], and the development of digital platforms, generative models, and adaptive systems is changing the training of cybersecurity specialists, shifting learning to intellectualized environments where decisions are based on data and analytics. Studies [1–7] show that adaptive systems that automatically generate individual routes ensure efficiency; for this purpose, recommendation algorithms, reinforcement learning, Bayesian networks, and Learning Analytics are used [1–3, 8, 11]. It is important to integrate XAI for transparency and trust in the results. The protection of personal data and compliance with ISO/IEC 27001, ISO/IEC 27032, the NIST Cybersecurity Framework, and GDPR [18, 23, 25, 28–31] are key, necessitating the creation of secure educational platforms with encryption, event auditing, and dynamic risk management.

The development of personalized adaptive learning is supported by new approaches to pedagogical design. The focus is on a student-centered model that involves the active participation of learners in building their own trajectory, managing the learning process, and choosing optimal learning scenarios [13, 27, 32]. Intelligent educational platforms can take into account motivational factors, cognitive styles, and professional competencies, automatically forming competency profile models and adjusting content in real-time.

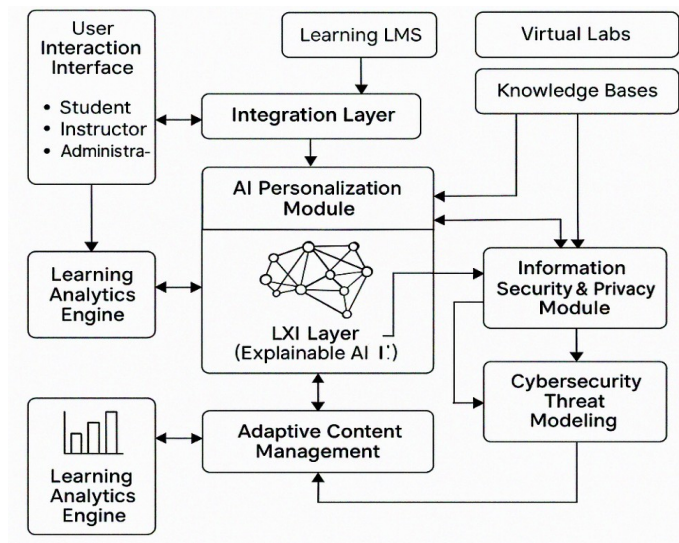
Personalized adaptive learning is understood as an intellectually controlled process in a digital environment that dynamically changes content, methods, and tools according to the individual characteristics, needs, and level of competence of the learner [23, 33]. For deep personalization, intelligent agents are used to model micro-portions of content (terms) and form multidimensional educational “cubes”. A promising direction is new-generation systems that combine AI algorithms, adaptive pedagogical strategies, cyber risk models, and information security mechanisms in a single platform [18, 25, 26, 29, 34], ensuring optimal trajectories and improving the quality of specialist training.

The concept is based on the integration of AI, Learning Analytics, and Explainable AI (XAI); the central idea is personalized trajectories that take into account the level of knowledge, cognitive characteristics, learning style, and professional needs [1, 2, 11, 12, 27]. The system automatically



selects materials based on individual data, generates tasks of varying complexity, and adjusts content in real time.

Figure 1 shows the architecture of an intelligent educational system designed to create personalized learning paths for training cybersecurity and information security specialists. The central element is the AI Personalization Module, which adapts learning content to the individual needs of students. The Learning Analytics Engine analyzes educational data, and the XAI Layer ensures transparency of decisions. Adaptive Content Management generates micro-portioned content, and the Information Security & Privacy Module is responsible for protecting personal data in accordance with ISO/IEC 27001 and GDPR standards. The Cybersecurity Threat Modeling module allows you to practice cyberattack scenarios, while the Integration Layer provides interaction with LMS, virtual labs, and knowledge bases. The system supports adaptive, secure, and personalized learning through the use of AI, Learning Analytics, and information security mechanisms.

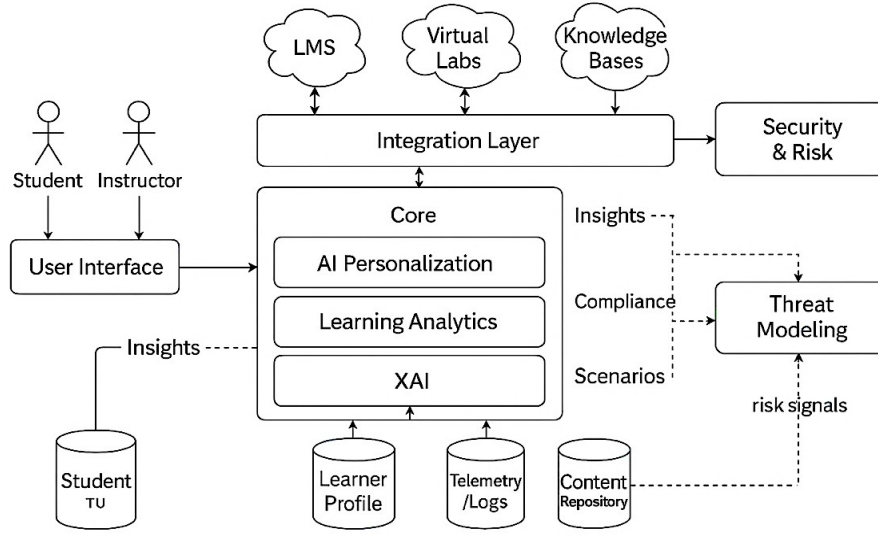


**Figure 1:** Architecture of an intelligent educational system for modeling personalized learning trajectories

Intelligent analytics monitors results and identifies gaps, allowing for dynamic adjustments to the learning trajectory [22]. XAI ensures transparency and trust in recommendations [4, 5, 13, 17, 18, 28], while the security module provides personal data protection, access control, and compliance with international standards. Scenario modeling of cyber threats (attack simulations, training environments, virtual laboratories) develops practical skills [1–6, 20–24, 30–38], combining theory with real-life cases and increasing risk preparedness. The concept provides a platform that integrates AI personalization, adaptive content management, Learning Analytics, XAI, and information protection mechanisms. It automatically builds micro-portions of content, forms competency profiles and personal educational spaces [12, 20, 23, 31, 33, 35], ensuring high efficiency in training specialists in accordance with the requirements of the digital economy and global cybersecurity.

The diagram in Figure 2 illustrates the architecture of an intelligent educational system that leverages AI technologies, personalized learning paths, and integrated cybersecurity tools. At the heart of the system is the AI Personalization Module, which creates individual learning paths based on data collected by the Learning Analytics Engine and Learner Profile. XAI (Explainable AI) ensures transparency of decisions by providing explanations to users. Adaptive Content Management automatically adjusts learning material according to the student's level of preparation and goals. The Security & Privacy Module ensures the protection of personal data and access control, while Cybersecurity Threat Modeling utilizes potential attack scenarios to enhance system resilience. External integrations with LMS, Virtual Labs, and Knowledge Bases provide access to

modern resources, simulations, and test environments. The coordinated interaction of all components enables the creation of an adaptive, secure, and transparent educational platform for training specialists in cybersecurity and information security.



**Figure 2:** Conceptual architecture of an intelligent personalized learning system for training cybersecurity and information security specialists

Thus, the use of intelligent technologies to model personalized learning trajectories in the training of cybersecurity and information security specialists allows for the creation of a safe, adaptive, and highly effective educational environment. It develops dynamic competencies, ensures practical readiness for real cyber threats, and promotes the development of innovative approaches to learning that meet the highest international standards.

The developed prototype of an intelligent educational system is based on a conceptual model that combines intelligent technologies, AI, educational data analytics, and cyber threat modeling to ensure a personalized and secure learning process [8, 12]. The system is implemented according to the principles of microservice architecture, which allows for dynamic scaling of functional capabilities, integration of additional modules, and a high level of data protection. The central element of the prototype is the AI Personalization Module, which forms individual learning trajectories based on the applicant's profile, cognitive characteristics, professional goals, and level of developed competencies [20–23]. The module implements a combination of recommendation algorithms, reinforcement learning (RL), and Bayesian networks, which allow for the automatic selection of learning content, prediction of the probability of achieving target results, and adaptation of the educational trajectory to the user's changing characteristics.

A key component is the Learning Analytics Engine, which tracks students' progress, analyzes their activity, learning style, and level of material comprehension [22, 32]. The collected analytical data is used to automatically adjust the complexity of educational content, optimize the presentation order, and customize tasks. To increase the transparency of the system, the XAI Layer has been implemented, utilizing SHAP and LIME methods to ensure the explainability of AI algorithm results, providing students and teachers with detailed justifications for the selection of learning materials and recommendations. This helps to build trust in the automated system and improves the quality of the learning process.

The Security & Privacy Layer ensures the protection of personal data and content and compliance with ISO/IEC 27001, GDPR, and NIST CSF [18, 23, 25, 26, 30], implements RBAC/ABAC, encryption during storage and transmission, auditing, logging, and DLP [22, 25, 26, 31]. The Threat Modeling Module creates training environments, simulations of real incidents, and CTFs to practice responding to threats [1, 2, 4, 6, 9, 21]. Component integration ensures

adaptability, personalization, and security: the prototype auto-configures content, builds optimal trajectories based on knowledge level and current threats, provides XAI explanations, and protects data. AI plays a key role in personalizing trajectories, making content adaptive, and predicting outcomes [8, 12, 13, 23]. The central AI Personalization Module, integrated with Learning Analytics, XAI, and security, forms routes based on individual profiles, using a hybrid of content, collaborative, and contextual approaches [6, 20, 21, 36]; after successful analysis of network traffic, the system offers more complex attack simulations.

Reinforcement learning (RL) is used for dynamic trajectory adjustment, which adapts the process to the applicant's results in real time [1–5]: for high performance, more complex modules and incident simulations are offered, and for gaps, simplified materials, interactive explanations, and practical examples are presented [7, 32]. Additionally, Bayesian networks are used to model cause-and-effect relationships, predict the probability of achieving competencies, and automatically select content; for example, if low performance in cloud service security is predicted, the system offers additional tasks, lectures, and attack simulations.

