Enhancing Cybersecurity Curriculum Development: AI-Driven Mapping and Optimization Techniques

Petr Dzurenda
dzurenda@vut.cz
Brno University of Technology
Brno, Czech Republic

Sara Ricci
ricci@vut.cz
Brno University of Technology
Brno, Czech Republic

Marek Sikora
marek.sikora@vut.cz
Brno University of Technology
Brno, Czech Republic

Michal Stejskal
xstejs34@vut.cz
Brno University of Technology
Brno, Czech Republic

Imre Lendak
lendak@uns.ac.rs
University of Novi Sad
Novi Sad, Serbia

Pedro Adão
pedro.adao@tecnico.ulisboa.pt
Instituto Superior Técnico, ULisboa,
and Instituto de Telecomunicações
Lisboa, Portugal

ABSTRACT

Cybersecurity has become important, especially during the last decade. The significant growth of information technologies, internet of things, and digitalization in general, increased the interest in cybersecurity professionals significantly. While the demand for cybersecurity professionals is high, there is a significant shortage of these professionals due to the very diverse landscape of knowledge and the complex curriculum accreditation process. In this article, we introduce a novel AI-driven mapping and optimization solution enabling cybersecurity curriculum development. Our solution leverages machine learning and integer linear programming optimization, offering an automated, intuitive, and user-friendly approach. It is designed to align with the European Cybersecurity Skills Framework (ECSF) released by the European Union Agency for Cybersecurity (ENISA) in 2022. Notably, our innovative mapping methodology enables the seamless adaptation of ECSF to existing curricula and addresses evolving industry needs and trend. We conduct a case study using the university curriculum from Brno University of Technology in the Czech Republic to showcase the efficacy of our approach. The results demonstrate the extent of curriculum coverage according to ECSF profiles and the optimization progress achieved through our methodology.

CCS CONCEPTS

• **Security and privacy** → Human and societal aspects of security and privacy; • **Human-centered computing** → **Human computer interaction (HCI)**; • **Social and professional topics** → *Computing education*; • **Applied computing** → **Education**.

KEYWORDS

Curricula Design, ECSF framework, Methodology, Cybersecurity Education

ARES 2024, July 30–August 02, 2024, Vienna, Austria
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1718-5/24/07
https://doi.org/10.1145/3664476.3670467

ACM Reference Format:
Petr Dzurenda, Sara Ricci, Marek Sikora, Michal Stejskal, Imre Lendak, and Pedro Adão. 2024. Enhancing Cybersecurity Curriculum Development: AI-Driven Mapping and Optimization Techniques. In *The 19th International Conference on Availability, Reliability and Security (ARES 2024), July 30–August 02, 2024, Vienna, Austria.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3664476.3670467

1 INTRODUCTION

Cybersecurity breaches pose a significant threat to governments, businesses, and individuals worldwide, particularly as digital technologies become increasingly integral to economies. The COVID-19 pandemic further accelerated society's reliance on digital platforms, with remote work becoming a fundamental for operational continuity during lockdowns. However, this shift exposed individuals and organizations to unprecedented cybersecurity risks (World Economic Forum, 2022 [15]). Despite the cybersecurity workforce reaching a record high of 5.5 million professionals, a global shortage of 3.9 million workers persists, with Europe facing a deficit of over 347,000 cybersecurity professionals (International Information System Security Certification Consortium (ISC2), 2023 [9]). Moreover, certain sectors, like governments and central banks, struggle to attract highly skilled professionals compared to industries such as finance (European Union Agency for Cybersecurity (ENISA), 2021 [11]). ENISA warned of a broader shortage of cybersecurity skills in the labor market, emphasizing the increasing demand for professionals with updated knowledge across various cybersecurity domains, including legal and policy frameworks (ENISA, 2021 [11]; ENISA, 2023 [5]). Effective cybersecurity training on all levels is essential to address both professional shortages and the general population's limited cybersecurity knowledge.

1.1 Contributions and Paper Organization

This paper contributes to advancing cybersecurity education by offering a systematic and data-driven approach to curriculum development that addresses evolving industry needs and trends. The main objective of this article is to enhance the development and availability of cybersecurity courses by unifying and synchronizing curriculum and training programs. This aims to provide the workforce with reliable and standardized options for upskilling and reskilling. To achieve this goal, we address the following Research Questions (RQ):

(RQ1) *How can ECSF profiles be made comprehensible to end users, and what practical methods exist for mapping academic courses to these profiles?* To address RQ1, we proposed a novel grouping methodology for ECSF key skills and knowledge. This grouping establishes a connection between NICE competencies (from which they are derived) and the ECSF framework, thereby assigning competencies to profiles and enhancing the readability and usability of the framework.

(RQ2) *How can a user-friendly AI-driven cybersecurity curriculum development tool be created, and which computational intelligence tools and method are suitable for this purpose?* To reflect RQ2, we proposed a novel cybersecurity curriculum development methodology based on machine learning and integer linear programming optimization mechanisms.

(RQ3) *How can such a tool be used by end users and what are its limitations?* To address RQ3, we provide implementation details for our CSProfiler online web tool, which incorporates our AI-driven cybersecurity curriculum development solution. Additionally, we conduct a case study using the existing university curriculum from Brno University of Technology in the Czech Republic to demonstrate the effectiveness of our approach.

The rest of the article is organized as follows. Section 2 reviews the key components of our AI-driven cybersecurity curriculum development solution. Section 3 presents our curriculum development methodology, addressing (RQ1). Section 4 provides the details about our CyberSecurity Profiler tool which implements the AI- and ILP-based solution and discusses (RQ2). Section 5 reports our experimental results. Section 6 sums up some open problems and potential extensions of our tool. Both sections reflect (RQ3). The final section contains the conclusions.

2 KEY COMPONENTS OF CYBERSECURITY SKILL DEVELOPMENT

In this section, we present key components of our Artificial Intelligence (AI)-driven cybersecurity curriculum development solution. Namely, we present the National Institute of Standards and Technology (NIST) Workforce Framework for Cybersecurity (NICE framework) and the European Cybersecurity Skills Framework (ECSF) as the basic building blocks of our solution. We introduce the REWIRE skills grouping enabling practical mapping of ECSF to actual curricula, professional training courses, and job market requirements. Then, we describe AI and Integer Linear Programming (ILP) methodologies that we use to process and evaluate cybersecurity curricula in our solution. The last section describes the Curricula Designer web application as our solution builds on this application.

2.1 National Initiative for Cybersecurity Education (NICE) framework

The National Initiative for Cybersecurity Education (NICE) in the United States, led by the National Institute of Standards and Technology (NIST), serves as a collaborative platform involving government, academia, and the private sector. NICE facilitates cybersecurity training and education by developing standards and best practices. The NICE Framework [14] delineates 1) work roles, 2) Knowledge, Skills, and Abilities (KSAs), and 3) tasks required for various cybersecurity roles.