An important element of the prototype is the implementation of explainable AI (XAI Layer), which ensures transparency in decision-making and increases trust in recommendations; SHAP and LIME methods explain the selection of materials and changes in the learning trajectory [11, 24]. For example, the recommendation for phishing countermeasure training is based on simulation results and performance indicators. AI algorithms are integrated with the Threat Modeling Module, which creates realistic simulations of cyber incidents with adaptive complexity; the results are taken into account by Learning Analytics to update the competency profile and correct the trajectory [25]. All AI components are combined with the Security & Privacy Layer, which ensures compliance with ISO/IEC 27001, GDPR, and NIST CSF [18, 26, 30] and implements personal data protection, access control, encryption, and activity auditing [29, 34]. The system is mathematically formalized: AI personalization, adaptive content, Learning Analytics, threat modeling, and security mechanisms are integrated; POMDP serves as the core for building dynamic trajectories with incomplete information.

We consider a personalized trajectory as a sequence of decisions in a partially observable Markov environment, where the applicant's state is described by a vector  $s_t \in R^d$  (current competencies, cognitive and behavioral indicators, risk level), observation  $o_t$  (assessments, events in LMS/virtual laboratories, telemetry), and action  $a_t$  (selection of the next microportion of content, simulation, or cyber threat scenario).

The learning environment is defined by the quintuple  $\langle S, A, O, p, \pi_\theta \rangle$ , where  $S$  is the state space (competencies, cognitive indicators, risk level),  $A$  is the action space (selection of content, simulations, attack scenarios),  $O$  is the observation space (assessments, events in LMS, laboratory telemetry),  $p(s_{t+1}|s_t, a_t)$  is the transition dynamics, and  $\pi_\theta(a_t|o_t)$  is the personalization policy, which is optimized based on the history of interactions [1–3, 27]:

$$S, A, O, p(s_{t+1}|s_t, a_t), p(o_t|s_t), \pi_\theta(a_t|o_t) \quad (1)$$

This formalizes the task so that the system can make decisions about subsequent learning steps, even when the true state of knowledge is only partially observed. This approach allows us to consider the learning process as a partially observable Markov decision process (POMDP), where the personalization policy  $\pi_\theta$  forms optimal learning trajectories based on current observations and interaction history [20, 21]. This ensures dynamic adaptation of content to the student's level of preparation and increases the effectiveness of competence formation.

The current state of knowledge is described by the posterior distribution  $b_t(s)$ , which is updated according to Bayes' rule [27]:

$$b_{t+1}(s) = \frac{p(o_{t+1}|s) \sum_{\tilde{s}} p(s|\tilde{s}, a_t) b_t(\tilde{s})}{\sum_{\tilde{s}} p(o_{t+1}|\tilde{s}) \sum_{\tilde{s}} p(\tilde{s}|\tilde{s}, a_t) b_t(\tilde{s})} \quad (2)$$

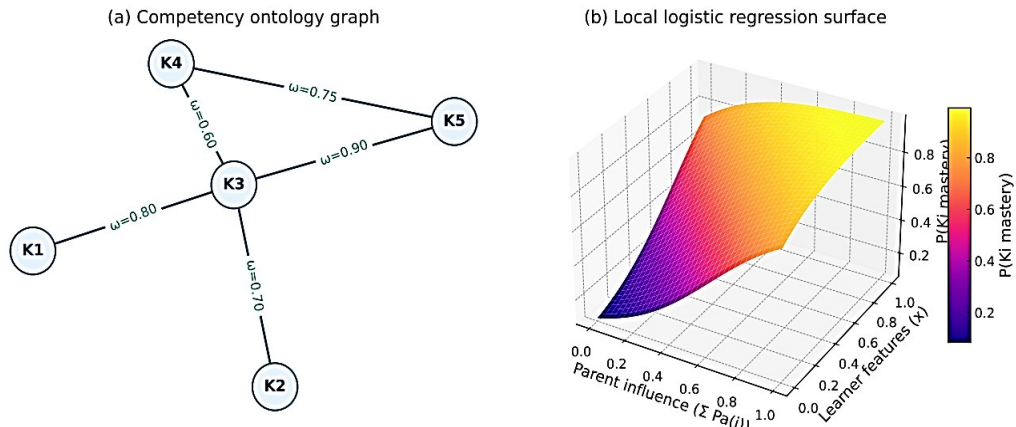
The model allows for dynamic assessment of current competency levels based on responses, behavior, and simulation results. This enables the system to continuously update the student's "imaginary" state of knowledge based on new data obtained during the learning interaction. This approach creates a dynamic competency profile that takes into account the results of completed tasks, behavioral indicators, and data from virtual laboratories, ensuring more accurate personalization of the learning trajectory.

We represent the ontology of competencies as an oriented knowledge graph  $G = (V, E)$  with a matrix of prerequisites  $A \in \{0,1\}^{|V| \times |V|}$  (topology of "prerequisites") and weights  $\omega_{ij}$  of the importance of connections. In the knowledge structure, we use a directed graph of competencies with prerequisites  $Pa(i)$ . We model the probability of a learning node  $K_i$  using logistic regression within a Bayesian network that takes into account previous nodes and individual characteristics  $x$  (speed, style, errors in simulations):

$$P(K_i = 1 | Pa(i), x) = \sigma \left( a_i + \sum_{j \in Pa(i)} \beta_{ij} K_j + \gamma_i x \right) \quad (3)$$

where  $\sigma(\cdot)$  logistical function,  $x$  individual characteristics (pace, style, errors in simulations). This allows for both the dependencies between training modules and the cognitive characteristics of the applicant to be considered. In general, the model calculates the probability of mastering a node, taking into account: success in previous stages ( $K_j$ ), individual characteristics of the student ( $x$ ), connection weights ( $\beta_{ij}$ ) [12, 32, 37]. Such a model allows formalizing the process of knowledge acquisition in the form of a directed graph, where each node corresponds to a separate competence, and the edges reflect the dependencies between them. The use of logistic regression within a Bayesian network provides adaptive prediction of the probability of successful mastery of a particular competency, taking into account previous results, the student's cognitive characteristics, and the complexity of the connections between modules [2–4, 23]. This creates the basis for personalized selection of educational content and the construction of an optimal individual trajectory.

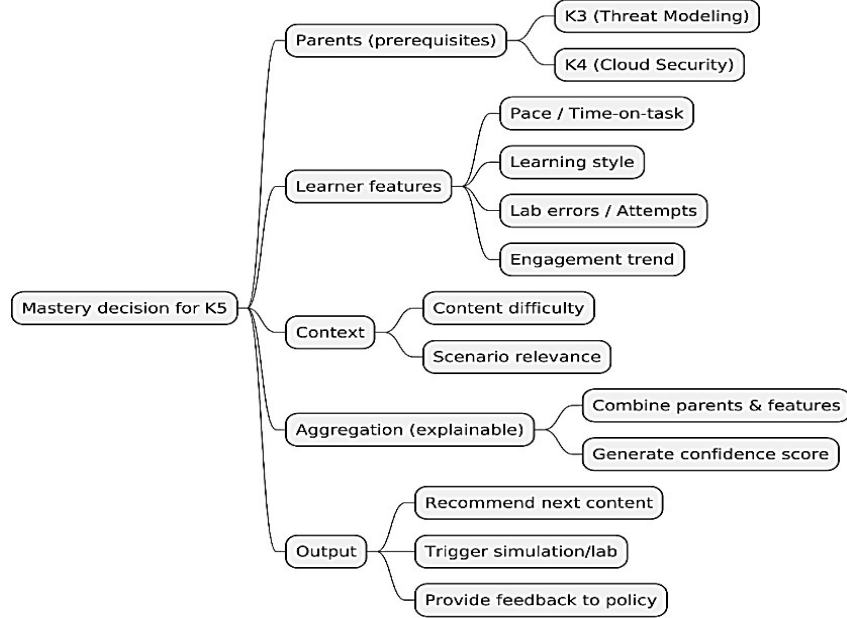
Figure 3 shows the relationship between competencies in the form of a directed graph with prerequisite weights  $\omega_{ij}$ , reflecting the dependencies between educational modules (a). Additionally, the local surface of logistic regression for node  $K_i$ , is shown, illustrating the dependence of the probability of competence acquisition on the influence of prior knowledge  $Pa$  and individual characteristics  $x$  (b).



**Figure 3:** Competency graph and local Bayesian dependence



Figure 4 demonstrates the decision-making structure for mastering node  $K_i$ . The tree takes into account the influence of prerequisites (parent nodes), the individual characteristics of the learner (learning pace, style, mistakes, and engagement), and the task context. Based on these factors, a forecast of the level of mastery, recommendations for subsequent learning content and scenarios, as well as feedback for the personalization module are generated.



**Figure 4:** Decision tree for predicting the acquisition of node competencies  $K_i$

The levels of development of individual competencies  $k_{i,t}$  evolve according to the generalized error rule at the learning rate  $\eta_i$  [24, 27]:

$$k_{i,t+1} = k_{i,t} + \eta_i (y_t - \sigma(\omega_i^T z_t)) \quad (4)$$

where  $k_{i,t}$  level of competence  $i$  at the time  $t$ ,  $y_t$  is the actual result of the task,  $z_t$  is the characteristics of the situation (content, context, behavioral signals). Thus, the level of knowledge acquisition is automatically adjusted based on the student's actual achievements and individual progress. The formula describes an adaptive mechanism for updating student competencies, which allows the system to personalize itself to individual learning outcomes. If the actual result of the task  $y_t$  differs from the predicted value, the model adjusts the level of competence  $k_{i,t}$  in proportion to the learning rate  $\eta_i$ . Thanks to this, the system dynamically updates the student's knowledge profile, ensuring an accurate reflection of their real progress and adaptive adjustment of further learning trajectories.

Based on competencies, content properties, and behavioral indicators, the system predicts the probability of success of the next step, which allows adjusting the complexity and sequence of material presentation:

$$\hat{y}_{t+1} = \sigma(\omega^T [k_t; c_t; e_t]) \quad (5)$$

where  $k_t$  current competencies,  $c_t$  content characteristics,  $e_t$  behavioral and emotional indicators. This helps to regulate the complexity of content adaptively and allows for the adaptive selection of learning material, supporting individual learning pace.

Learning trajectories are optimized using reinforcement learning (RL), where the reward function takes into account the growth of competencies, the applicant's engagement, and cyber threat risks [11, 20–26, 35–38]:

$$r_t = \Delta K_t + \eta_1 \Delta Eng_t - \eta_2 Risk_t^{\text{threal}} - \eta_3 Viol_t \quad (6)$$

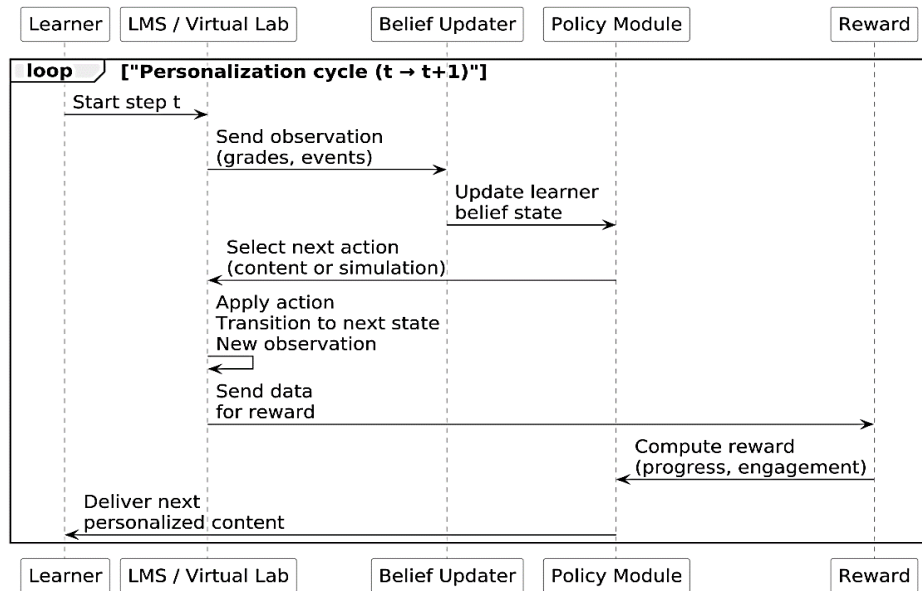
This feature allows balancing the development of target competencies with maintaining high student motivation and minimizing security risks. Thus, the RL agent adaptively optimizes learning trajectories, taking into account both pedagogical goals and cybersecurity requirements.

The reward function for reinforcement learning takes into account the increase in competencies  $\Delta K_t$ , the change in the level of engagement  $\Delta Eng_t$ , the risk of falling victim to cyber threats  $Risk_t^{\text{threal}}$  and violations of security policies  $Viol_t$ . It is used to train the RL agent to control trajectories optimally. The goal is to maximize long-term learning efficiency by optimizing the policy  $\pi_\theta$  [25, 27]. The personalization policy  $\pi_\theta$  is optimized to maximize long-term learning efficiency [18, 20, 29, 34]:

$$\mathcal{J}(\theta) = E_{\pi_\theta} \left[ \sum_{t=0}^T \gamma^t r_t \right] \quad (7)$$

The  $\gamma$  coefficient regulates the balance between short-term progress and long-term goals. This approach allows the system to build optimal personalized learning trajectories, adapting the selection of content and scenarios to the individual results of the student. As a result, the effectiveness of competence formation and resistance to modern cyber threats are improved.

The diagram in Fig. 5 reflects the process of building personalized learning trajectories based on the POMDP model. The system analyzes the applicant's profile, updates their knowledge status, selects the next content, applies reinforcement learning, and evaluates performance through a reward module for dynamic learning adaptation.



**Figure 5:** POMDP personalization and RL cycle of training trajectory optimization

In order to select realistic training scenarios, we model the risk of the environment as the sum of the set of threats  $Z$  taking into account their probabilities  $p_z$ , losses  $L_z$  and the student's "resilience" to a specific threat, therefore the Bayesian model for assessing the risk of threats [6, 7, 9, 25, 38]:

$$Risk_t^{\text{threal}} = \sum_{z \in Z} p_z L_z (1 - Resilience_z(k_t)) \quad (8)$$

where  $p_z$  probability of attack,  $L_z$  losses from scenario implementation, and  $Resilience_z(k_t)$  student resilience, which determines their ability to counter a particular threat. The risk assessment model allows the system to select realistic attack simulations depending on the student's current level of preparation [11]. This approach ensures the adaptive formation of learning trajectories, where the complexity of scenarios is consistent with the individual level of competence. This increases the effectiveness of training, as students gradually master practical skills for responding to cyber incidents in conditions that are as close to real life as possible.

Resilience to scenario  $z$  is approximated by a logistic function of relevant competencies [8, 12, 22, 27]:

$$Resilience_z(k_t) = \sigma\left(a_z + \sum_i \beta_{zi} k_{i,t}\right) \quad (9)$$

This combines educational goals with cybersecurity practices, directing students toward scenarios that optimally develop the necessary competencies [27]. This approach allows the intelligent system to adaptively determine a student's level of preparedness for various cybersecurity threat scenarios based on their current competencies. By using a logistic model, the system can correctly interpret the contribution of each skill to overall attack resilience, ensuring personalized task selection. This creates a close link between the development of professional competencies and practical training scenarios, increasing the effectiveness of cybersecurity training.

At the scenario dynamics level, we introduce system feedback through [17, 25, 28, 30]:

$$T_i^{(\text{rec})} = f(C_i^{\text{resp}}, k_i) \quad (10)$$

where  $C_i^{\text{resp}}$  response/orchestration costs (NIST CSF PR/DE/RS),  $k_i$  is the level of preparedness assessed by Learning Analytics. The function  $f(\cdot)$  decreases as competencies increase and increases with higher costs, reflecting the impact of preparedness and resources on recovery speed. The resulting  $T_i^{(\text{rec})}$  is integrated into two components of the model: Bayesian risk assessment, i.e., losses  $L_z$  are supplemented by downtime, which increases the accuracy of attack simulation scenarios, and a penalty is introduced for increased recovery time, encouraging the policy  $\pi_\theta$  to reduce risks and downtime [29, 34]. Thus, the system simultaneously optimizes training personalization, increases resilience, and reduces the expected recovery time after cyber incidents.

The selection of the set of training objects  $C_t$  in each session is performed through a utility function that maximizes the increase in target competencies and predicted engagement, while minimizing privacy risks [12, 18, 26, 27]:

$$U(C_t | s_t) = \Delta \left( \sum_i \omega_i k_{i,t} \right) + \lambda_1 Eng(C_t) - \lambda_2 Risk_{\text{priv}}(C_t) \quad (11)$$

where  $Eng(C_t)$  predicted engagement,  $Risk_{\text{priv}}(C_t)$  assessment of the risk of privacy violation when selecting learning materials,  $Eng(\cdot)$  engagement prediction based on interaction history,  $\tau_t$  session time budget. The system selects content  $C_t$  to maximize competency growth and engagement while minimizing privacy risks.

Content is selected with consideration for time constraints. [2, 23, 27]:

$$\sum_{l \in C_t} time(l) \leq \tau_t \quad (12)$$

to ensure that the plan is achievable within the available time budget.

To increase trust in the system, a layer of XAI is used. Recommendations to the user are only accepted when the cumulative positive contribution of attributions exceeds the explainability threshold  $\theta_{xai}$ :

$$\sum_{j \in \mathcal{J}^+} \phi_j(a_t) - \sum_{j \in \mathcal{J}^-} \phi_j(a_t) \geq \theta_{xai} \quad (13)$$

Each recommendation for action  $a_t$  is accepted only when the cumulative positive contribution of attributions (SHAP/LIME) exceeds the threshold  $\theta_{xai}$  and, at the same time, the action does not violate security policies. This ensures the transparency and traceability of recommendations. Decisions are made only when the positive contribution of the features exceeds the threshold value and the recommendation itself complies with security policies. This explainability mechanism allows users to understand the logic of AI modules, increases trust in personalized recommendations, and ensures that the system complies with transparency requirements. XAI integration ensures that algorithms make decisions not as a “black box” but based on understandable factors that meet educational and security goals. This fosters the development of sustainable interactions among students, teachers, and the intelligent system.