The NICE Framework provides a comprehensive structure for understanding cybersecurity work roles, knowledge, skills, and tasks. With 7 categories and 33 Specialty Areas, the framework delineates the diverse aspects of cybersecurity work. The defined 52 Work Roles are the most detailed groupings of cybersecurity and related work, which include KSAs and tasks performed in each role. In addition, a Competencies classification is derived from the NICE Cybersecurity Skills Framework [19]. These competencies are of particular interest for our analyses since they allow "education and training providers to be responsive to employer or sector needs" as mentioned in [19].

2.2 European Cybersecurity Skills Framework

The European Cybersecurity Skills Framework (ECSF) [4] is the result of a collaborative effort between ENISA and the ENISA ad-hoc working group on cybersecurity skills framework. The ECSF is designed to establish a standardized understanding of roles, competencies, skills, and knowledge in the field of cybersecurity, facilitating skills recognition and aiding in the development of cybersecurity training programs. This framework categorizes cybersecurity roles into 12 profiles, each delineating key tasks, skills, knowledge, and competencies. Notably, the ECSF encompasses a total of 84 key skills and 69 key knowledge areas, providing a comprehensive taxonomy for academia and guiding the content of cybersecurity-related education. The 12 ECSF profiles are: Chief Information Security Officer (CISO), Cyber Incident Responder, Cyber Legal, Policy & Compliance Officer, Cyber Threat Intelligence Specialist, Cybersecurity Architect, Cybersecurity Auditor, Cybersecurity Educator, Cybersecurity Implementer, Cybersecurity Researcher, Cybersecurity Risk Manager, Digital Forensics Investigator, and Penetration Tester.

2.3 REWIRE Mission and Skills Grouping

The REWIRE project [16] is a European initiative that aims to address the shortage of cybersecurity experts by improving the availability, accessibility, and quality of cybersecurity courses and certifications. To achieve this goal, the project incorporates the ECSF framework as a key reference point in its actions. Upon analyzing the ENISA framework, it becomes apparent that the key skills and knowledge outlined to describe the profiles are expressed in unique ways, posing challenges in depicting interconnections among the profiles. To address this issue, a method was developed to group skills and knowledge representing similar concepts but expressed differently. This approach enables the clustering of skills and knowledge based on their descriptions, resulting in the identification of 31 distinct groups. This grouping not only simplifies the readability and usability of the profiles but also enhances their comprehensibility, particularly for individuals lacking technical knowledge in the field. We refer to Section 3.2 for more details.

2.4 Artificial Intelligence

Machine Learning (ML) is increasingly used for tasks like image analysis and time-series forecasting [1, 2]. Recurrent Neural Networks (RNN) are popular for tasks such as time series analysis [18] and Natural Language Processing (NLP) [20]. RNNs compress input data into a context vector. They exhibit strong performance in sentiment analysis, accurately categorizing sentence sentiment as positive or negative. This methodology can also be effectively applied to classify required skills in course descriptions. Our model uses a pre-built word-piece tokeniser for Bidirectional Transformer (BERT) [1]. With the tokens from the word-piece tokeniser, a model is used to concatenate the mean of all intermediate context vectors and the last context vector. The concatenated vector is fed to an Artificial Neural Network (ANN) to classify security skills. The hidden unit of Long Short-Term Memory (LSTM) is 128, and the hidden layer of ANN is 256. The model training involved testing its ability to predict required security skills using a labeled dataset, with the Long Short-Term Memory (LSTM) model trained solely on the training dataset to prevent bias.

The ML learning approach offers several advantages. It reduces the labor and time required for textual analysis, allows for nuanced analysis of sentiments, and increases flexibility by enabling easy addition of new data to the dataset. Compared to survey-based approaches, accessing publicly available course description allows for obtaining more data with less difficulty.

2.5 Integer Linear Programming (ILP)

To determine the optimal combinations of curriculum courses to achieve the desired ECSF profile, it is essential to first identify a suitable search algorithm. This algorithm should be capable of identifying the best combination of the courses based on user-defined input conditions, such as the number of years of the study program, the maximum number of courses per semester, the total number of European Credit Transfer System (ECTS) credits, and the desired ECSF profiles to be covered by the curriculum. In addition, the algorithm should maintain a balance between the distribution of ECTS credits load between individual study years and semesters. These restrictions necessitate solving an optimization problem. One potential approach to addressing this challenge is to use Integer Linear Programming (ILP) [13], specifically 0-1 Integer Linear Programming, as the decision variables (i.e., whether to select a course or not) are binary in nature. Although this problem is known to be Nondeterministic Polynomial (NP)-hard, existing solvers demonstrated efficient handling of problems with thousands of decision variables, which is well beyond the scope of our application.

2.6 Curricula Designer

The SPARTA project established a robust community dedicated to collaboratively define, develop, and disseminate solutions aimed at strengthening cybersecurity measures. As part of the SPARTA initiative, a novel free web application named Cybersecurity Curricula Designer [7] was developed. This tool enables the creation of cybersecurity curricula tailored to meet specific job requirements.

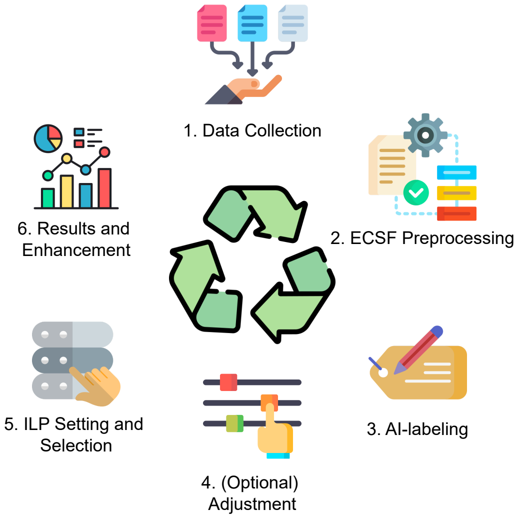


Figure 1: Curriculum development methodology.

Leveraging the SPARTA framework, which is derived from the National Institute of Standards and Technology (NIST) Workforce Framework for Cybersecurity (NICE framework), the application employs an insertion mechanism to analyze and align generated curricula with the NIST NICE work roles. Furthermore, the application was expanded to incorporate the key skills and knowledge outlined in the ENISA framework [8].

3 CURRICULUM DEVELOPMENT METHODOLOGY

We propose a novel AI-driven mapping and optimization techniques for cybersecurity curriculum development, establishing a connection between academy and industry. Syllabi serve as detailed documents outlining course information, structure, objectives, learning outcomes, and content, providing a comprehensive framework for understanding the skills and knowledge students are expected to acquire. In our approach, we analyze courses using an RNN algorithm trained on job ads, which represent market demand and contain skills groups from ECSF profiles. This analysis correlates the skills and knowledge described in the ECSF profiles with course contents gathered from the syllabi. Subsequently, we apply an ILP algorithm to select courses that match the required skills. These techniques facilitate the alignment of academic offerings with industry needs, ensuring that cybersecurity curricula are tailored to address current demands and trends in the field.

The proposed methodology of AI-driven curriculum development is depicted in Figure 1. This methodology involves six key steps: data collection from various sources such as university websites and course catalogs; definition of an efficient mapping methodology to link identified courses with ECSF profiles; AI-labelling of courses for alignment with ECSF profiles; human adjustment of AI output if needed; selection of the most suitable combination of processed courses based on user-defined conditions; and presentation of the results on curriculum development/enhancement. These steps allow for the construction of a newly developed cybersecurity curriculum by combining available academic courses or adapting existing curricula to align with desired ECSF profiles.