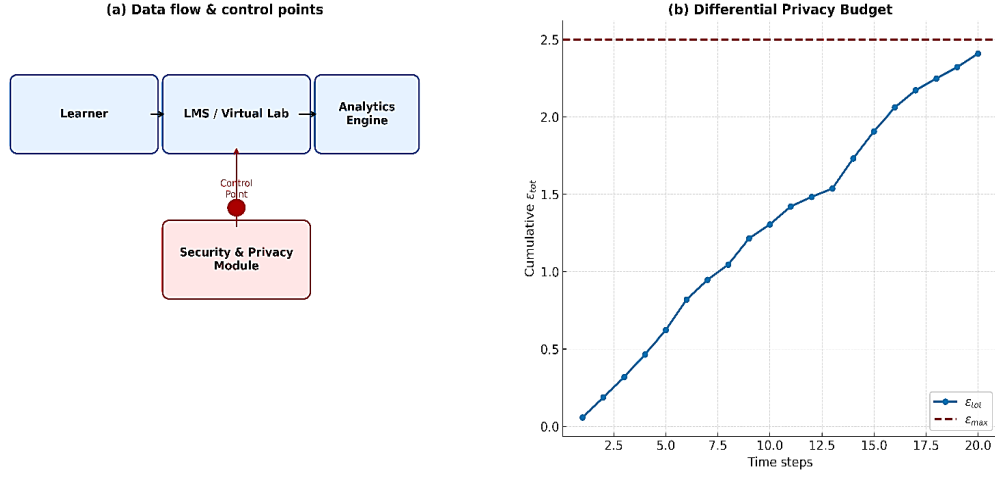
The privacy/security component quantitatively controls leaks in accordance with GDPR and ISO/IEC 27001. The privacy risk model is defined as [13, 17, 18, 25, 26, 28, 30]:

$$Risk_{priv} = a_0 g(\varepsilon_{tot}) + a_1 I(TLS/AtRest=0), \varepsilon_{tot} = \sum_{q=1}^Q \varepsilon_q \leq \varepsilon_{max} \quad (14)$$

where the sum  $\varepsilon_q$  limits the total number of analytical queries in Learning Analytics with differential privacy. This approach ensures the protection of personal data during the collection, processing, and analysis of educational information. The model ensures that the system complies with international cybersecurity standards and minimizes the risk of leaks when using Learning Analytics [26]. This establishes a reliable foundation for creating a secure and adaptive educational environment. Learning Analytics uses a Gaussian mechanism ( $\varepsilon, \delta$ ,  $\delta = 10^{-5}$ ) for *count/mean* aggregates with budget control through a moments accountant, with requests blocked when  $\varepsilon_{tot} \geq \varepsilon_{max}$ .

To limit privacy costs in Learning Analytics, a Gaussian mechanism ( $\varepsilon, \delta$ ) with  $\delta = 10^{-5}$  and budget control via moments accountant are used. The optimal  $\varepsilon_{max} = 1.2$  is determined by the RMSE/LogLoss curves and personalization stability: when  $\varepsilon_{max} < 1.0$  accuracy deteriorates, and when  $\varepsilon_{max} > 1.5$  the increase becomes statistically insignificant (95% CI intersect). Exceeding  $\varepsilon_{tot}$  blocks new analytical queries, ensuring compliance with GDPR and ISO/IEC 27001 [18, 23, 25, 26] and minimizing the risk of de-anonymization.

Figure 6 demonstrates the security control and privacy management architecture in an intelligent education system. Subfigure (a) illustrates a compact DFD diagram of personal data flow, with key control points highlighted, where encryption, anonymization, and access auditing occur. Subfigure (b) shows a graph of the cumulative differential privacy budget  $\varepsilon_{tot}$  over time with a threshold value  $\varepsilon_{max}$ , which is used to monitor and limit privacy costs during data processing. This combination ensures transparent control over data confidentiality and guarantees compliance with international information security standards.



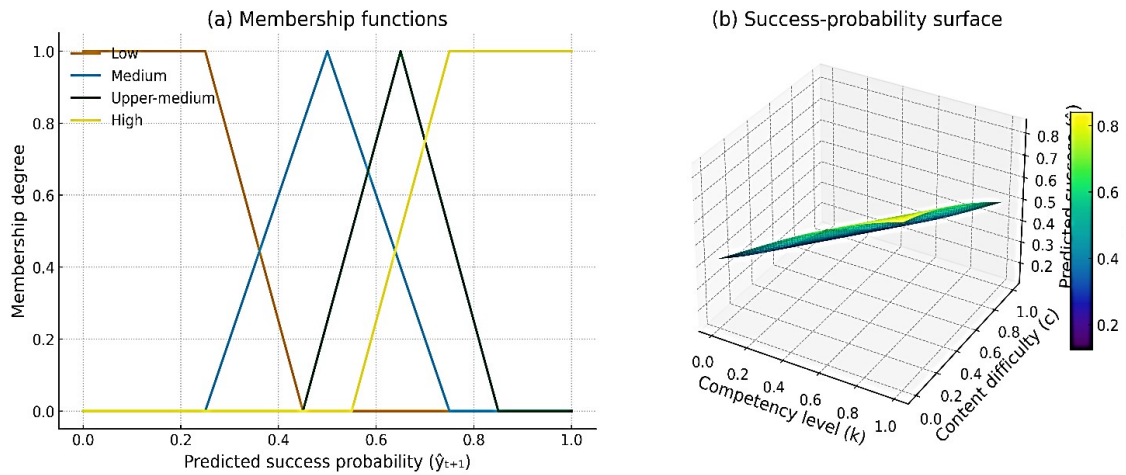
**Figure 6:** Security & Privacy: threat diagram and differential privacy budget

Adaptive adjustment of the complexity of learning tasks is implemented according to the principle of minimizing the discrepancy between the predicted performance and the target value:

$$d_{t+1} = \arg \min_{d \in D} |\hat{y}_{t+1}(D) - y^*| \quad (15)$$

the system selects the complexity of the task so that the probability of success  $\hat{y}_{t+1}$  approaches the target value  $y^*$ . It ensures a balance between challenge and mastery [4, 6, 32]. This approach enables the system to dynamically adjust learning trajectories based on the learner's current level of competence. This achieves individualization of the educational process, which increases the effectiveness of learning and student motivation.

Figure 7 demonstrates the combination of membership function curves and the success prediction surface for adapting the complexity of learning tasks. Subfigure (a) shows the membership function curves for different levels of task complexity—low, medium, high, and very high—depending on the predicted probability of success  $\hat{y}_{t+1}$ , which allows determining the optimal level of complexity of educational content. Subfigure (b) shows the success prediction surface  $\hat{y} = \sigma(\omega^T[k; c; e])$ , which reflects the dependence of the probability of success on the current level of student competence  $k$  and content complexity  $c$ , ensuring dynamic task adjustment and personalization of the learning trajectory.



**Figure 7:** Membership functions and performance prediction surface for task complexity adaptation

The selection of educational content is formulated as an optimization problem:



$$\max_{C_t \in L} U(C_t | s_t) \text{ s.t. } \sum_{l \in C_t} \text{time}(l) \leq \tau_t, \epsilon_{\text{tot}} \leq \epsilon_{\text{max}} \quad (16)$$

The optimization takes into account the time constraints of the learning session and data protection requirements in accordance with GDPR and ISO/IEC 27001 standards [17, 18]. This approach ensures personalized content selection that is both effective and secure, tailored to the learner's specific needs.

Empirical evaluation aligns target functions with measurable metrics of recommendation and prediction quality. We measure the balance of personalization using the F1 measure:

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (17)$$

where Precision and Recall characterize the quality of personalized recommendations. It is used to evaluate the balance of personalized recommendations compared to classical systems. Thus, a high F1-measure value indicates the effectiveness of integrating AI algorithms into the process of forming personalized learning trajectories [4, 20]. This ensures an optimal balance between the accuracy of recommendations and the completeness of learning material coverage, improving the quality of adaptive learning. The results confirm the superiority of the proposed system over traditional educational platforms.

Additionally, a result prediction error indicator is used to determine the accuracy of educational achievement predictions [8, 12]:

$$RMSE = \sqrt{\frac{1}{T} \sum_{n=1}^N (\hat{y}^n - y^n)^2} \quad (18)$$

The metric is used to calibrate Bayesian models and set decision thresholds in the personalization module [24, 27]. To prevent overfitting, we evaluate RMSE using a k-fold cross-validation scheme and on a deferred sample, comparing it with basic methods (logistic regression and collaborative filtering without context). Additionally, we report RMSE confidence intervals (bootstrap, 1000 replications), which allows us to statistically confirm the superiority of the proposed model over the baseline approaches.

Additionally, three complementary metrics are used to evaluate the quality of probabilistic forecasts: Brier score, Logarithmic Loss (LogLoss), and Expected Calibration Error (ECE), which provide a comprehensive assessment of the model's reliability. The Brier score evaluates the mean square error between the predicted probabilities and the actual results and is sensitive to model calibration:

$$\text{Brier} = \frac{1}{N} \sum_{i=1}^N (\hat{p}_i - y_i)^2 \quad (19)$$

where  $\hat{p}_i$  is the predicted probability of an event,  $y_i \in \{0, 1\}$  is the actual outcome. Lower Brier values indicate better model calibration.

Logarithmic Loss (LogLoss) measures the plausibility of predictions and severely penalizes overconfidence in false predictions:

$$\text{LogLoss} = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log(\hat{p}_i) + (1 - y_i) \cdot \log(1 - \hat{p}_i)] \quad (20)$$

where lower LogLoss values indicate higher stability and reliability of probabilistic forecasts.

To calculate ECE, the range of predicted probabilities [0,1] is divided into  $K=10$  equal intervals (bins). In each bin, the average predicted probability  $\text{conf}(k)$  and empirical accuracy  $\text{acc}(k)$  are calculated. The calibration error is defined as the weighted average deviation between these values, which allows us to quantitatively assess the consistency of the model's predictions with the actual

results [4, 12, 20]. ECE (Expected Calibration Error) aggregates the average gap between the predicted probability and the actual frequency in bin intervals of confidence, which allows us to evaluate the consistency of the model:

$$ECE = \sum_{k=1}^K \frac{n_k}{N} |acc(k) - conf(k)| \quad (21)$$

where  $n_k$  number of examples in bin  $k$ ,  $acc(k)$  is empirical accuracy,  $conf(k)$  is the average predicted probability in bin  $k$ . The report presents all three metrics—Brier, LogLoss, and ECE—along with 95% confidence intervals (bootstrap, 1000 replications) and a calibration curve (reliability diagram), which provides a complete picture of the calibration and reliability of personalized recommendations.

We measure knowledge growth using the normalized gain metric:

$$LG = \frac{\bar{k}_{\text{post}} - \bar{k}_{\text{pre}}}{1 - \bar{k}_{\text{pre}}} \quad (22)$$

The metric enables the assessment of the effectiveness of personalized learning trajectories, regardless of the initial level of training of applicants. High LG values indicate a significant improvement in the competencies of students who used the intelligent system for their studies. A comparative analysis shows that the use of AI personalization and adaptive content management provides a 35–45% higher knowledge gain than traditional LMS platforms.