3.1 Data Collection

The data used for curriculum development and analysis were extracted from existing course syllabi. A syllabus can be set out by an examination board or tailored by the instructor who teaches or controls the course. These syllabi serve as comprehensive guides, providing essential information for curriculum development. They include details on course information, structure, purpose, objectives, learning outcomes, and content, offering a comprehensive overview of the skills and knowledge expected to be acquired by students. Specifically, a course syllabus includes:

- **Course Information**: provides basic details such as the course title, code, type, level, and the academic year/semester it is offered.
- **Instructor Information**: where the instructor's names are listed.
- **Course Structure**: outlines the workload distribution, including ECTS credits, weekly lectures and hours, as well as laboratory sessions, if applicable.
- **Course Purpose and Objectives**: are stated to clarify the aim and expected outcomes for students.
- **Learning Outcomes**: describe the knowledge, skills, and competencies students are expected to gain upon completing the course successfully.
- **Course Prerequisites and Co-requisites**: required for enrollment in the course.
- **Course Content**: provides a detailed breakdown of the topics and subjects covered throughout the duration of the course.
- **Teaching Methodology**: describes the approach used for delivering the course content, such as face-to-face instruction or online delivery.
- **Bibliography**: lists required readings, textbooks, and additional resources.
- **Assessment**: includes the weighting of examinations, assignments, and class participation.
- **Language**: is specified to ensure students are aware of the medium of communication used in lectures and materials.

Language, course information, and structure sections provide fundamental details used for filtering courses, such as ECTS credits and semester allocation. Conversely, the content of sections on course purpose and objectives, learning outcomes, and course content were collected and fed into the AI-labeling process. In fact, these sections detail the specific skills and knowledge that students are expected to acquire.

3.2 ECSF Preprocessing

The ECSF describes 12 profiles with a total of 84 skills and 69 knowledge, often expressed in unique phrasing. This presents a challenge in illustrating the relationships among profiles through the connections of identical skills and knowledge. One approach to address this challenge is to group together skills and knowledge that convey the same concept but are phrased differently. Consequently, skills and knowledge can be clustered based on their descriptions. For instance, Table 1 shows one such cluster along with the corresponding ECSF key skills and knowledge. Additionally, it displays the profile associated with each skill or knowledge.

Table 1: "Software Development" skill group with ECSF key skills and knowledge linked to the associated ECSF profile.

ECSF Skills and Knowledge	ECSF Profile
Skill. Configure solutions according to the organisation's security policy	Cybersecurity Implementer
Skill. Develop codes, scripts and programmes	Penetration Tester
Knowledge. Review codes assess their security	Penetration Tester
Knowledge. Secure development lifecycle	Cybersecurity Architect
Knowledge. Computer programming	Cyber Threat Intelligence Specialist, Cybersecurity Implementer, Penetration Tester
Knowledge. Secure coding recommendations and best practices	Cybersecurity Implementer

It is noteworthy that when we initiated our analysis of skills needs in cybersecurity, the ECSF had not yet been established as a European classification of skills. Consequently, we relied on the NICE NIST competencies framework as our starting point due to its comprehensive structure and alignment with other existing frameworks [6]. The following steps were undertaken:

(1) Some of the NICE competencies required adjustments to better suit the European (EU) market, were irrelevant for our analysis, or could be merged. Hence, REWIRE cybersecurity experts proposed a total of 29 REWIRE skills groups derived from the NICE competencies, which were then validated through a survey involving cybersecurity stakeholders. We refer to REWIRE Deliverable R2.2.2 [10] for more details.
(2) The 29 REWIRE skills were used as the foundation for grouping the ECSF key skills and knowledge, with strict adherence to the descriptions provided by the REWIRE skills.
(3) Three REWIRE skills were renamed to more accurately reflect the groups and their newly defined scope based on ECSF skills and knowledge. Specifically, "Communication" was modified to "Collaborate and Communicate," "Enterprise Architecture" became "Enterprise Architecture and Infrastructure Design," and "Information Technology Assessment" was updated to "Information Security Controls Assessment."
(4) Two new skills were identified as missing: "Problem Solving & Critical Thinking" and "Technology Fluency." The definition of these new groups was informed by the descriptions provided in the NIST NICE competencies.
(5) Following the creation of the initial draft of groups, REWIRE experts collaborated to ensure consistency in the grouping process.

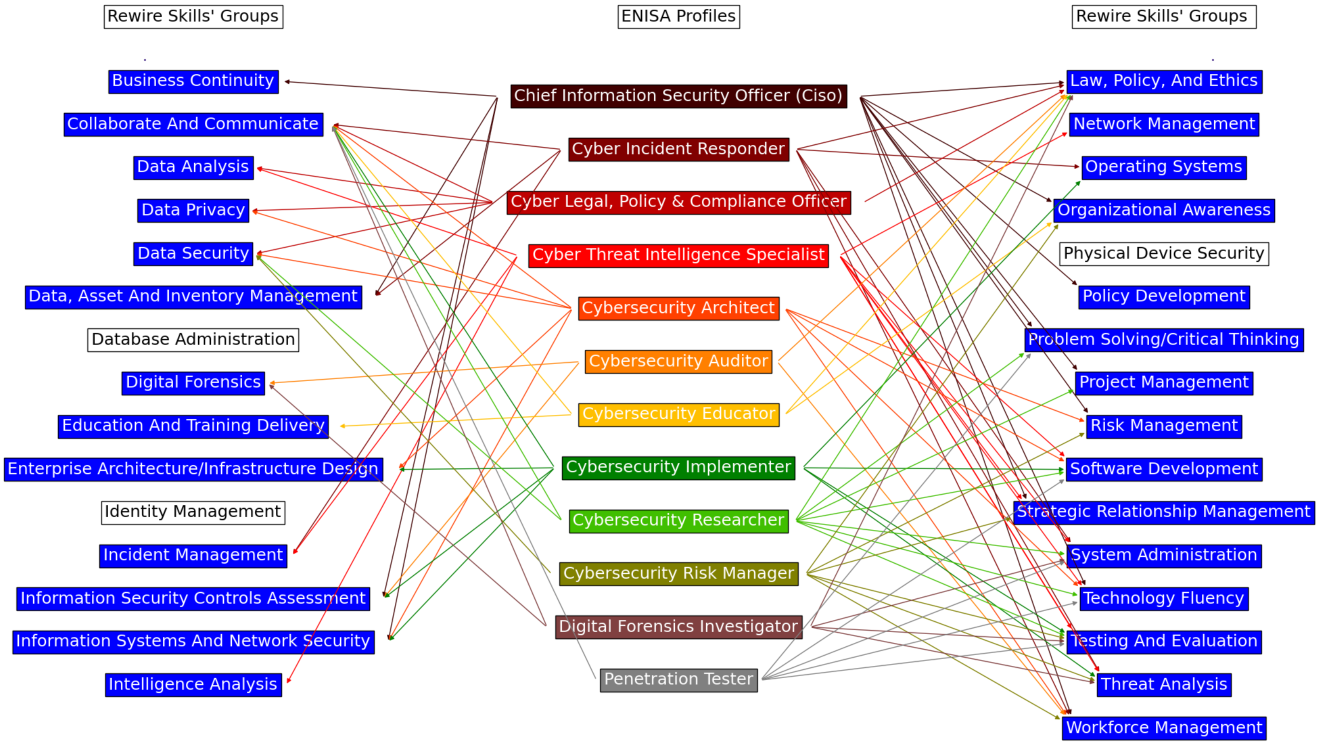


Figure 2: Mapping REWIRE Skills Groups to ECSF profiles.

Figure 2 illustrates the correlation between REWIRE skills groups and ECSF profiles. It is noteworthy that the proposed grouping establishes a connection between NICE competencies and the ENISA framework, thereby assigning competencies to profiles and enhancing the readability and usability of the profiles.

Furthermore, the grouping strategy for key skills and knowledge identified a discrepancy in the ECSF framework, which may warrant future improvements. Specifically, several REWIRE skill groups lack assigned skills, knowledge, or both. These groups were initially identified independently from the ECSF framework and subsequently expanded with associated key skills and knowledge from the framework. This approach strengthened the definition of cybersecurity skills and identified skills overlooked in both REWIRE groups and the ECSF framework. Potential improvements include adding missing descriptions to existing profiles and considering the inclusion of any missing profiles in the ECSF framework. For further details, please refer to REWIRE Deliverable R3.4.1 [3].

3.3 AI-labeling

Our application utilizes a RNN to analyze the text inserted into the system, specifically targeting the extraction of relevant skills. RNNs are particularly adept at processing sequential data, making them well-suited for tasks such as NLP. In our approach, sentences are interpreted as sequences of words, which are then converted into integer representations and fed into the RNN model. This methodology, commonly used for tasks like sentiment analysis, is adapted to classify the required skills in the given text.

Before the RNN can effectively analyze text and extract skills, it needs to be trained on a dataset of labeled examples. The training consisted of testing the model's ability to predict the required security skills using a test dataset. The ground truths were manually labeled. Furthermore, the LSTM model was only trained on the training dataset to prevent the model from learning the test dataset and thereby biasing the training results. The training dataset comprised 937 cybersecurity job advertisement descriptions. Note that employing the same algorithm in job ads and course description datasets ensure:

- **Consistency in Skill Identification**: Employing the same algorithm ensures consistency in identifying cybersecurity-related skills across both datasets, enabling a unified approach to skill extraction and analysis.
- **Cross-Domain Skill Mapping**: By applying the algorithm to both job ads and course descriptions, it becomes possible to map the skills demanded in the job market to the skills taught in academic courses. This facilitates alignment between industry needs and educational offerings.
- **Feedback Loop for Curriculum Enhancement**: Analysis of both job ads and course descriptions can establish

a feedback loop for curriculum enhancement, allowing institutions to adapt their programs in response to evolving industry requirements and technological advancements.

After training the RNN, each course description collected in Section 3.1 was input into the RNN algorithm, which then automatically generated the list of identified skill groups.

3.4 (Optional) Adjustment

The next step involves an optional manual review of the assigned skill groups for each course. This process requires a cybersecurity expert to examine the course descriptions provided as input to the RNN algorithm and verify the accuracy of the assigned skill groups. The manual review of the assigned skill groups serves as valuable feedback to the RNN results, allowing cybersecurity experts to validate and refine the accuracy of the skill group assignments.

3.5 ILP Setting and Selection

The objective of this section is to give a detailed description of the ILP algorithm providing course selection, and therefore, curriculum design itself. The algorithm was configured to match the courses to chosen ECSF profiles so that all skill groups included in the profile are represented by these courses. As we theorised that a single course would not be able to match all skills and knowledge, multiple courses were planned to be selected. The found combination of courses is optimal if it matches the profile and the selected parameters set up by the user. It is also essential to get an optimal solution in the shortest possible time. The filtering applied by the user includes important numerical values such as the number of study years, number of ECTS credits, and included skills groups of the courses. Furthermore, the algorithm maintains a balance between the distribution of ECTS credit load between individual study years and semesters. These values, together with the maximum number of courses per semester, are subjected to optimization to find the best possible result. The ILP optimization method helps us to solve this task. The process of optimization is following:

(1) First, we remove REWIRE skills groups from the selected ECSF profile that we are unable to cover with our set of courses. The subset of the skills groups is denoted $\mathcal{R}$.

(2) Second, we run the ILP algorithm. The algorithm takes as input: 1) the subset of skills groups $\mathcal{R}$, 2) number of ECTS credits C and its tolerance, 3) a number of study years Y, 4) the set of available courses M with the associated skills groups they cover, and 5) the number of maximum courses per semester $\mathcal{NC}$.

(3) The ILP algorithm settings is defined by Equation 1. The equation says that we are looking for the smallest subset of courses b_i that meet the given conditions.

$$F_{min} = \sum_{i=0}^{M} b_i, \tag{1}$$

where $F_{\min}$ is a minimization of the objective function where b_i represents a binary variable deciding whether the course will be included or not.

(4) The following equations describe the constraints defining the range of tolerance of the number of ECTS credits in the entire study program. The $C_{\min}$ and $C_{\max}$ values defined in equation (2) refer to the credits range in which the sum of the courses should move.

$$\sum_{i=0}^{M}(c_i b_i) \geq C_{min} \qquad \sum_{i=0}^{M}(c_i b_i) \leq C_{max}, \tag{2}$$

where c_i defines number of credits per course b_i, $C_{\min}$ is the lower bound, and $C_{\max}$ is the upper bound.

(5) The condition defined by Equation 3 ensures an even distribution of courses between semesters. Namely, $S_{\min}$ and $S_{\max}$ define the minimum and maximum difference between the distribution of ECTS credits between semesters. The C bounds are set by the user input, while S bounds are set by default to $S_{\min} = -1$ and $S_{\max} = 1$.

$$\sum_{i=0}^{M}(s_i b_i) \geq S_{min} \qquad \sum_{i=0}^{M}(s_i b_i) \leq S_{max}, \tag{3}$$

where s_i defines semester type ($s_i = 1$ for summer semester and $s_i = -1$ for winter semester).

(6) The condition defined by Equation 4 guarantees that on average there are $\mathcal{NC}$ per semester.

$$\sum_{i=0}^{M} b_i \geq \mathcal{NC} \cdot 2Y, \tag{4}$$

where $2Y$ defines the total number of semesters in the whole study program.

(7) The last constraint defined by Equation 5 searches for courses that include the skills groups from the subset $\mathcal{R}$. In the equation, x_j stands for skills groups in $\mathcal{R}$ and n is a number of skills group in $\mathcal{R}$. The skills group x_j takes the value 1 if it is covered by the course b_i, and 0 otherwise.