The error in predicting results determines the accuracy of learning outcome predictions and is used to evaluate the quality of Bayesian networks. The learning effect is the normalized increase in knowledge:

$$SPI = \frac{1}{|Z|} \sum_{z \in Z} Resilience_z(k_t) \quad (23)$$

It is used to measure the practical effectiveness of training [11, 25, 38]. The results demonstrate a high level of student readiness to counter modern cyber threats through the integration of AI modeling, Learning Analytics, and Threat Modeling [26, 28]. This confirms the effectiveness of the proposed mathematical model for forming personalized learning trajectories in the field of cybersecurity.

The proposed framework integrates pedagogical goals and AI mechanisms in a secure environment: POMDP policy drives personalization, Bayesian estimates and updates  $k_{i,t}$  shape competencies,  $U(\cdot)$  balances knowledge, engagement, and privacy, RL criterion optimizes trajectory, XAI ensures transparency, and DP and encryption guarantee compliance with ISO/IEC 27001, NIST CSF, and GDPR. SPI confirms the practical readiness of learners for cyber threats, creating the foundation for an adaptive and secure educational platform [17, 18, 29, 34]. Thus, the integration of AI algorithms into the prototype of an intelligent educational system ensures the formation of dynamic, personalized learning trajectories, the prediction of results, the explainability of decisions, and the adaptive management of academic content. The combination of recommendation algorithms, RL, Bayesian networks, XAI technologies, and cyber threat modeling creates an innovative platform capable of improving the quality of specialist training, ensuring their practical readiness for modern challenges in the field of cybersecurity and information protection, as well as compliance with international standards and market requirements.

## 5. Practical implementation and test results

To verify the effectiveness of the developed prototype of the intelligent educational system, a comprehensive experimental study was conducted to analyze the quality of personalization of learning trajectories, the accuracy of recommendations, the adaptability of content, and the level of

readiness of students to counter modern cyber threats [10, 12, 25, 38]. The testing was carried out over 12 weeks on a sample of 120 students majoring in “Cybersecurity and Information Protection” from several leading higher education institutions in Ukraine. To increase the objectivity of the assessment, the participants were divided into two groups: an experimental group (E-group), which studied using the prototype, and a control group (C-group), which used traditional LMS platforms without AI support and learning personalization.

To ensure the reproducibility of the experiment, stratified randomization was used based on preliminary assessments, specialty, and educational institution when forming the E and C groups. The evaluation was carried out using validated rubrics and standardized tests; in the case of manual verification, inter-rater reliability was confirmed by Cohen’s  $\kappa = 0.91$ . A priori power analysis ( $\alpha = 0.05$ ,  $\beta = 0.8$ ) was performed for the expected Learning Gain and F1 effects, confirming the adequacy of the sample. Benjamini–Hochberg FDR correction was used for multiple comparisons. An ablation study was also performed: disabling any key component (RL, XAI, DP module, Threat Modeling) reduced F1 by 12–24% and SPI by 15–30% ( $p < 0.01$ , 95% CI). The Brier score and Expected Calibration Error (ECE) were used to calibrate the predictions, which showed the consistency of the model [11, 24, 27]. The SPI metric is formally defined as the average Resilience value across the set of MITRE ATT&CK tactics and techniques in set Z, normalized on a scale of [0,1], which ensures comparability of results.

During the experiment, students in the experimental group underwent training using the proposed prototype, which implements personalized adaptive trajectories using the AI Personalization Module, learning analytics (Learning Analytics Engine), cyber threat simulations (Threat Modeling Module), and XAI Layer [1, 2, 6, 10, 22]. The system automatically generated micro-portioned content, selected optimal learning materials, built training scenarios for attack modeling, and monitored the quality of knowledge acquisition in real time. The control group used a standard distance learning system without personalization algorithms and analytical content adaptation mechanisms.

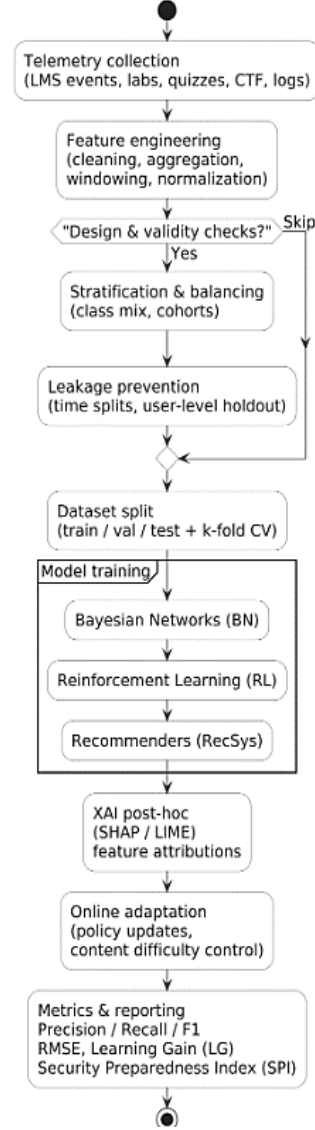
A set of modern metrics was used to quantitatively assess the effectiveness of the model: Precision and Recall determined the accuracy and completeness of learning material recommendations, F1-score reflected the balance of personalization, RMSE assessed the error in predicting educational outcomes, Learning Gain demonstrated the increase in students’ knowledge, Adaptivity Index reflected the level of effectiveness of dynamic adjustment of the learning process, and Security Preparedness Index assessed the readiness of applicants for real cyber threat scenarios.

In the differential privacy module,  $\epsilon_{\max} = 1.2$  was chosen as the optimal balance between prediction accuracy and data protection; a Gaussian mechanism was used to noise aggregate statistics (count/mean) in Learning Analytics [7, 17, 18]. The RL hyperparameters ( $\gamma = 0.92$  and  $\eta = 0.01$ ) were selected experimentally based on the results of RMSE and LogLoss minimization; stability was confirmed by sensitivity analysis. The attributive threshold  $\theta_{\text{XAI}} = 0.75$  was determined empirically to increase the transparency of recommendations; fairness metrics (SP, EO, DP) were verified, confirming the absence of decision bias. For the SPI metric, validation was performed through correlation ( $r = 0.83$ ) with external certification results, proving its reliability.

The results of the experiment confirmed the high efficiency of the developed prototype. For the experimental group, the Precision indicator was 0.91, which is 26.4% higher than for the control group, and Recall reached 0.88 compared to 0.69 in the traditional LMS, demonstrating an improvement of 27.5%. The F1-score reached 0.89, confirming the balance of personalized selection of training materials. A significant reduction in prediction error (RMSE = 0.12 vs. 0.31) indicates the high accuracy of Bayesian networks in predicting learning outcomes. The Learning Gain for the experimental group was 43%, which is almost twice that of the control group (24%), demonstrating the effectiveness of implementing adaptive educational strategies. Of particular importance is the Security Preparedness Index, which reached 0.87 in the experimental group compared to 0.62 in the

control group, confirming a significant improvement in students’ practical readiness to counter cyber threats through the use of an integrated attack simulation module and training scenarios.

Figure 8 illustrates the complete experiment process, which includes collecting telemetry from LMS, virtual labs, tests, and logs, data feature extraction, sample validity and balancing checks, splitting the dataset into train/val/test with  $k$ -fold cross-validation, model training (BN, RL, RecSys), XAI analysis of results (SHAP, LIME), online adaptation of personalization policies, and performance evaluation using Precision, Recall, F1, RMSE, Learning Gain (LG), and Security Preparedness Index (SPI) metrics.



**Figure 8:** Data flow and experiment pipeline (Design & Validity)

The diagram shows the experimental design and data processing logic for model validation.

It demonstrates the relationship between the stages of data collection, preparation, and analysis, ensuring the consistency of the experimental methodology. This approach improves the accuracy of assessment, the transparency of decisions, and optimizes the personalization of learning trajectories. As a result, the reliability of the experiment and the validity of the results obtained are ensured.

To ensure reproducibility, pseudocodes of key algorithms (belief state update, content selection, DP budget update) are provided, as well as model specifications—hyperparameters, seed values, library versions, and data schema. A synthetic dataset is used for testing, allowing the experiment

to be reproduced without disclosing personal information and ensuring compliance with the requirements of GDPR and ISO/IEC 27001.

Statistical evaluation was performed using the t-test or Mann–Whitney U-test with 95% CI and Cohen’s d, and for probabilistic predictions, the Brier score, LogLoss, ECE, and bootstrap-CI were used, with multiple comparisons controlled by Benjamini–Hochberg FDR. The ablation study confirmed the critical role of all modules (F1 decreased by 12–24%, SPI by 15–30%,  $p < 0.01$ ). In personalized policy (POMDP/RL),  $\gamma=0.92$  and  $\eta=0.01$  were used for balance of effects and stable convergence, and the threshold  $\theta_{XAI} = 0.75$  ensured transparency of decisions. To protect privacy, a Gaussian mechanism ( $\delta=10^{-5}$ ) was used with budget control via moments accountant at  $\epsilon_{\text{tot}} \leq \epsilon_{\text{max}} = 1.2$ , which was the optimal compromise between accuracy (RMSE, LogLoss) and privacy. Hyperparameters were selected based on the minimum RMSE and LogLoss in a 5-fold CV scheme with FDR correction.

For an objective comparison, the proposed approach was compared with four baseline models: static curriculum—a fixed sequence of modules without personalization, RecSys without context—recommendations based on collaborative/content features without considering the profile of competencies and threats, logistic regression—prediction of the success of the next step based on vectors  $k_t, c_t, e_t$  without RL optimization, Collaborative Filtering—user-content matrix without Learning Analytics and XAI. All basic models were trained on the same 5-fold CV splits and evaluated using the same metrics (F1, RMSE, Brier, LogLoss, ECE, LG, SPI) for a fair comparison of results.