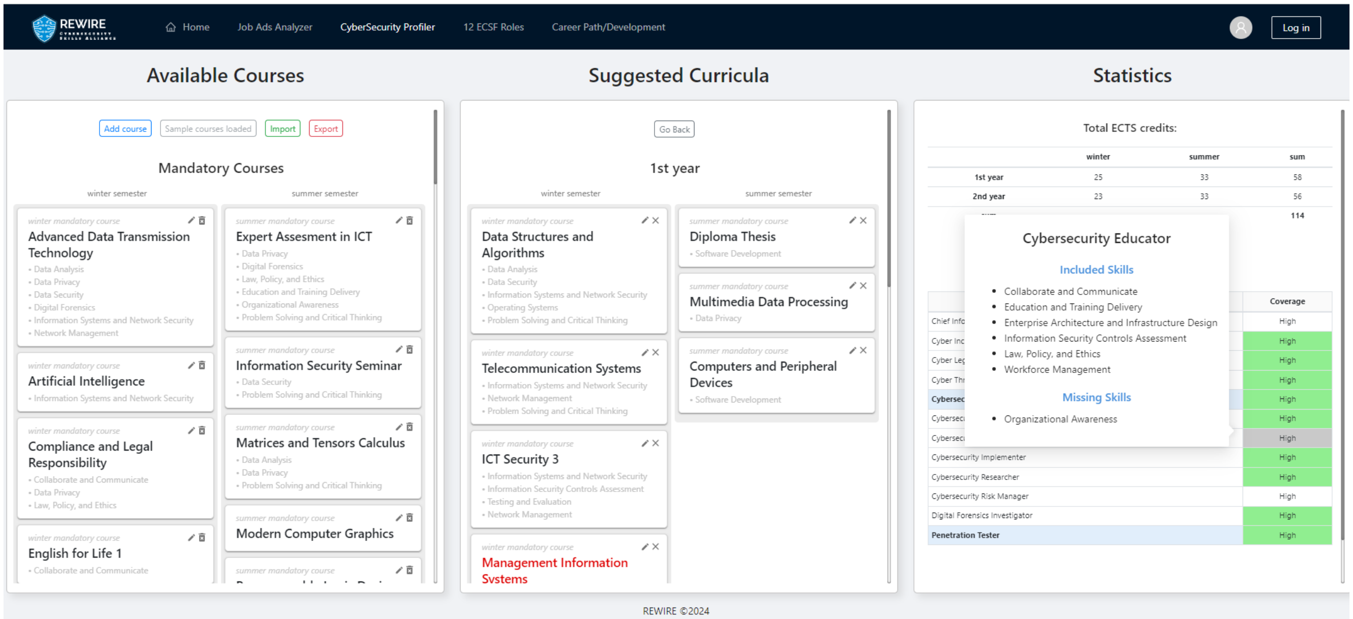
$$\sum_{i=0}^{M}(x_j\, b_i) \geq 1 \text{ for } j = 1, \ldots, n \tag{5}$$

(8) The algorithm is run twice. In the first run, an optimal solution is found. This solution aims to fulfill all parameters like credits, semesters, and matching the profile. All courses in the solution are given priority two. The algorithm is then executed again just with the courses already included in the solution and the output is only necessary courses to fulfill the profile criteria and also the semester balance is maintained. These courses are then given priority one.

(9) In the final phase, the selected courses in the winter and summer semesters are evenly distributed into semesters. This spreads the selected courses over the entire length of the study program and maintain an even degree of difficulty between semesters. Note that we do not consider how deep each course covers each skill group.

3.6 Results and Enhancement

The final step of the curriculum development presents the evaluation of the designed curriculum. The user gets visual information on key building courses of the curriculum which are essential to cover selected ECSF profiles, the number and distribution of ECTS credits across semesters, the level of coverage of all ECSF profiles, and information on missing skills groups to cover the required ECSF profile. The user is also given the option to replace non-key courses of the curriculum with other available courses, and thereby achieve an even higher adaptation of the curriculum to his/her own needs



Available Courses

Mandatory Courses

winter semester

winter mandatory course
Advanced Data Transmission Technology
- Data Analysis
- Data Privacy
- Data Security
- Digital Forensics
- Information Systems and Network Security
- Network Management

winter mandatory course
Artificial Intelligence
- Information Systems and Network Security

winter mandatory course
Compliance and Legal Responsibility
- Collaborate and Communicate
- Data Privacy
- Law, Policy, and Ethics

winter mandatory course
English for Life 1
- Collaborate and Communicate

summer semester

summer mandatory course
Expert Assesment in ICT
- Data Privacy
- Digital Forensics
- Law, Policy, and Ethics
- Education and Training Delivery
- Organizational Awareness
- Problem Solving and Critical Thinking

summer mandatory course
Information Security Seminar
- Data Security
- Problem Solving and Critical Thinking

summer mandatory course
Matrices and Tensors Calculus
- Data Analysis
- Data Privacy
- Problem Solving and Critical Thinking

summer mandatory course
Modern Computer Graphics

summer mandatory course

Suggested Curricula

1st year

winter semester

winter mandatory course
Data Structures and Algorithms
- Data Analysis
- Data Security
- Information Systems and Network Security
- Operating Systems
- Problem Solving and Critical Thinking

winter mandatory course
Telecommunication Systems
- Information Systems and Network Security
- Network Management
- Problem Solving and Critical Thinking

winter mandatory course
ICT Security 3
- Information Systems and Network Security
- Information Security Controls Assessment
- Testing and Evaluation
- Network Management

winter mandatory course
Management Information Systems

summer semester

summer mandatory course
Diploma Thesis
- Software Development

summer mandatory course
Multimedia Data Processing
- Data Privacy

summer mandatory course
Computers and Peripheral Devices
- Software Development

Statistics

Total ECTS credits:

	winter	summer	sum
1st year	25	33	58
2nd year	23	33	56
			114

	Coverage
Chief Info	High
Cyber Inc	High
Cyber Leg	High
Cyber Thr	High
Cybersec	High
Cybersec	High
Cybersec	High
Cybersecurity Implementer	High
Cybersecurity Researcher	High
Cybersecurity Risk Manager	High
Digital Forensics Investigator	High
Penetration Tester	High

Cybersecurity Educator

Included Skills

- Collaborate and Communicate
- Education and Training Delivery
- Enterprise Architecture and Infrastructure Design
- Information Security Controls Assessment
- Law, Policy, and Ethics
- Workforce Management

Missing Skills

- Organizational Awareness

Figure 3: Graphical user interface of the CSProfiler tool.

e.g., by adding a higher importance to preferable skills groups or expanding the scope of the curriculum to cover more ECSF profiles.

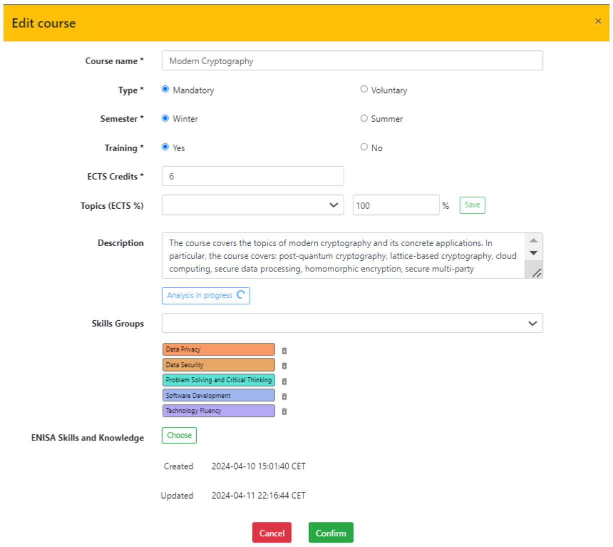
4 CYBERSECURITY PROFILER: PURPOSE, TARGETED USERS, AND USABILITY

The REWIRE CyberSecurity Profiler (CSProfiler) is an open-source, freely available, dynamic web application that serves as a comprehensive tool for designing and analyzing curricula, professional training, and certification schemes in the cybersecurity domain. Building upon the foundation of the SPARTA Curricula Designer, it extends its capabilities by allowing dynamic analysis of courses and their optimized selection for the creation of specific cybersecurity curricula. This comprehensive tool enables users to design their curricula, certifications, or training programs while accessing real-time statistics in compliance with ENISA ECSF and NICE frameworks. Specifically, the CSProfiler maps existing curricula, trainings, and certifications to specific cybersecurity roles defined by ENISA profiles, identifies recommended courses for these roles, and allows users to design study programs accordingly. Through its integration with the 31 REWIRE group, users can easily navigate and maintain a database of existing cybersecurity educational resources.

4.1 Using the Tool

The CSProfiler tool fully integrates the AI-driven cybersecurity curriculum development concept presented in the previous sections. The tool is divided into three main sections, see Figure 3: 1) the left section (Available Courses) allows users to define new courses or modify the existing ones, 2) the middle section (Suggested Curricula) allows the composition of a study program from defined courses with help of AI, and 3) the right section (Statistics) provides the statistical data and compliance of the designed program with the requirements.

When users add courses, they are required to input various data, including the course name, type, semester, and number of ECTS credits, along with a text describing the course aims, objectives, and outcomes. This is where the AI algorithm plays a crucial role. By analyzing the provided course description, the algorithm determines the skill groups covered by the course as shown in Figure 4.



Course name * Modern Cryptography

Type * ● Mandatory ○ Voluntary

Semester * ● Winter ○ Summer

Training * ● Yes ○ No

ECTS Credits * 6

Topics (ECTS %) 100 %

Description: The course covers the topics of modern cryptography and its concrete applications. In particular, the course covers: post-quantum cryptography, lattice-based cryptography, cloud computing, secure data processing, homomorphic encryption, secure multi-party

Skills Groups
- Data Privacy
- Data Security
- Problem Solving and Critical Thinking
- Software Development
- Technology Fluency

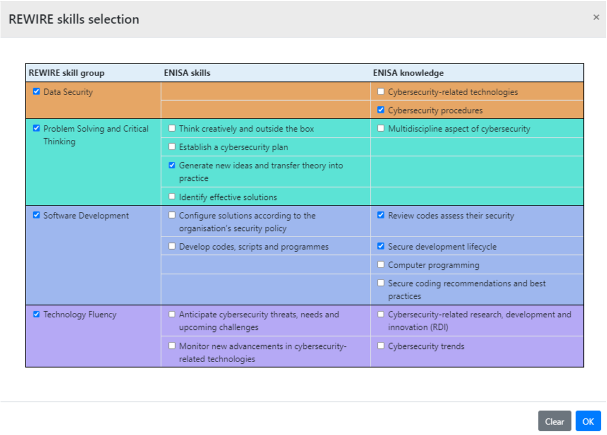
ENISA Skills and Knowledge

Created 2024-04-10 15:01:40 CET

Updated 2024-04-11 22:16:44 CET

Figure 4: Course edit menu in the CSProfiler tool.

If users do not agree with the AI recommendation, they can easily remove or add skills groups or even more, they can use more detailed classification of their courses by using the specific ECSF skills and knowledge as depicted in Figure 5.

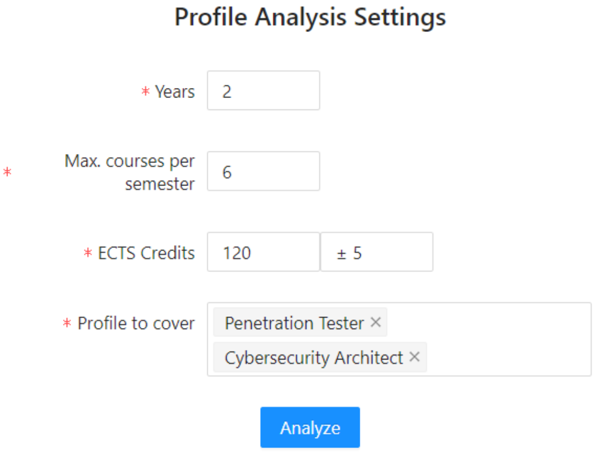


REWIRE skills selection ×

REWIRE skill group	ENISA skills	ENISA knowledge
☑ Data Security		☐ Cybersecurity-related technologies
		☑ Cybersecurity procedures
☑ Problem Solving and Critical Thinking	☐ Think creatively and outside the box	☐ Multidiscipline aspect of cybersecurity
	☐ Establish a cybersecurity plan	
	☑ Generate new ideas and transfer theory into practice	
	☐ Identify effective solutions	
☑ Software Development	☐ Configure solutions according to the organisation's security policy	☑ Review codes assess their security
	☐ Develop codes, scripts and programmes	☑ Secure development lifecycle
		☐ Computer programming
		☐ Secure coding recommendations and best practices
☑ Technology Fluency	☐ Anticipate cybersecurity threats, needs and upcoming challenges	☐ Cybersecurity-related research, development and innovation (RDI)
	☐ Monitor new advancements in cybersecurity-related technologies	☐ Cybersecurity trends

Figure 5: Course classification with skills groups.

When a user creates all needed courses, then he/she can run the automatized curriculum development driven by the CSProfiler. To do so, the user has to fill in his/her requirements on the curriculum and run the analysis process in the corresponding CSProfiler section as depicted in Figure 6.



Profile Analysis Settings

* Years	2
* Max. courses per semester	6
* ECTS Credits	120 ± 5
* Profile to cover	Penetration Tester × Cybersecurity Architect ×

Figure 6: Curricula setting Menu.

The result of the analysis is the selection of the most optimal combination of courses from the set of available courses covering the selected ECSF profiles and their division into individual semesters (Figure 3, middle). The tool highlights (red text font of the courses titles) those courses that are key to the selected ECSF profiles. Notably, key courses are a minimal combination of courses with needed skills for the selected profile. The other courses have a supplementary character and deepen knowledge in the already covered ECSF profiles with regard to the maximum number of credits for the entire study. In the right column of Figure 3 (Statistics), the tool displays the statistics of the created curriculum, i.e., the total number of ECTS credits, covered ECSF profiles, and missing and covered skills' groups. If necessary, users can move courses between the left column (Available Courses) and the middle column (Suggested Curricula) and thus optimize the proposed curriculum according to their needs.