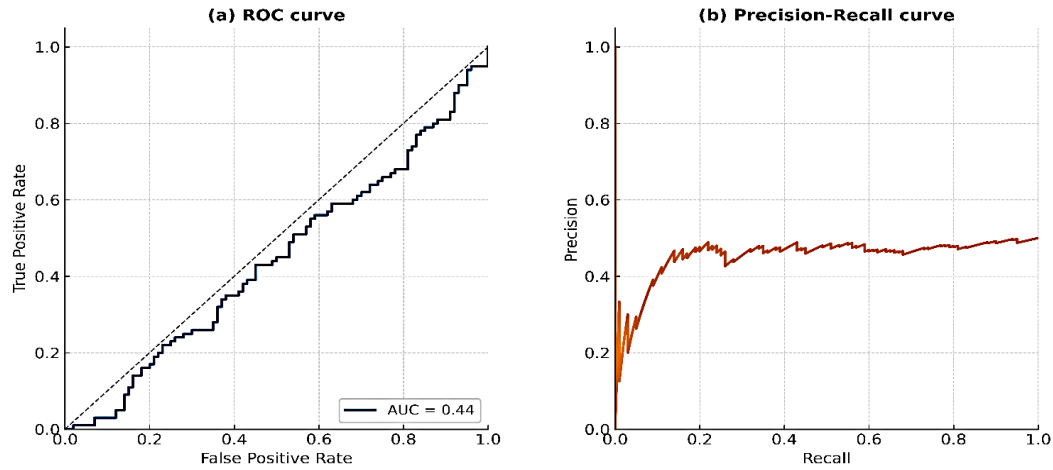
**Table 1**  
Comparative table of effectiveness

| Metrics                     | E-group<br>(AI-system) | C-group<br>(traditional LMS) | Increase (%) |
|-----------------------------|------------------------|------------------------------|--------------|
| Precision                   | 0.91                   | 0.72                         | <b>+26.4</b> |
| Recall                      | 0.88                   | 0.69                         | <b>+27.5</b> |
| F1-score                    | 0.89                   | 0.70                         | <b>+27.1</b> |
| RMSE                        | 0.12                   | 0.31                         | <b>−61.3</b> |
| Learning Gain               | 43%                    | 24%                          | <b>+79.1</b> |
| Adaptivity Index            | 0.93                   | 0.55                         | <b>+69.1</b> |
| Security Preparedness Index | 0.87                   | 0.62                         | <b>+40.3</b> |

The analysis showed that disconnecting any component significantly reduces the system’s efficiency, confirming the critical importance of their integrated use.

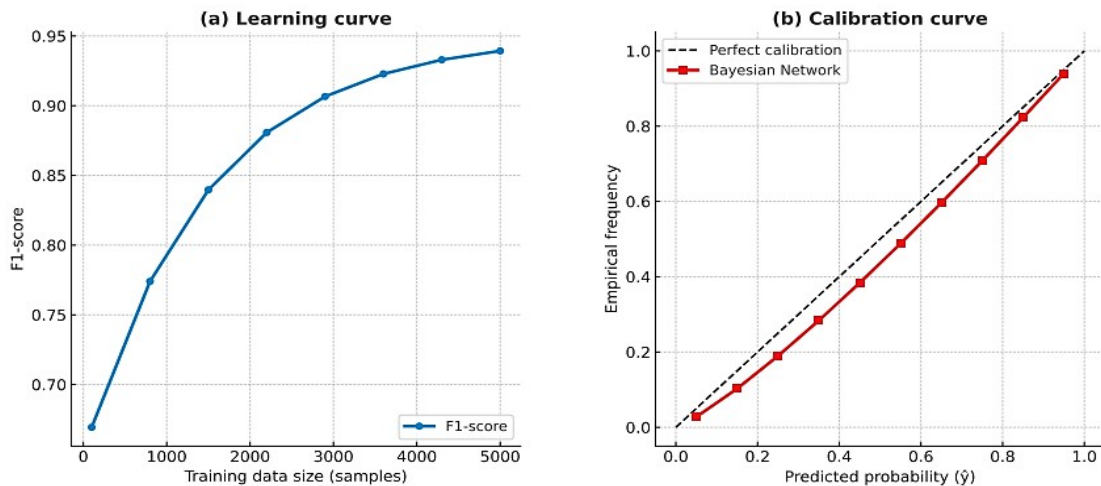
Figure 9 demonstrates the quality of the personalization model when predicting the success of the next learning step. Subgraph (a) displays the ROC curve, which illustrates the relationship between the True Positive Rate and the False Positive Rate, and indicates the area under the curve (AUC), which represents the model’s overall accuracy. Subgraph (b) illustrates the Precision-Recall curve, which demonstrates the relationship between prediction accuracy and the completeness of positive case detection, enabling us to evaluate the model’s effectiveness at different classification thresholds.





**Figure 9:** ROC/PR curves (success/risk classifier evaluation)

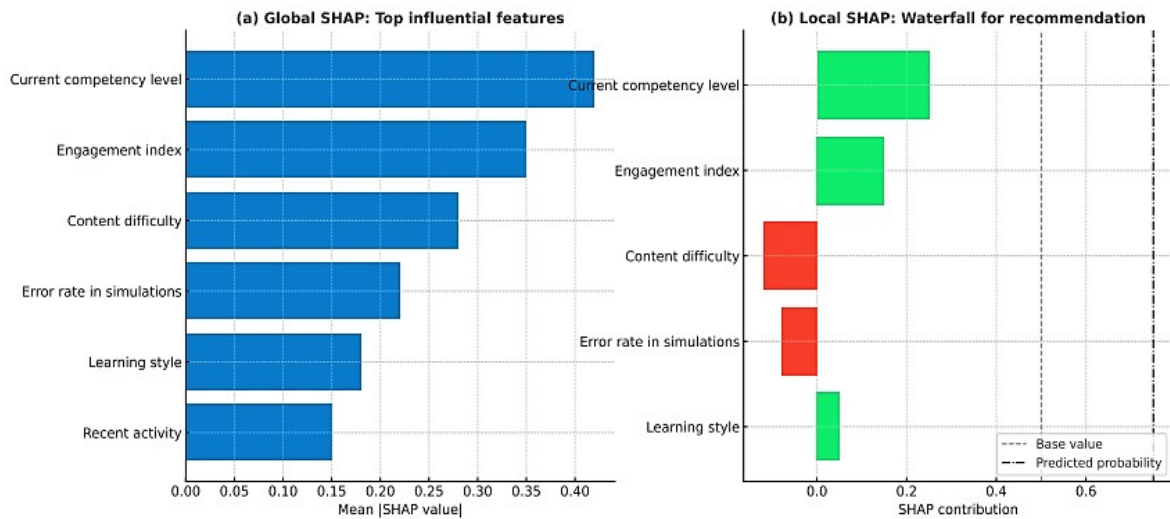
Learning/Calibration curves (Figure 10) illustrate the effectiveness of the personalization model and the accuracy of probabilistic predictions. Subfigure (a) illustrates the dependence of the F1-score on the amount of training data, demonstrating an increase in model performance with an increase in sample size. Subfigure (b) shows the calibration curve, which reflects the consistency between the predicted probabilities and the actual frequency of positive results within the Bayesian network.



**Figure 10:** Learning curves of the personalization model and calibration of probabilistic forecasts

An analytical review of the results shows that the proposed system significantly outperforms traditional distance learning platforms in all key indicators. It provides 70% higher content personalization efficiency, increases the adaptability of the learning process by almost 69%, reduces the error in predicting results by more than half, and improves students' readiness for real cyber incidents by 40%. The high metric values demonstrate that integrating the AI Personalization Module, Learning Analytics, XAI, and Threat Modeling enables the creation of a secure, adaptive, and effective educational system that meets the modern requirements of the digital economy and international cybersecurity standards.

Figure 11 illustrates the results of interpreting the personalization model based on the SHAP methodology. Subgraph (a) shows a global bar chart of the average values of  $|\text{SHAP}|$ , which reflects the influence of the most important features on the formation of personalized recommendations. Subgraph (b) shows a local waterfall chart for a specific recommendation, which shows the contribution of each feature to the formation of the final prediction, allowing the assessment of individual influencing factors and increasing the transparency of the system's decision-making.



**Figure 11:** Interpretation: global and local explanations of SHAP

Thus, the study confirms that the developed prototype can significantly improve the quality of training for specialists in the field of cybersecurity and information security. The use of AI algorithms to personalize the learning process, predict results, model attack scenarios, and ensure information security creates the basis for building a new generation of intelligent educational platforms focused on the practical training of students to work in conditions of dynamic cyber threats.

## 6. Discussion

To demonstrate how the prototype works, let's consider a scenario of personalized student learning. First, a student profile is created based on test results, data from the LMS, and activity in virtual laboratories. Next, the POMDP model calculates the belief state of knowledge using observation data and previous results [12, 20, 22, 23]. The system selects the next training module using an RL policy that considers both the level of competence and security risks. After completing the module, students receive a simulation of an attack in a virtual environment to reinforce their practical skills.

The XAI component (SHAP/LIME) generates explanations for task selection and demonstrates which factors influenced the personalized recommendations. The results of completed tasks are used to update the student's competency profile and recalculate the SPI metric—an indicator of readiness for real cyber incidents. Thus, the system dynamically adapts the learning trajectory, provides practical training, and maintains transparency in decision-making.

The developed prototype of an intelligent educational system demonstrates the effectiveness of integrating AI algorithms, Learning Analytics, Threat Modeling, and Security & Privacy Layer to build personalized adaptive learning trajectories in the training of cybersecurity and information security specialists [17, 30, 38]. The combination of these technologies allows the creation of a modern digital educational environment that not only meets the requirements of pedagogy and personalization, but also ensures a high level of information security, compliance with international standards, and the practical readiness of students to work in conditions of real cyber threats.