4.2 Technical Implementation

As shown in Figure 7, the implementation of the web application is divided into two parts: 1) client and 2) server with database. The client is a Javascript application that is downloaded and launched in the user's web browser after the first loading of the website. This application contains a Graphical User Interface (GUI) and mediates all interaction with the user in real time. In the case of working with database data or processing user data, the application connects to the server, which mediates communication with the database and provides computational analysis tools.

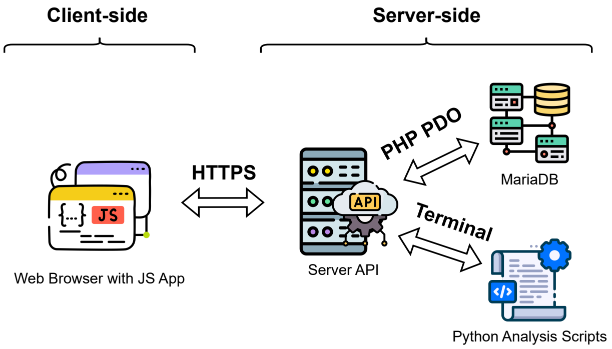


Figure 7: Application structure.

4.2.1 Client-side. This part represents the front-end of the entire system. It is a dynamic Javascript web application based on the React framework. The code is executed directly in the user's web browser, allowing for immediate feedback to the user. The following programs and libraries were also used:

- Antd – Graphical design and components,
- React Bootstrap – Graphical components, and structure,
- Apexcharts – Graphs and plots,
- Crypto – Cryptography features,
- React Beautiful DnD – Drag'n'Drop mechanism,
- Node-sass – Style preprocessor,
- Axios – Connection to the data server,
- NPM - Node Packet Manager.

The CSProfiler application is part of a larger web application called the CyberAbility Platform[1] which is open-source and available on the public GitLab repository[2]. The CSProfiler is implemented as a React component that inherits several variables and callback functions from the parent application within component

[1] https://cyberability-platform.informacni-bezpecnost.cz/

[2] https://gitlab.com/brno-axe/rewire/cyberability_platform

attributes called *props*. These mainly concern the management of user accounts and data, connection to the server, various dialog windows and functions for displaying content. CSProfiler itself also contains its own set of variables called *state*, in which it stores data related to its own content.

4.2.2 Server-side. On the server side, there is a server Application Programming Interface (API) created in Hypertext PreProcessor (PHP). This API accepts requests from the client application and returns confirmation of the performed action or requested data. Individual actions are distinguished based on the requested URL and depending on the received data from the client application. The data is stored in the MariaDB SQL server on the same server machine. Access to the database is done through the PHP API. In order to keep the data safe from unauthorized manipulation through the web application, an authentication and authorization mechanism using user tokens is also implemented in the API.

The server also offers analysis tools that allow users to analyze their data. These tools are implemented as Python scripts and executed on the same machine where the API server and the database are running. These analysis tools are more detailed in the next subsection.

4.2.3 Server Analysis Tools. The server includes two Python scripts to perform 2 different analysis. These Python scripts are run using the input provided by the client, and upon successful termination, the result of the analysis is processed by the PHP API and sent to the client. The first analysis tool is for skill prediction using a machine learning algorithm. This tool consists of a tokenizer and an Long short-term memory (LSTM) model. The entire tool is implemented in Python 3 and also allows retraining of the model. This analysis tool was introduced and described in [17]. The second analysis tool is used to design an optimal curriculum based on the required professional profile from available study courses. This analysis is based on the integer linear programming algorithm described in Section 3.

5 CASE STUDY

The proposed curriculum development methodology was applied to the "Information Security" master's study program [12] offered at Brno University of Technology in the Czech Republic. This program comprises a total of 60 courses, consisting of 15 compulsory, 17 compulsory-elective, and 28 elective courses. Compulsory courses provide fundamental knowledge, compulsory-elective courses guide students towards specialization in desired work roles, and elective courses offer opportunities to broaden students' general knowledge. The master's program spans two years, during which students are required to earn a minimum of 120 ECTS credits from a total of 271 credits offered.

The "Information Security" master's program offers an interdisciplinary education, encompassing technical fields such as applied mathematics, cryptography, informatics, and humanities, with a special emphasis on law. This interdisciplinary approach provides students with a comprehensive understanding of cybersecurity from various perspectives, making it an ideal candidate for applying our curriculum development methodology. With the program's emphasis on cybersecurity, we aim to explore how students can tailor their studies to align with specific ECSF profiles, thereby enhancing their readiness for roles in the cybersecurity sector.

Table 2: Adjustment results of the AI-labeling.

♯ Courses	Added Skills	♯ Courses	Deleted Skills
23	0	37	<5
23	1	2	5
8	2	21	>5
6	3 or 4		

The data collection phase involved downloading the syllabi of each master's course and extracting the content from sections such as *course description*, *aims*, and *syllabus* of lectures and laboratories. Subsequently, each course underwent AI-labeling followed by human adjustment. The AI-labeling process was executed by simply clicking a button in the CSProfiler, while manual adjustments were carried out by REWIRE cybersecurity experts.

To track the changes made during the adjustment process, an Excel file was created to monitor the addition and removal of skills identified through AI-labeling. Table 2 provides an overview of the number of courses that had specific amounts of skills added or removed. Notably, the RNN algorithm demonstrated good performance, with only a small number of the expected skill groups missing from the courses. Specifically, only 4 courses required the addition of 3 skills each, and only 2 courses required the addition of 4 skills each. However, several courses necessitated the deletion of incorrectly assigned skill groups. It is noteworthy that courses in law and languages exhibited a higher frequency of skill group deletions. This can be attributed to text descriptions that, while mentioning technical topics, do not deepen into details.

After inputting the first-year compulsory courses into the CSSProfiler, it became apparent that the curriculum already covers the requirements for the Cyber Legal, Policy & Compliance Officer profile. This profile necessitates proficiency in four skill groups: "Collaborate and Communicate," "Data Analysis," "Data Privacy," and "Law, Policy, and Ethics," all of which are adequately addressed by existing language, data processing, and law courses. Additionally, the master's program offers a foundational coverage of essential skill groups for various other profiles, including Cyber Incident Responder, Cybersecurity Architect, Cybersecurity Auditor, Cybersecurity Educator, Cybersecurity Implementer, Cybersecurity Researcher, Cybersecurity Risk Manager, and Digital Forensics Investigator.

Furthermore, the ILP algorithm can assist in structuring a curriculum tailored to specific profiles, such as the Cyber Incident Responder. For this purpose, we configured the ILP with the following parameters: a two-year academic timeline, a maximum of six courses per semester, an ECTS range of 120 ± 5, and a focus on the Cyber Incident Responder profile. A comparison between the results obtained from the ILP selection and those following the compulsory/elective credit requirements reveals an overlap of 8 out of 19 courses. This discrepancy arises from the ILP's ability to select courses based on the presence of one or more skill groups required for the profile, without consideration of compulsory and elective course constraints. The ILP analysis results can be used to recommend adjustments to the curriculum structure, such as making certain courses compulsory to ensure comprehensive coverage of essential skill groups for specific profiles.