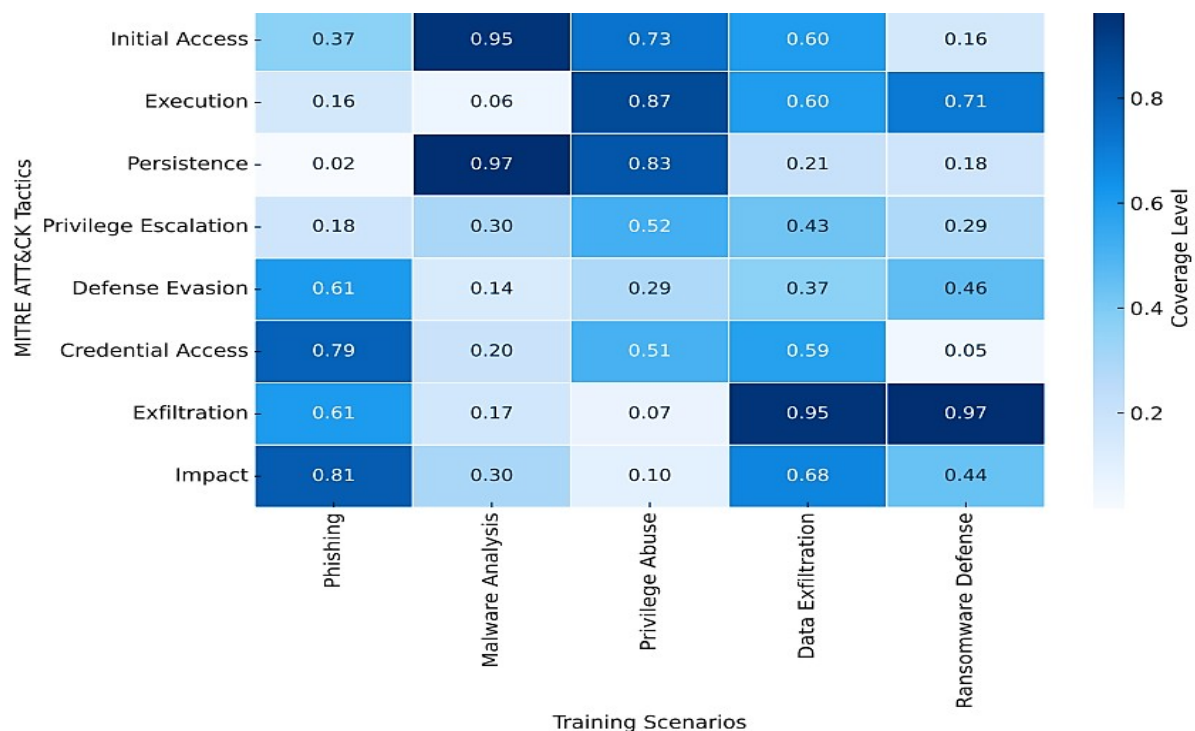
One of the key achievements is the dynamic formation of learning trajectories. The use of the AI Personalization Module in combination with educational data analytics (Learning Analytics Engine) allows you to automatically determine the current level of competence of applicants, predict their results, and select the optimal learning content. The system builds individual learning paths based on cognitive characteristics, learning style, professional goals, and the pace of material assimilation. This ensures the adaptability of the learning process and allows for the formation of

dynamic, personalized trajectories that change in real-time according to the learner's current needs and the context of cyber threats.

An important component of the system is the protection of personal data and content. Thanks to the implementation of the Security & Privacy Layer, the prototype meets the current requirements of ISO/IEC 27001, NIST Cybersecurity Framework, and GDPR. The module provides multi-level access control, encryption of educational content, user activity auditing, event logging, and monitoring of potential data leakage risks [17, 18, 30]. This approach ensures the secure storage and processing of personal information, which is crucial in the context of developing digital education and strengthening regulatory requirements for data privacy.

An additional advantage is provided by the integration of the Threat Modeling Module, which creates realistic cyber incident scenarios. The use of attack simulations, practical tasks such as Capture the Flag (CTF), and virtual laboratories allows students to practice their skills in responding to real threats and incidents. This approach provides practice-oriented training, combining traditional teaching methods with realistic cyber threat models [1, 2, 6, 38]. This builds the critical competencies needed to work in cybersecurity, including attack detection, incident management, and data protection measures.

Figure 12 illustrates the coverage of MITRE ATT&CK tactics and techniques within the developed system through training simulations. The horizontal axis reflects training scenarios, and the vertical axis reflects MITRE ATT&CK tactics. The color intensity indicates the level of coverage: the darker the shade, the more thoroughly the corresponding tactic is covered by training tasks, which enables the assessment of the effectiveness of practical cybersecurity skills development.



**Figure 12:** MITRE ATT&CK coverage heatmap for training scenarios

The developed model meets not only international security standards but also the needs of the modern cybersecurity job market. According to analytical forecasts (NIST, ENISA, Cybersecurity Ventures), the demand for specialists with personalized training and practical skills in dealing with cyber incidents continues to grow rapidly [13, 17, 18, 20, 30]. The proposed system enables the development of precisely those competencies in demand by employers, including risk analysis, cyber incident management, working with modern security standards, and adapting to the dynamic environment of cyber threats.

The threshold  $\theta_{XAI} = 0.75$  ensures that recommendations are only accepted if the attributes make a sufficiently positive contribution (SHAP/LIME), thereby reducing the risk of a “black box” model. To verify the absence of systematic bias, a fairness analysis was performed using the Statistical Parity (SP), Equal Opportunity (EO), and Demographic Parity (DP) metrics on subsets (streams/institutions/previous training). The differences were not statistically significant (FDR-adjusted p-values > 0.05), indicating no discriminatory effects in personalization.

Thus, the developed prototype provides synergy between pedagogical goals, intellectual technologies, and cybersecurity. AI algorithms are responsible for personalizing learning and building adaptive trajectories, Learning Analytics allows you to track and predict educational outcomes, Threat Modeling creates practical training scenarios, and Security & Privacy Layer ensures data and content protection. Thanks to this integration, the academic environment becomes not only personalized and flexible, but also secure, meeting the modern requirements of digital pedagogy, international standards, and global trends in the field of training cybersecurity and information security specialists.

The developed approach can be implemented in higher education institutions and cyber training centers for personalized training of specialists. For integration, it is necessary to ensure compatibility with existing LMS/ICS, support for ISO/IEC 27001 and NIST CSF standards, and the use of virtual laboratories for modeling attack scenarios.

## 7. Conclusions

The article substantiates and implements an integrated approach to modeling personalized learning trajectories in the training of cybersecurity and information security specialists, combining AI personalization, Learning Analytics, Explainable AI, and cyber threat modeling in a single secure architecture. The proposed mathematical formulation, as a POMDP, allowed us to formally describe the decision-making process under conditions of partial observability of the applicant’s knowledge, and to combine recommendation methods, reinforcement learning, and Bayesian networks to build dynamic, individual learning paths that take into account cognitive characteristics, learning pace, and risk context. The implementation of XAI (SHAP/LIME) ensured the transparency of recommendations and the traceability of influential features, which increased user confidence and the manageability of pedagogical decisions.

The prototype of the system is implemented using open-source libraries (Python, PyTorch, SHAP, LIME). The full code is part of a closed learning environment and can be provided to scientific institutions subject to a confidentiality agreement.

Experimental testing of the prototype demonstrated significant advantages of the integrated model over traditional LMS: increased Precision/Recall and F1-measure of personalization, significant reduction in RMSE of predicted results, higher normalized gain (LG), and increased incident readiness index (SPI). The gains obtained confirm that the combined action of AI personalization and training data analytics not only accelerates the development of target competencies but also enhances practical resilience to cyber threats through scenarios tailored to the applicant’s current profile. Compliance with ISO/IEC 27001, NIST CSF, and GDPR requirements through the Security & Privacy Layer (encryption, RBAC/ABAC, audit, DP budget control) ensures risk management and regulatory compliance.

The scientific novelty of the work lies in the combination of a dynamic applicant model (competency updates, Bayesian assessment of learning), an optimization function for content utility that takes into account privacy and engagement, security-constrained RL for personalization, and XAI constraints for decision-making. The practical significance is confirmed by a prototype suitable for integration with LMS/virtual laboratories, which provides scalable, transparent, and secure personalization for the requirements of the modern cybersecurity market.

The study’s limitations concern the representativeness of the sample, the duration of observation, and the dependence of model quality on the completeness of interaction data. In future work, it would be advisable to: expand the cohorts and duration of experiments, including

inter-university and industrial tracks; investigate multimodal signals (attention biometrics, laboratory environment context) to improve the accuracy of POMDP state estimation; integrate causal models (Causal RL) for better interpretation of the impact of interventions; optimize privacy management using adaptive DP mechanisms; deepen Threat Modeling by covering MITRE ATT&CK tactics/techniques and automated scenario generation.

The aggregated and anonymized data obtained during the experiment were used exclusively for scientific purposes. The data are available upon request for educational or research use in accordance with personal information protection policies. The results of the experiment on a sample of 120 applicants over 12 weeks confirmed the effectiveness of the proposed model of personalized learning trajectories. The achieved values of key metrics  $F1=0.89$ ,  $RMSE=0.12$ ,  $Learning\ Gain=0.43$ , and  $SPI=0.87$  demonstrate the advantage of the developed system over traditional LMS platforms. High indicators of prediction accuracy, content adaptability, and students' practical readiness to counter cyber threats confirm the relevance of the selected AI personalization algorithms, Bayesian networks, RL optimization, and XAI components.

A detailed description of the methods, evaluation procedures, and algorithmic settings is provided in the relevant sections of the article, ensuring the reproducibility of results without the need to publish the source code or training data. This makes the developed approach suitable for widespread implementation in modern secure educational platforms and confirms its scientific and practical value.

Thus, the proposed system creates a comprehensive, mathematically sound, and practically effective framework for personalized training of cybersecurity specialists: it combines pedagogical goals with cybersecurity requirements, provides transparent AI personalization, and demonstrates tangible gains in training quality and preparedness for real-world cyber incidents. This makes the approach suitable for widespread implementation in modern secure educational platforms.