6 DISCUSSION

The CSProfiler and its AI-driven cybersecurity curriculum development approach empower users with a user-friendly and intuitive tool for creating their cybersecurity curricula. Our results show that the results of AI are usable and further research and development can enhance the accuracy and usability of the tool.

For example, the high quality of artificial intelligence course classification requires having a high-quality training dataset containing courses labeled with corresponding skills groups. Furthermore, it is necessary to have a training dataset of at least 1000 records. Therefore, our further research will focus on improving the quality of the dataset, including the possibilities of using AI systems such as ChatGPT for a more accurate mapping of the curriculum to ECSF profiles, and therefore, their more accurate design, finer classification using ECSF skills and knowledge and the indication of the percentage representation of skills groups in courses can help. This information can be passed to the ILP module for a more accurate evaluation of the final solution. Furthermore, our model considers all or nothing, whereas we could be interesting on assigning a non-integer value, say 1/3 of the topics of a profile were covered by the course.

User-friendliness and simplicity are key features for end-user acceptance of the tool. Therefore, we plan to give the user additional features, allowing an easy and intuitive way to optimize their curricula. For example, users could select which courses they prefer for curriculum design (course weighting). For instance, the tool could suggest the substitution of curriculum courses by other available courses covering the same skills groups.

7 CONCLUSIONS

In this work, we presented an innovative tool for the AI-driven cybersecurity curriculum development that is in line with the current ECSF released by ENISA. The tool is based on three main parts: 1) mapping methodology allowing practical adaptation of ECSF to existing curricula, 2) machine learning algorithm ensuring categorization/labeling of courses with our skills groups, and 3) integer linear programming algorithm allowing to find the most suitable courses to fulfill the required ECSF profiles and considering additional user requirements. Our skills grouping methodology proves its suitability for the practical adaptation of the ECSF framework. Without the grouping methodology, it would be very difficult if not impossible to map academic or professional courses to ECSF profiles. In the case of the AI-driven approach, our results show that the results of AI are usable.

ACKNOWLEDGMENTS

The work was supported by the ERASMUS+ Sector Skills Alliance (SSA) programme of the European Union (grant 621701-EPP-1-2020-1-LT-EPPKA2-SSA-B 'REWIRE'), by the Ministry of the Interior of the Czech Republic under grant VJ03030003 under program IMPAKT 1, by Fundação para a Ciência e a Tecnologia (FCT) under project UIDB/50008/2020 (Instituto de Telecomunicações), and the project no. TKP2021-NVA-29 supported by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme.

REFERENCES

[1] Stuart Berg, Dominik Kutra, Thorben Kroeger, Christoph N Straehle, Bernhard X Kausler, Carsten Haubold, Martin Schiegg, Janez Ales, Thorsten Beier, Markus Rudy, et al. 2019. Ilastik: interactive machine learning for (bio) image analysis. *Nature Methods* 16, 12 (2019), 1226–1232.

[2] Gianluca Bontempi, Souhaib Ben Taieb, and Yann-Aël Le Borgne. 2012. Machine learning strategies for time series forecasting. In *European business intelligence summer school*. Springer, 62–77.

[3] Petr Dzurenda and Sara Ricci. 2022. R3.4.1 Mapping the framework to existing courses and schemes. https://rewireproject.eu/wp-content/uploads/2023/03/REWIRE_R3.4.1_Deliverable-v8-Final-EC-Check.pdf

[4] ENISA. 2022. European Cybersecurity Skills Framework Role Profiles. https://www.enisa.europa.eu/publications/european-cybersecurity-skills-framework-role-profiles

[5] ENISA. 2023. Communication on the Cybersecurity Skills Academy. https://digital-strategy.ec.europa.eu/en/library/communication-cybersecurity-skills-academy

[6] Cyber2yr2020 Task Group. 2020. *Cybersecurity curricular guidance for associate-degree programs*. Association for Computing Machinery.

[7] Jan Hajny, Sara Ricci, Edmundas Piesarskas, and Marek Sikora. 2021. Cybersecurity Curricula Designer. In *Proceedings of the 16th International Conference on Availability, Reliability and Security* (Vienna, Austria) *(ARES 21)*. Association for Computing Machinery, New York, NY, USA, Article 144, 7 pages. https://doi.org/10.1145/3465481.3469183

[8] Jan Hajny, Marek Sikora, Athanasios Vasileios Grammatopoulos, and Fabio Di Franco. 2022. Adding European Cybersecurity Skills Framework into Curricula Designer. In *Proceedings of the 17th International Conference on Availability, Reliability and Security* (Vienna, Austria) *(ARES '22)*. Association for Computing Machinery, New York, NY, USA, Article 82, 6 pages. https://doi.org/10.1145/3538969.3543799

[9] ICS2. 2023. Cybersecurity Workforce Study. https://www.isc2.org/Research

[10] Jan Jerabek, Edmundas Piesarskas, Imre Lendák, Viktor Varga, Gyorgy Dan, Sara Ricci, Paulius Pakutinskas, and Donatas Alksnys. 2022. R2.2.2 Cybersecurity Skills Needs Analysis. https://rewireproject.eu/wp-content/uploads/2022/04/R2.2.2-Cybersecurity-Skills-Needs-Analysis_FINAL_v1.1.pdf

[11] Jason RC Nurse, Konstantinos Adamos, Athanasios Grammatopoulos, and Fabio Di Franco. 2021. Addressing the eu cybersecurity skills shortage and gap through higher education. *European Union Agency for Cybersecurity (ENISA) Report* (2021).

[12] Brno University of Technology. 2023-2024. curriculum "Information Security". https://www.vut.cz/en/students/programmes/programme/8369

[13] Christos H Papadimitriou and Kenneth Steiglitz. 1998. *Combinatorial optimization: algorithms and complexity*. Courier Corporation.

[14] Rodney Petersen, Danielle Santos, Matthew C. Smith, Karen A. Wetzel, and Greg Witte. 2020. NIST Special Publication 800-181 Revision 1: Workforce Framework for Cybersecurity (NICE Framework). available online at https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-181r1.pdf. https://doi.org/10.6028/NIST.SP.800-181r1

[15] A Pipikaite, G Bueermann, A Joshi, and J Jurgens. 2022. Global Cybersecurity Outlook 2022. In *Geneva: World Economic Forum*.

[16] REWIRE. 2020. REWIRE: Cybersecurity Skills Alliance - A new Vision for Europe. https://rewireproject.eu/

[17] Sara Ricci, Marek Sikora, Simon Parker, Imre Lendak, Yianna Danidou, Argyro Chatzopoulou, Remi Badonnel, and Donatas Alksnys. 2022. Job Adverts Analyzer for Cybersecurity Skills Needs Evaluation. In *Proceedings of the 17th International Conference on Availability, Reliability and Security*. 1–10.

[18] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. 1986. Learning representations by back-propagating errors. *nature* 323, 6088 (1986), 533–536.

[19] Karen Wetzel. 2021. *NICE framework competencies: assessing learners for cybersecurity work*. Technical Report. National Institute of Standards and Technology.

[20] Wenpeng Yin, Katharina Kann, Mo Yu, and Hinrich Schütze. 2017. Comparative study of CNN and RNN for natural language processing. *arXiv preprint arXiv:1702.01923* (2017).