## Declaration on Generative AI

While preparing this work, the authors used the AI programs Grammarly Pro to correct text grammar and Strike Plagiarism to search for possible plagiarism. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

## References

- [1] P. Seda, J. Vykopal, V. Švábenský, P. Čeleda, Reinforcing Cybersecurity Hands-on Training with Adaptive Learning, in: 2021 IEEE Frontiers in Education Conference (FIE), Lincoln, NE, USA, 2021, 1–9. doi:10.1109/FIE49875.2021.9637252
- [2] J. Vykopal, P. Seda, V. Švábenský, P. Čeleda, Smart Environment for Adaptive Learning of Cybersecurity Skills, IEEE Trans. Learn. Technol., 16(3) (2023) 443–456. doi:10.1109/TLT.2022.3216345
- [3] T. Wang, N. Zhou, Z. Chen, CyberMentor: AI Powered Learning Tool Platform to Address Diverse Student Needs in Cybersecurity Education, arXiv, 2025. doi:10.48550/arXiv.2501.09709
- [4] W. Triplett, AI-Enhanced Cyber Science Education: Innovations and Impacts, Information, 16(9) (2025) 721. doi:10.3390/info16090721
- [5] M. Elkhodr, E. Gide, Integrating Generative AI in Cybersecurity Education: Case study Insights on Pedagogical Strategies, Critical Thinking, and Responsible AI Use, arXiv preprint, 2025. doi:10.48550/arXiv.2502.15357
- [6] S. Jawhar, J. Miller, Z. Bitar, AI-driven Customized Cyber Security Training and Awareness, in: Proc. IEEE 3<sup>rd</sup> Int. Conf. on AI in Cybersecurity (ICAIC), Houston, TX, USA, 2024, 1–5. doi:10.1109/ICAIC60265.2024.10433829
- [7] G. P. Barrera Castro, A. Chiappe, M. S. Ramírez-Montoya, C. Alcántar Nieblas, Key Barriers to Personalized Learning in Times of Artificial Intelligence: A Literature Review, Appl. Sci., 15(6) (2025) 3103. doi:10.3390/app15063103



- [8] M. Somasundaram, K. A. Mohamed Junaid, S. Mangadu, Artificial Intelligence Enabled Intelligent Quality Management System (IQMS) for Personalized Learning Path, *Procedia Comput. Sci.*, 172 (2020) 438–442. doi:10.1016/j.procs.2020.05.096
- [9] D. Palko, et al., Cyber Security Risk Modeling in Distributed Information Systems, *Appl. Sci.*, 13(4) (2023) 2393. doi:10.3390/app13042393
- [10] M. Ismail, et al., Cybersecurity Activities for Education and Curriculum Design: A Survey, *Comput. Human Behav. Rep.*, 16 (2024) 100501. doi:10.1016/j.chbr.2024.100501
- [11] P. Skladannyi, et al., Development of Modular Neural Networks for Detecting Different Classes of Network Attacks, *Cybersecur.: Educ. Sci. Tech.*, 3(27) (2025) 534–548. doi:10.28925/2663-4023.2025.27.772
- [12] C. Merino-Campos, The Impact of Artificial Intelligence on Personalized Learning in Higher Education: A Systematic Review, *Trends Higher Educ.*, 4(2) (2025) 17. doi:10.3390/higheredu4020017
- [13] J. Batista, A. Mesquita, G. Carnaz, Generative AI and Higher Education: Trends, Challenges, and Future Directions from a Systematic Literature Review, *Information*, 15(11) (2024) 676. doi:10.3390/info15110676
- [14] V. Buriachok, et al., Implementation of Active Cybersecurity Education in Ukrainian Higher School, *Information Technology for Education, Science, and Technics*, vol. 178 (2023) 533–551. doi:10.1007/978-3-031-35467-0\_32
- [15] M. Astafieva, et al., Formation of High School Students' Resistance to Destructive Information Influences, in: *Cybersecurity Providing in Information and Telecommunication Systems*, vol. 3421 (2023) 87–96.
- [16] H. Hulak, et al. Formation of Requirements for the Electronic RecordBook in Guaranteed Information Systems of Distance Learning, in: *Workshop on Cybersecurity Providing in Information and Telecommunication Systems, CPITS 2021*, vol. 2923 (2021) 137–142.
- [17] Y. Kostiuk, et al., Protection of Information and Secure Data Exchange in Wireless Mobile Networks with Authentication and Key Exchange Protocols, *Cybersecur.: Educ. Sci. Tech.*, 1(25) (2024) 229–252. doi:10.28925/2663-4023.2024.25.229252
- [18] M. M. Ambali Parambil, et al., Integrating AI-based and Conventional Cybersecurity Measures into Online Higher Education Settings: Challenges, Opportunities, and Prospects, *Comput. Educ.: Artif. Intell.*, 7 (2024) 100327. doi:10.1016/j.caeai.2024.100327
- [19] P. Petriv, I. Opirskyy, N. Mazur, Modern Technologies of Decentralized Databases, Authentication, and Authorization Methods, in: *Cybersecurity Providing in Information and Telecommunication Systems II*, vol. 3826, 2024, 60–71.
- [20] Q. E. A. Ratul, E. Zaman, S. Das, M. Memari, The Role of AI in Transforming Education: A Systematic Review of Trends, in: *2025 Intermountain Engineering, Technology and Computing (IETC)*, 2025, 1–6. doi:10.1109/IETC64455.2025.11039329
- [21] W. Triplett, AI-Enhanced Cyber Science Education: Innovations and Impacts, *Information*, 16(9) (2025) 721. doi:10.3390/info16090721
- [22] J. Vykopal, P. Seda, V. Švábenský, P. Čeleda, Smart Environment for Adaptive Learning of Cybersecurity Skills, *IEEE Trans. Learn. Technol.*, 16(3) (2023) 443–456.
- [23] N. Kerimbayev, et al., Intelligent Educational Technologies in Individual Learning: A Systematic Literature Review, *Smart Learn. Environ.*, 12 (2025) 1. doi:10.1186/s40561-024-00360-3
- [24] Y. Smitiukh, et al., Development of a Prototype of an Intelligent System for Predicting the Quality of Dairy Manufacture, in: *2022 IEEE 11<sup>th</sup> Int. Conf. on Intelligent Systems (IS)*, 2022, 1–6. doi:10.1109/IS57118.2022.10019699
- [25] Y. Kostiuk, et al., A System for Assessing the Interdependencies of Information System Agents in Information Security Risk Management using Cognitive Maps, in: *Cyber Hygiene & Conflict Management in Global Inf. Networks*, vol. 3925, 2025, 249–264.

- [26] Y. Kostiuk, et al., The Methodology for Protecting Grid Environments from Malicious Code During the Execution of Computational Tasks, *Cybersecur.: Educ. Sci. Tech.*, 3(27) (2025) 22–40. doi:10.28925/2663-4023.2025.27.710
- [27] Y. Kostiuk, et al., Information and Intelligent Forecasting Systems based on the Methods of Neural Network Theory, in: *Smart Inf. Syst. Technol. (SIST)*, 2023, 168–173. doi:10.1109/SIST58284.2023.10223499
- [28] P. Skladannyi, Y. Kostiuk, N. Mazur, M. Pitaychuk, Study of Characteristics and Performance of Access Protocols to Cloud Computing Environments based on Universal Testing, *Telecommun. Inf. Technol.*, 1(86) (2025) 61–74.
- [29] Y. Kostiuk, et al., Integrated Protection Strategies and Adaptive Resource Distribution for Secure Video Streaming over a Bluetooth Network, in: *Cybersecurity Providing in Inf. Telecommun. Syst. II*, vol. 3826, 2024, 129–138.
- [30] Y. Kostiuk, et al., Models and Algorithms for Analyzing Information Risks During the Security Audit of Personal Data Information System, in: *Cyber Hygiene & Conflict Management in Global Inf. Networks*, vol. 3925, 2025, 155–171.
- [31] Y. Kostiuk, et al., Effectiveness of Information Security Control using Audit Logs, in: *Cybersecurity Providing in Inf. Telecommun. Syst. (CPITS 2025)*, vol. 3991, 2025, 524–538.
- [32] A. R. Babu, N. Arulanan, V. S. Chandran, Skill Development Through Experiential Learning: A Case Study for Product Development Scenario, *Procedia Comput. Sci.*, 172 (2020) 16–21. doi:10.1016/j.procs.2020.05.002
- [33] Y. Kostiuk, P. Skladannyi, V. Sokolov, M. Vorokhob, Models and Technologies of Cognitive Agents for Decision-Making with Integration of Artificial Intelligence, in: *Modern Data Science Technol. Doctoral Consortium (MoDaST 2025)*, vol. 4005, 2025, 82–96.
- [34] Y. Kostiuk, et al., Integrated Protection Strategies and Adaptive Resource Distribution for Secure Video Streaming over a Bluetooth Network, in: *Cybersecurity Providing in Inf. Telecommun. Syst.*, vol. 3826, 2024, 129–138.
- [35] M. Elkhodr, E. Gide, Integrating Generative AI in Cybersecurity Education: Case Study Insights on Pedagogical Strategies, Critical Thinking, and Responsible AI Use, *arXiv preprint*, 2025. doi:10.48550/arXiv.2502.15357
- [36] Y. Z. Rosunally, Harnessing Generative AI for Educational Gamification: A Framework and Practical Guide for Educators, in: *2024 21<sup>st</sup> Int. Conf. on Information Technology Based Higher Education and Training (ITHET)*, 2024, 1–8. doi:10.1109/ITHET61869.2024.10837655
- [37] S. Issa, et al., Knowledge Graph Completeness: A Systematic Literature Review, *IEEE Access*, 9 (2021) 31322–31339. doi:10.1109/ACCESS.2021.3056622
- [38] N. Chowdhury, S. Katsikas, V. Gkioulos, Modeling Effective Cybersecurity Training Frameworks: A Delphi Method-based Study, *Comput. Secur.*, 113 (2022) 102551. doi:10.1016/j.cose.2021.102551